CONTENTS

ABSTRACTS

Emre Can, Mehmet Bozca
Optimising the Geometric Parameters of a Gear in a Tractor Transmission under Constraints Using KISSsoft

We optimise the speed gears in a tractor transmission with KISSsoft software under three constraints: input power torque, transmission system volume and the gear ratio for each speed. This study aimed to optimise the module, face width, gear quality, centre distance, number of teeth, helix angle, addendum modification coefficient and pressure angle for each speed while considering the above constraints based on an optimisation chart. Tooth bending stress, tooth contact stress, contact ratio and specific sliding were considered during optimisation. Additionally, the effects of changes in a module on the gear profiles, overlap ratio, number of teeth and weight of the gear pair were examined. Strength calculations of gear pairs that were optimised and defined for all geometric parameters with KISSsoft were calculated with the mathematical model described in ISO 6336, and results were then compared. Finally, backlash was minimised for all gear pairs as defined with geometric parameters, and all dimensions and tolerances were determined for gear inspection after manufacturing. A concept design was also presented. We conclude that both the KISSsoft results and mathematical model results are within the range of the target value.

Muzafar A. Kalwar
Geological and Geotechnical Assesment of Aggregates used in Nagar Parkar district Tharparkar Sindh Pakistan

This study aims to determine the geology of granite and evaluate the engineering properties of the samples to make recommen-dations for the construction industry. The study area is situated in the Nagar Parker complex in Pakistan, which is located in the extreme south-east of the Thar District and the desert of the Sindh Province, near the Run of Kutch (24° 15′–35 30′ N, 70° 40′–58 07′ E), and it covers ca. 500–1,000 km2. In this region, several Quaternary deposits, subordinate and dispersed Jurassic–Tertiary sandstones and clays are overlying the Nagar Igneous Complex basement. According to international standards, there are various possible aggregate sources. However, only a few of them have been reviewed for suitability reasons. Six quarries in Nagar Parker, Pakistan, were selected for evaluation as coarse aggregate in concrete construction and civil engineering works in this research. Although the aggregates from the six quarries are specified and already widely used in the Sindh Province, there is a lack of studies on their geological properties. The results of the presented research revealed that samples from Dhedvero, Karai, Nagarparkar, Mokrio, Dinsi and Wadlai meet all of the international standard requirements for aggregates. Geotechnical, petrographic and geochemistry laboratory tests were conducted in this research and included bulk density, water absorption, specific gravity test, index of flakiness and elongation, soundness aggregate test, crushing value aggregate, impact value aggregate and abrasion value of Los Angeles. Furthermore, chemical alkali-silica reaction potential test and petrographic examination were tested. As a result, we evaluated the properties of granite, which is a crystalline igneous rock with a visibly crystalline structure and texture, made up of feldspar, i.e., potash feldspar and oligoclase. The evaluated minerals are compatible with the standards of civil engineering works and can be used as a concrete aggregate. The evaluated three types of minerals included Dhedvero simple intrusion, Nagar pink granite and grey granite.

Thomas Koch
Microbiology of Metalworking Fluids: What We Know and Lessons to be Learnt

Water-miscible metalworking fluids are an essential component of many manufacturing processes. During their lifetime they are subject to per-manent changes in their physical and chemical characteristics. Due to their high content of water and their chemical composition in use, metalworking fluids (MWF) are prone to microbial life, i.e. the proliferation of bacteria and fungi. The microbial activity leads to significant changes in the chemical composition of the MWF, which can result in the loss of their technical properties. This paper briefly discusses the influences of microbial contamination on the technical quality of MWF and presents common monitoring systems for the detection of microorganisms. Finally, measures are described that can be taken to protect MWF from damage caused by high microbial loads in daily practice. In a short outlook, alternative research approaches are mentioned that aim at sustainable use of MWF.

Leander Marquardt, Heiner-Joachim Katke, Andreas Reinke, Niklas Kockskämper
Influence of Valve-Seat Angles to Operation Values and Emissions of Medium-Speed Diesel Engines

For the development of gas exchange for large diesel engines, a compromise has to be found between efficient valve-flow and the time between overhauls. On the one hand, large effective flow areas, especially during valve-overlap, are demanded. On the other hand, there are limitations of cylinder bore regarding the maximum diameter of inlet and outlet valves and the minimum distance (dead space) between valves and piston, as well as wear-related smaller seat angles. For large medium-speed diesel engines, a valve-seat angle of $\beta = 30°$ for inlet and outlet valves is a standard application. For engine-operation with clean fuels, a valve-seat lubrication (gasoil) or smaller seat angles (natural gas) need to be applied. With this presentation, the basic influence of different valve-seat angles on the operation values and emissions will be considered for the example of the single-cylinder research engine FM16/24. Using a self-developed testbed, experimental investigations into effective flow areas as a function of valve-lift at inlet and outlet valves have to be executed. With this input, different cycle calculations including T/C have to be carried out to determine deviances in specific fuel-oil consumption, exhaust-gas temperatures, NOx emissions and air/fuel ratio. The results will be discussed critically.

Krzysztof Oprzędkiewicz, Wojciech Mitkowski, Maciej Rosół
Fractional Order, State Space Model of the Temperature Field in the PCB Plate

In the paper the fractional order, state space model of a temperature field in a two-dimensional metallic surface is addressed. The proposed model is the two dimensional generalization of the one dimensional, fractional order, state space of model of the heat transfer process. It uses fractional derivatives along time and length. The proposed model assures better accuracy with lower order than models using integer order derivatives. Elementary properties of the proposed model are analysed. Theoretical results are experimentally verifed using data from industrial thermal camera.

Tadeusz Kaczorek, Andrzej Ruszewski
Standard and Fractional Discrete-Time Linear Systems with Zero Transfer Matrices

The transfer matrix of the standard and fractional linear discrete-time linear systems is investigated. Necessary and sufficient conditions for zeroing of the transfer matrix of the linear discrete-time systems are established. The considerations are illustrated by examples of the standard and fractional linear discrete-time systems.

Khalissa Saada, Salah Amroune, Moussa Zaoui, Amin Houari, Kouider Madani, Amina Hachaichi
Experimental and Numerical Study of the Effect of the Presence of a Geometric Discontinuity of Variable Shape on the Tensile Strength of an Epoxy Polymer

The presence of geometric discontinuity in a material reduces considerably its resistance to mechanical stresses, therefore reducing the service life of materials. The analysis of structural behaviour in the presence of geometric discontinuities is important to ensure the proper use, especially if it is regarding a material of weak mechanical properties such as a polymer. The objective of the present work is to analyse the effect of the notch presence of variable geometric shapes on the tensile strength of epoxy-type polymer specimens. A series of tensile tests were carried out on standardised specimens, taking into account the presence or absence of a notch. Each series of tests contains five specimens. Two notch shapes were considered: circular (hole) and elliptical. The experimental results in terms of stress–strain clearly show that the presence of notches reduces considerably the resistance of the material, where the maximum stress for the undamaged specimen was 41.22 MPa and the lowest stress for the elliptical-notched specimen was 11.21 MPa. A numerical analysis by the extended finite element method (XFEM) was undertaken on the same geometric models; in addition, the results in stress–strain form were validated with the experimental results. A remarkable improvement was obtained (generally an error within 0.06%) for strain, maximum stress, Young's modulus and elongation values. An exponential decrease was noted in the stress, strain, and Young's modulus in the presence of a notch in the material.

Krzysztof Magnucki, Joanna Kustosz, Damian Goliwąs
Effective Shaping of a Stepped Sandwich Beam with Clamped Ends

The aim of this work is to propose a sandwich beam with stepped layer thickness in three parts along its length. The total depth, width of the cross-section and its mass are constant. The beam is under a uniformly distributed load. The system of two equilibrium equations was formulated for each part based on the literature. This system was analytically solved for the successive parts of the beam and the functions of the shear effect and deflection were determined in them. The effective stepped layer thicknesses was determined on the basis of the adopted criterion for minimizing the maximum deflection of the beam. The example calculations were made for two elected beams. The effective shapes of these beams are shown in the figures. Moreover, FEM numerical calculations of the deflections of these beams are performed.

Zbyszko Klockiewicz, Grzegorz Ślaski
The Influence of Friction Force and Hysteresis on the Dynamic Responses of Passive Quarter-Car Suspension
with Linear and Non-Linear Damper Static Characteristics

Vehicle passive suspensions consist of two major elements generating force – spring and passive damper. Both possess non-linear characteristics, which are quite often taken into account in simulations; however, the friction forces inside the hydraulic damper and the damping force's hysteresis are usually left out. The researchers in this paper present the results of examination of the influence of using complex damper models – with friction and hysteresis; and with linear and non-linear static characteristics – on the chosen dynamic responses of a suspension system for excitations in the typical exploitation frequency range. The results from the simulation tests of the simplified and advanced versions of the damper model – different transfer functions and their relation to the reference model's transfer functions – are compared. The main conclusion is that friction and hysteresis add extra force to the already existing damping force, acting similar to damping increase for the base static characteristics. But this increase is not linear – it is bigger for smaller frequencies than for higher frequencies. The research shows the importance of including non-linear characteristics and proposed modules in modelling passive dampers.

Michal Macias, Dominik Sierociuk
Finite Length Triple Estimation Algorithm and its Application to Gyroscope MEMS Noise Identification

The noises associated with MEMS measurements can significantly impact their accuracy. The noises characterised by random walk and bias instability errors strictly depend on temperature effects that are difficult to specify during direct measurements. Therefore, the paper aims to estimate the fractional noise dynamics of the stationary MEMS gyroscope based on finite length triple estimation algorithm (FLTEA). The paper deals with the state, order and parameter estimation of fractional order noises originating from the MEMS gyroscope, being part of the popular Inertial Measurement Unit denoted as SparkFun MPU9250. The noise measurements from x, y and z gyroscope axes are identified using a modified triple estimation algorithm (TEA) with finite approximation length. The TEA allows a simultaneous estimation of the state, order and parameter of fractional order systems. Moreover, as it is well-known that the number of samples in fractional difference approximations plays a key role, we try to show the influence of applying the TEA with various approximation length constraints on final estimation results. The validation of finite length TEA in the noise estimation process coming from MEMS gyroscope has been conducted for implementation length reduction achieving 50% of samples needed to estimate the noise with no implementation losses. Additionally, the capabilities of modified TEA in the analysis of fractional constant and variable order systems are confirmed in several numerical examples.

Mateusz Pietrala, Piotr Leśniewski, Andrzej Bartoszewicz
An ITAE Optimal Sliding Mode Controller for Systems with Control Signal and Velocity Limitations

In this paper, a sliding mode controller, which can be applied for second-order systems, is designed. Robustness to external dis-turbances, finite regulation time and a good system's behaviour are required for a sliding mode controller. In order to achieve the first two of these three goals, a non-linear, time-varying switching curve is introduced. The representative point (state vector) belongs to this line from the very beginning of the control process, which results in elimination of the reaching phase. The stable sliding motion along the switching curve is provided. Natural limitations such as control signal and system's velocity constraints will be taken into account. In order to satisfy them, the sliding line parameters will be properly selected. However, a good dynamical behaviour of the system has to be provided. In order to achieve that, the integral time absolute error (ITAE) quality index will be introduced and minimised. The simulation example will verify theoretical considerations.

Jamil A. Haider, Sana Gul, Jamshaid U. Rahman, Fiazud D. Zaman
Travelling Wave Solutions of the Non-Linear Wave Equations

This article focuses on the exact periodic solutions of nonlinear wave equations using the well-known Jacobi elliptic function expansion method. This method is more general than the hyperbolic tangent function expansion method. The periodic solutions are found using this method which contains both solitary wave and shock wave solutions. In this paper, the new results are computed using the closed-form solution including solitary or shock wave solutions which are obtained using Jacobi elliptic function method. The corresponding solitary or shock wave solutions are compared with the actual results. The results are visualised and the periodic behaviour of the solution is described in detail. The shock waves are found to break with time, whereas, solitary waves are found to be improved continuously with time.

Nidhish K. Mishra
Computational Analysis of Soret and Dufour Effects on Nanofluid Flow Through a Stenosed Artery in the Presence of Temperature-Dependent Viscosity

In this study, the Soret and Dufour effects in a composite stenosed artery were combined with an analysis of the effect of varying viscosity on copper nanofluids in a porous medium. Blood viscosity, which changes with temperature, is taken into account using the Reynolds viscosity model. The finite difference approach is used to quantitatively solve the governing equations. For use in medical applications, the effects of the physical parameters on velocity, temperature and concentration along the radial axis have been investigated and physically interpreted. The results are graphically displayed and physically defined in order to facilitate comprehension of the various phenomena that occur in the artery when nanofluid is present. It is observed that the Soret effect increases the rate of heat transfer but decreases the rate of mass transfer. The new study enhances knowledge of non-surgical treatment options for stenosis and other abnormalities, hence reducing post-operative complications. Additionally, current research may have biomedical applications such as magnetic resonance angiography (MRA), which provide a picture of an artery and enable identification of any anomalies, and thus may be useful

Adam Szcześniak, Zbigniew Szcześniak, Leszek Cedro
Synthesis of Pneumatic Systems in the Control of the Transport Line of Rolling Elements

This paper presents the synthesis of a pneumatic control system for a selected configuration of the transport path for the delivery of rolling elements to spiral storage in inter-operational transport. The sequential control system sets the state of the manifolds to ensure a flow of workpieces to serve the subsequent storage. The essential module of the control system is the memory block. It is developed based on a storage filling sequence graph. The filling level of the storages can be monitored in one or two points using sensors. The rolling element displacement control sensors work together with appropriately designed systems to execute the delay of the rising and falling edge input signal. By using a two-level control of the filling level of the storages, it is possible to control the emptying status of the storages as a function of the technological time of removal of the items from the storage between the two control points. Control systems were synthesised and verified using Festo's FluidSim computer programme.

Aastha Aastha, Khem Chand
Soret and Dufour Effects on Chemically Reacting and Viscous Dissipating Nanofluid Flowing Past a Moving Porous Plate in the Presence of a Heat Source/Sink

This study performed a numerical investigation of the Soret and Dufour effects on unsteady free convective chemically reacting nanofluid flowing past a vertically moving porous plate in the presence of viscous dissipation and a heat source/sink. The equations directing the flow are non-dimensionalised, modified to ordinary differential equations and emerging equations are resolved computationally by using the bvp4c function in MATLAB software. The results obtained from this analysis indicate that the resulting velocity of the nanofluid increases with increasing Grashof number, mass Grashof number and porosity parameter. An increase in the Dufour number increases the fluid temperature, whereas the concentration profile declines with the increase in the Schmidt number. It is also observed that the skin friction coefficient, Nusselt number and Sherwood number increase with increasing magnetic field parameter, Eckert number and Schmidt number, respectively. The present study reveals the impact of Soret and Dufour effects on heat and mass transfer rates in chemically reacting and viscous dissipating nanofluids.

Roman Król, Kazimierz Król
Multibody Dynamics Model of the Cycloidal Gearbox, Implemented in Fortran for Analysis of Dynamic Parameters Influenced by the Backlash as a Design Tolerance

In this study, dynamical parameters of the cycloidal gearbox working at the constant angular velocity of the input shaft were investigated in the multibody dynamics 2D model implemented in the Fortran programming language. Time courses of input and output torques and forces acting on the internal and external sleeves have been shown as a function of the contact modelling parameters and backlash. The analysis results in the model implemented in Fortran were compared with the results in the 3D model designed using MSC Adams software. The values of contact forces are similar in both models. However, in the time courses obtained in MSC Adams there are numerical singularities in the form of peaks reaching 500 N for the forces at external sleeves and 400 N for the forces acting at internal sleeves, whereas in the Fortran model, there are fewer singularities and the maximum values of contact forces at internal and external sleeves do not exceed 200 N. The contact damping and discretisation level (the number of discrete contact points on the cycloidal wheels) significantly affect the accuracy of the results. The accuracy of computations improves when contact damping and discretisation are high. The disadvantage of the high discretization is the extended analysis time. High backlash values lead to a rise in contact forces and a decrease in the force acting time. The model implemented in Fortran gives a fast solution and performs well in the gearbox optimisation process. A reduction of cycloidal wheel discretisation to 600 points, which still allows satisfactory analysis, could reduce the solution time to 4 min, corresponding to an analysis time of 0.6 s with an angular velocity of the input shaft of 52.34 rad/s (500 RPM).

Zbigniew Kneba, Jacek Kropiwnicki, Jakub Hadrzyński, Maciej Ziółkowski
Forecasting Biogas Formation in Landfills

The aim of the present research was to develop a mathematical model for estimating the amount of viscous gas generated as a function of weather conditions. Due to the lack of models for predicting gas formation caused by sudden changes in weather conditions in the literature, such a model was developed in this study using the parameters of landfills recorded for over a year. The effect of temperature on landfill gas production has proved to be of particular interest. We constructed an algorithm for calculating the amount of the produced gas. The model developed in this study could improve the power control of the landfill power plant.

Michaela Zeißig, Frank Jablonski
Numerical Investigation of Production-Related Characteristics Regarding their Influence on the Fatigue Strength of Additively Manufactured Components

In order to further enhance the application of additive manufacturing (AM) processes, such as the laser powder bed fusion (L-PBF) process, reliable material data are required. However, the resulting specimen properties are significantly influenced by the process parameters and may also vary depending on the material used. Therefore, the prediction of the final properties is difficult. In the following, the effect of residual stresses on the fatigue strength of 316L steel, a commonly used steel in AM, is investigated using a Weibull distribution. The underlying residual stress distributions as a result of the building process are approximated for two building directions using finite element (FE) models. These imply significantly different distributions of tensile and compressive residual stresses within the component. Apart from the residual stresses, the impact of the mean stress sensitivity is discussed as this also influences the predicted fatigue strength values.

Lukas Peters, Rüdiger Kutzner, Marc Schäfer, Lutz Hofmann
Ability of Black-Box Optimisation to Efficiently Perform Simulation Studies in Power Engineering

In this study, the potential of the so-called black-box optimisation (BBO) to increase the efficiency of simulation studies in power engineering is evaluated. Three algorithms (“Multilevel Coordinate Search” (MCS) and “Stable Noisy Optimization by Branch and Fit” (SNOBFIT) by Huyer and Neumaier and “blackbox: A Procedure for Parallel Optimization of Expensive Black-box Functions” (blackbox) by Knysh and Korkolis) are implemented in MATLAB and compared for solving two use cases: the analysis of the maximum rotational speed of a gas turbine after a load rejection and the identification of transfer function parameters by measurements. The first use case has a high computational cost, whereas the second use case is computationally cheap. For each run of the algorithms, the accuracy of the found solution and the number of simulations or function evaluations needed to determine the optimum and the overall runtime are used to identify the potential of the algorithms in comparison to currently used methods. All methods provide solutions for potential optima that are at least 99.8% accurate compared to the reference methods. The number of evaluations of the objective functions differs significantly but cannot be directly compared as only the SNOBFIT algorithm does stop when the found solution does not improve further, whereas the other algorithms use a predefined number of function evaluations. Therefore, SNOBFIT has the shortest runtime for both examples. For computationally expensive simulations, it is shown that parallelisation of the function evaluations (SNOBFIT and blackbox) and quantisation of the input variables (SNOBFIT) are essential for the algorithmic performance. For the gas turbine overspeed analysis, only SNOBFIT can compete with the reference procedure concerning the runtime. Further studies will have to investigate whether the quantisation of input variables can be applied to other algorithms and whether the BBO algorithms can outperform the reference methods for problems with a higher dimensionality.

Jacek Kropiwnicki, Tomasz Gawłas
Estimation of the Regenerative Braking Process Efficiency in Electric Vehicles

In electric and hybrid vehicles, it is possible to recover energy from the braking process and reuse it to drive the vehicle using the batteries installed on-board. In the conditions of city traffic, the energy dissipated in the braking process constitutes a very large share of the total resistance to vehicle motion. Efficient use of the energy from the braking process enables a significant reduction of fuel and electricity consumption for hybrid and electric vehicles, respectively. This document presents an original method used to estimate the efficiency of the regenerative braking process for real traffic conditions. In the method, the potential amount of energy available in the braking process was determined on the basis of recorded real traffic conditions of the analysed vehicle. The balance of energy entering and leaving the battery was determined using the on-board electric energy flow recorder. Based on the adopted model of the drive system, the efficiency of the regenerative braking process was determined. The paper presents the results of road tests of three electric vehicles, operated in the same traffic conditions, for whom the regenerative braking efficiency was determined in accordance with the proposed model. During the identification of the operating conditions of the vehicles, a global positioning system (GPS) measuring system supported by the original method of phenomenological signal correction was used to reduce the error of the measured vehicle's altitude. In the paper, the efficiency of the recuperation process was defined as the ratio of the accumulated energy to the energy available from the braking process and determined for the registered route of the tested vehicle. The obtained results allowed to determine the efficiency of the recuperation process for real traffic conditions. They show that the recuperation system efficiency achieves relatively low values for vehicle No. 1, just 21%, while the highest value was achieved for vehicle No. 3, 77%. Distribution of the results can be directly related to the power of electric motors and battery capacities of the analysed vehicles.

DOI 10.2478/ama-2023-0016

OPTIMISING THE GEOMETRIC PARAMETERS OF A GEAR IN A TRACTOR TRANSMISSION UNDER CONSTRAINTS USING KISSSOFT

Emre CAN*, Mehmet BOZCA**

*Graduate School of Science and Engineering, Yildiz Technical University, Yildiz, 34349 Istanbul, Turkey
**Mechanical Engineering Faculty, Machine Design Division, Yildiz Technical University, Yildiz, 34349 Istanbul, Turkey

canemrecan@hotmail.com, mbozca@yildiz.edu.tr

received 15 July 2022, revised 29 October 2022, accepted 13 November 2022

Abstract: We optimise the speed gears in a tractor transmission with KISSsoft software under three constraints: input power torque, transmission system volume and the gear ratio for each speed. This study aimed to optimise the module, face width, gear quality, centre distance, number of teeth, helix angle, addendum modification coefficient and pressure angle for each speed while considering the above constraints based on an optimisation chart. Tooth bending stress, tooth contact stress, contact ratio and specific sliding were considered during optimisation. Additionally, the effects of changes in a module on the gear profiles, overlap ratio, number of teeth and weight of the gear pair were examined. Strength calculations of gear pairs that were optimised and defined for all geometric parameters with KISSsoft were calculated with the mathematical model described in ISO 6336, and results were then compared. Finally, backlash was minimised for all gear pairs as defined with geometric parameters, and all dimensions and tolerances were determined for gear inspection after manufacturing. A concept design was also presented. We conclude that both the KISSsoft results and mathematical model results are within the range of the target value.

Key words: gears, geometric parameters, optimisation, KISSsoft

1. INTRODUCTION

Gears are typically used in many mechanical systems, particularly automotive systems, and can be designed to be more reliable, lighter and quieter via optimisation studies. Additionally, gears can be more cost-competitive during optimisation. Thus, many studies have investigated the optimisation of gears' geometric parameters.

Within the scope of the study, the results obtained with the mathematical model are compared with the results found with KISSsoft. In this study, both the mathematical model and KISSsoft analyses are carried out. This is the main difference that distinguishes this study from the existing studies in the literature.

The design and contact stress of helical gears in lightweight cars have been analysed in various studies. High stresses that cause pitting decrease as gear width increases. Contact failure in gears can be predicted by calculating contact stress [1].

The combined effects of the gear ratio, helix angle, face width and module on the bending and compressive stress of steel alloy helical gears have also been investigated. Increasing the module, face width and helix angle results in decreased tooth-root stress [2].

The effect of gear design variables on the dynamic stress of multistage gears has also been analysed. Increasing the module results in higher dynamic stress, increasing the pressure angle markedly increases stress levels, and increasing the contact ratio increases bending stress [3].

The effects of the module and pressure angle on contact stresses in spur gears have also been investigated. Studies have shown a decrease in contact fatigue life with an increase in the module and pressure angle [4].

The effect of backlash on bending stresses in spur gears has also been investigated. Both stresses and deformations increase as a result of increased backlash [5].

The effects of gear parameters on the surface durability of gear flanks have also been investigated. The optimum parameters of cylindrical gear pairs are determined in terms of specific sliding and the contact stresses on the flanks [6].

The effects of sliding speed and specific sliding of the interval meshing gears have also been analysed. Increasing the profile shift coefficient decreases specific sliding but also decreases the contact ratio [7].

The effects of profile shifts in helical gear mechanisms with analytical and numerical methods have also been investigated. Both tooth-root stress and tooth contact stress decrease with a positive profile shift coefficient [8].

The optimisation of addendum modification for the bending strength of involute spur gears has also been studied. Increasing both the profile shift coefficient and pressure angle decreases tooth-root stress [9].

Finite element analysis of the contact stress and bending stress in the helical gear pair has also been performed. Increasing the helix angle increases both tooth-root stress and tooth contact stress [10].

The optimisation of the geometric parameters of gears under variable loading conditions has also been performed and show that tooth-root stresses increase due to negative profile shifts and decrease by positive profile shifts [11].

Effective design parameters have been optimised for an automotive transmission gearbox to reduce tooth bending stress. Both tooth-root stress and tooth contact stress decrease with decreasing contact ratio via an increasing pressure angle [12].

Empirical model-based optimisation of gearbox geometric design parameters to reduce rattle noise in an automotive transmission was presented and showed that by optimising the geometric parameters of the gearbox, it is possible to obtain a lightweight gearbox structure and minimise rattling noise [13].

Torsional vibration model-based optimisation of gearbox geometric design parameters to reduce rattle noise in an automotive transmission was studied and showed that by optimising the geometric parameters of the gearbox, it is possible to obtain a lightweight gearbox structure and minimise rattling noise [14].

Transmission error model-based optimisation of the geometric design parameters of an automotive transmission gearbox to reduce gear-rattle noise showed that by optimising the geometric parameters of the gearbox, it is possible to obtain a gear structure with high bending and contact strengths, and to minimise the torsional vibration, transmission error and gear-rattling noise [15].

The remainder of the present study is structured as follows: Section 2 discusses the calculation of the load-carrying capacity of the helical gear; Section 3 the optimisation concept; Section 4 the optimisation steps with KISSsoft; Section 5 provides the results and discussion; and Section 6 presents the conclusions.

2. CALCULATING THE LOAD CAPACITY OF HELICAL GEARS

Gears encounter tooth bending stress and tooth contact stress during power-torque transfers. Therefore, damage can occur on gears due to stresses on the gears, which must be considered during design.

2.1. Tooth bending stress

The tangential force creates a bending stress on the tooth, and the radial force creates a compressive stress on the tooth. These forces cause stress concentrations at the root of the tooth, and the stress concentration must be considered so that the load capacity of the gear can be calculated. The distribution of the forces on the gears is shown in Fig. 1, and the tooth bending stress according to the ISO 6336 standard is calculated as follows [16]:

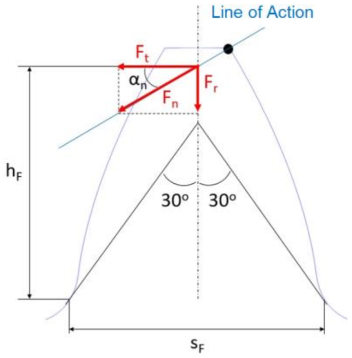


Fig. 1. Tooth bending stress

The real tooth-root stress, σ_F, is calculated as follows:

$$\sigma_F = \frac{F_t}{bm_n} Y_F Y_S Y_\varepsilon Y_\beta K_A K_V K_{F\beta} K_{F\alpha} \tag{1}$$

where F_t is the nominal tangential load (N), b is the face width (mm), m_n is the normal module (mm), Y_F is the form factor (-), Y_S is the stress correction factor (-), Y_ε is the contact ratio factor (-), Y_β is the helix angle factor (-), K_A is the application factor (-), K_V is the dynamic factor (-), $K_{F\beta}$ is the face load factor (-) and $K_{F\alpha}$ is the transverse load factor (-).

The permissible bending stress, σ_{FP}, is calculated as follows:

$$\sigma_{FP} = \sigma_{F\,lim.} Y_{ST} Y_N Y_\delta Y_R Y_X \tag{2}$$

where σ_{FLim} is the nominal stress (N/mm^2), Y_{ST} is the stress correction factor (–), Y_N is the life factor (–), Y_δ is the relative notch sensitivity factor (–), Y_R is the relative surface factor (–) and Y_X is the size factor (–).

The safety factor for bending stress S_F is calculated as follows:

$$S_F = \frac{\sigma_{FP}}{\sigma_F} \tag{3}$$

2.2. Tooth contact stress

The force that affects the surfaces in contact with each other in the gear pair creates a high surface pressure called Hertzian contact stress shown in Fig.2 due to the effect on a small area of the surface during power transmission. These stresses cause wear and pitting depending on material fatigue. The surface pressure that occurs on gears is calculated according to the ISO 6336 standard as follows [17]:

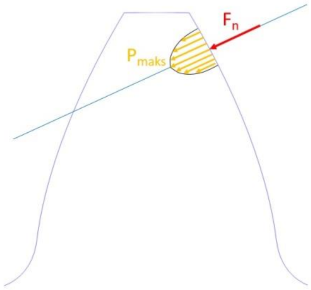


Fig. 2. Tooth contact stress

The real contact stress, σ_H, is calculated as follows:

$$\sigma_H = \sqrt{\frac{F_t(u+1)}{bm_n u}} Z_H Z_E Z_\varepsilon Z_\beta \sqrt{K_A K_V K_{H\beta} K_{H\alpha}} \tag{4}$$

where u is the gear ratio (–), Z_H is the zone factor (–), Z_E is the elasticity factor, Z_ε is the contact ratio factor (–), Z_β is the helix angle factor (–), $K_{H\beta}$ is the face load factor and $K_{H\alpha}$ is the transverse load factor (–).

The permissible contact stress, σ_{HP}, is calculated as follows:

$$\sigma_{Hp} = \sigma_{H\,lim.} Z_N Z_L Z_V Z_R Z_W Z_X \tag{5}$$

where σ_{Hlim} is the allowable stress (N/mm^2), Z_N is the life factor (–), Z_L is the lubrication factor (–), Z_V is the velocity factor (–), Z_R is the roughness factor (–), Z_W is the work hardening factor (–) and Z_X is the size factor (–).

The safety factor for contact stress, SH, is calculated as follows:

$$S_H = \frac{\sigma_{Hp}}{\sigma_H} \tag{6}$$

DOI 10.2478/ama-2023-0016

3. OPTIMISATION

Gears are currently used in many mechanical systems. Both safe and cheaper gear systems can be designed via optimisation, and efficient systems can be created to determine requisite geometric parameters. During optimisation, all geometric parameters of the gear pairs are determined step by step, as indicated in the flow chart below:

1. Face width, gear quality and first-level module optimisation
2. Outputs: the module graphic, S_F, S_H; changes in the gear profiles; the overlap ratio; the teeth number; the gear ratio; and the system weight
3. Centre-distance optimization
4. Outputs: the tip-diameter graphic; the centre distance; and the S_F, S_H, centre-distance graphic
5. Last-level module optimisation
6. Number of teeth optimisation
7. Outputs: the specific-sliding graphic, the number of teeth, and the contact ratio
8. Helix angle optimisation
9. Outputs: axial forces and contact ratio
10. Addendum modification coefficient optimisation
11. Outputs: the angle-of-rotation graphic, the specific sliding and changes in the gear profiles
12. Pressure-angle optimisation
13. Comparison of the results for both the KISSsoft and mathematical models
14. Outputs: tooth bending stress, tooth contact stress and safety factors
15. Backlash optimisation
16. Outputs: dimensions and tolerances for inspection of all gears after manufacturing
17. Concept design
18. Outputs: conceptual 3D design

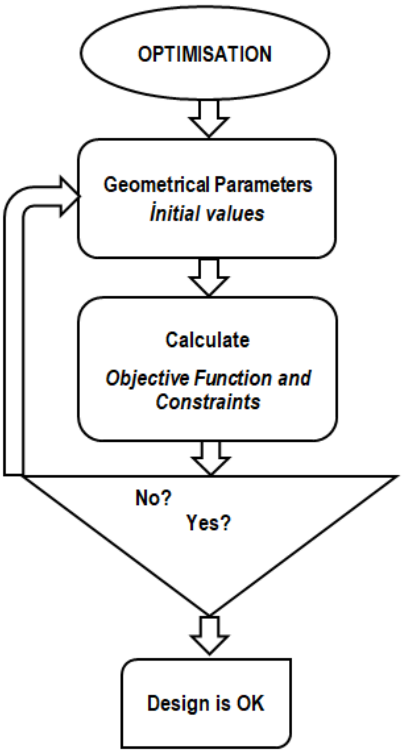


Fig. 3. Flow chart of optimisation

4. OPTIMISATION WITH KISSSOFT

Four speed gears of a tractor transmission shown in Fig.4 were optimised via KISSsoft software in this study. The input power was 50 kW, and the torque was 238 Nm for the speed gears that were optimised in this study. These four speed gears have ratios that are similar to those in Tab. 1 with tolerances of 4%.

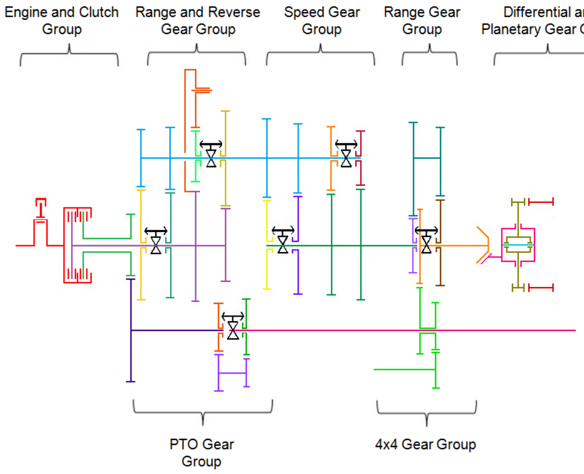


Fig. 4. Gear scheme of tractor transmission

Tab. 1. Ratios

Speed	Ratio (4%)
1	3.1
2	1.9
3	1.1
4	0.7

The volume that can be used for speed gears is limited due to other systems on both tractors and transmission. The maximum volume that can be used in transmission for these speed gear groups is shown in Fig. 5.

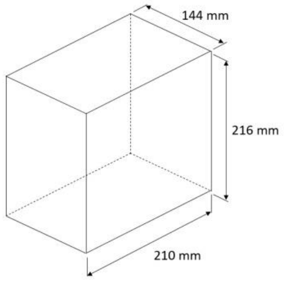


Fig. 5. Maximum volume

4.1. Gear width, gear quality and first-level module optimisation

First, suitable geometric parameters of gear width, gear quality and module were calculated while considering constraints with KISSsoft. Other geometric parameters were assumed to be constant, and these other geometric parameters were optimised in the subsequent steps to design an efficient system.

The vertical dimension was considered 186 mm (216– 30 mm) during optimisation because all gear pairs that assemble at the same centre distance can use space approximately 15 mm along the vertical dimension based on the number of teeth and module of gear pairs. The centre distance was considered 93 mm (186 mm/2 mm) during the first optimisation step so that two shafts could be placed at 186 mm. Volume measures are shown in Fig.6.

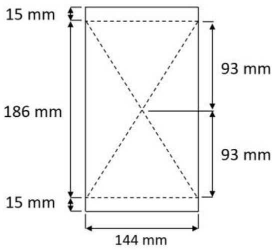


Fig. 6. Centre distance

The pressure angle was considered to be 20°, which is used as a standard for many manufacturers. The helix angle was considered to be 17°, which is the average value for similar tractor transmission. The material was defined as 16 MnCr5, which is used in similar systems. These parameters were accepted as specified initially but were later optimised.

First, optimisation was begun with the same gear width for four gear pairs, and then each gear width was defined according to the results. The horizontal dimension was 210 mm, as shown in Fig. 5 and Fig.7. Four gears (b_1, b_2, b_3 and b_4) and two synchromeshes (s_1 and s_3) must be placed 210 mm away from each other. Each synchromesh requires 50 mm in the horizontal direction, and there should be 5 mm spaces (s_2) between the second and third gears for safety due to production and assembly errors. The initial gears' widths are shown in Tab. 2.

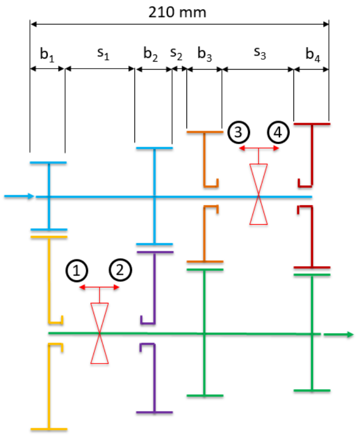


Fig. 7. Four speed gear group

Tab. 2. Gear widths

s_1 (mm)	s_2 (mm)	s_3 (mm)	b_1 (mm)	b_2 (mm)	b_3 (mm)	b_4 (mm)
50	5	50	26.25	26.25	26.25	26.25

Gear pairs that can be used for the first speed gear were calculated with KISSsoft considering 50 kW power, 238 Nm torque, 3.1 gear ratio, 20° pressure angle, 17° helix angle, 93 mm centre distance, 7 quality, 26.25 mm gear width, and a module between 1 mm and 5 mm. According to the results, 386 different solutions were found for the first speed gear pair, and all results are shown in Fig. 8 to evaluate the parameters. In the figure, the horizontal axis represents the module, the vertical axis represents the minimum root safety and the colour scale represents the minimum flank safety. Increasing the module results appears to increase the root safety. For the 1 mm module, root safety was calculated to be approximately 0.6, for the 4.8 mm module, it was approximately 2.1. Considering the colour scale, the colours change from red to blue with increasing modules; thus, increasing the module results in decreased flank safety.

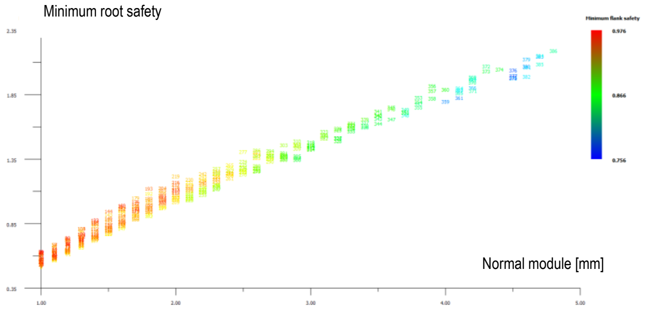


Fig. 8. Optimisation of the first speed gear pairs (b = 26.25 mm)

For this system, the root safety is between 1.3 and 1.5. There are some optimisation steps for this study; thus, root safeties that are between 1.1 and 1.9 are sufficient for this first optimisation considering that there can be changes in other steps. Additionally, the flank safeties must be at least 1, and average value of flank safety have to be 1.1 in the figure. According to Fig. 8, all flank safeties are <1; thus, these gear pairs cannot be used for this system. Gear width and quality were changed so that flank safeties could be increased. First, the gear width was increased to 35 mm and 40 mm, and then the gear quality was increased from 7 to 6. The gear width was not increased by >40 mm because the horizontal dimension was limited, and space is needed for other gear pairs. The results based on these new parameters are shown in Figs. 9–11 and Tab. 3.

Tab. 3. Optimisation of the first speed gear pairs with new parameters

Gear width (mm)	Gear quality	S_{Hmin}
26.25	7	0.756
35	7	0.919
40	7	0.982
40	6	1.018

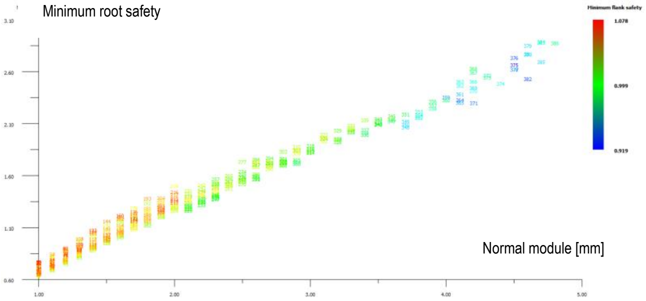


Fig. 9. Optimisation of the first speed gear pairs with new parameters (b = 35 mm)

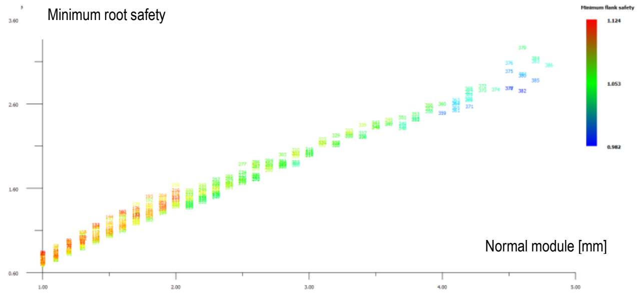


Fig. 10. Optimisation of the first speed gear pairs with new parameters (b = 40 mm, 7 quality)

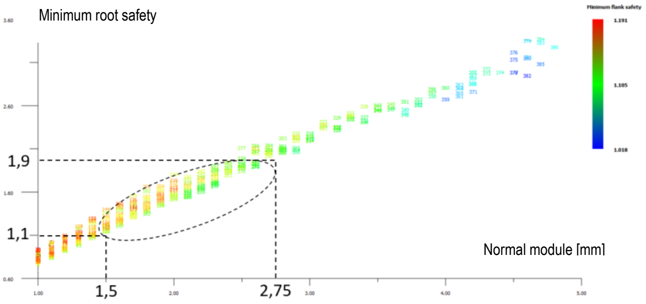


Fig. 11. Optimisation of the first speed gear pairs with new parameters (b = 40 mm, 6 quality)

According to the results, a 40 mm gear width and a quality of 6 yield the optimal first speed gear pair. Additionally, the module should be between 1.5 mm and 2.75 mm wide.

After optimising the first speed gear pairs, the second speed gear pairs were optimised. The input conditions were the same as in the first gear optimisation except for the ratio, which was set to 1.9. According to the results, 587 different solutions were found, and the results are shown in Tab. 4 and Fig. 12. According to the results, a 25 mm gear width and a quality of 7 yielded the optimal second speed gear pair. Additionally, the module should be between 1.5 mm and 2.75 mm wide.

Tab. 4. Optimisation of the second speed gear pairs with new parameters

Gear width (mm)	Gear quality	S_{Hmin}
30	7	1.130
25	7	1.016

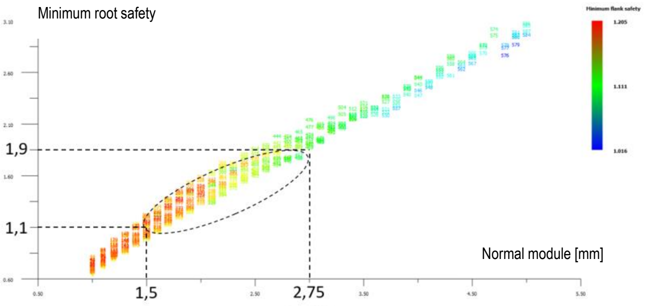


Fig. 12. Optimisation of the second speed gear pairs (b = 25 mm)

After optimisation of the first and second speed gear pairs, a 40 mm horizontal space remained as the horizontal limit of the system. Therefore, the gear width was set equal to 20 mm for both the third and fourth speed gear pairs.

Optimisation was performed for the third speed gear pair, and 774 different solutions were found. As shown in Tab. 5 and Fig. 13, a 20 mm gear width and a quality of 8 yielded the optimal third speed gear pair. Additionally, the module should be between 1.5 mm and 3 mm.

Finally, the fourth speed gear pair was optimised, and 716 different solutions were found. As shown in Tab. 6 and Fig. 14, a 20 mm gear width and a quality of 8 yielded the optimal fourth speed gear pair. Additionally, the module should be between 1.5 mm and 2.75 mm.

Tab. 5. Optimisation of the third speed gear pairs with new parameters

Gear width (mm)	Gear quality	S_{Hmin}
20	7	1.063
20	8	1.009

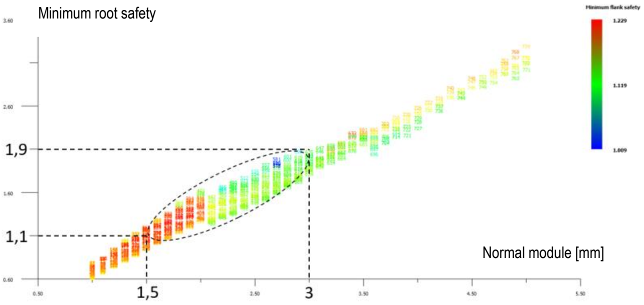


Fig. 13. Optimisation of the third speed gear pairs (8 quality)

Tab. 6. Optimisation of the fourth speed gear pairs with new parameters

Gear width (mm)	Gear quality	S_{Hmin}
20	7	1.186
20	8	1.111

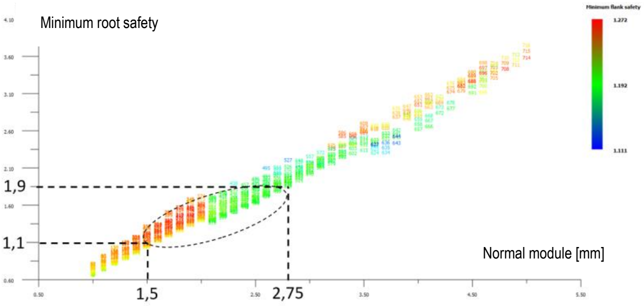


Fig. 14. Optimisation of the fourth speed gear pairs (8 quality)

4.2. Centre-distance optimisation

The centre distance was set equal to 93 mm, which is the centre of the useable vertical dimension of considering the gear width, quality and first-level module optimisation. In this step, the centre distance was optimised between reasonable values of 86 mm and 98 mm.

For the first speed gear pair, the optimisation was performed with the following assumptions: 50 kW power, 238 Nm torque, 3.1 gear ratio, 20° pressure angle, 17° helix angle, quality of 6, 40 gear width, and modules between 1.5 mm and 2.75 mm with standard measurements in steps of 0.25 mm. Based on these values, 1,006 different solutions were found. The tip diameters of gears are important to optimise the centre distance due to volume constraints. In Fig. 15, the horizontal axis represents the tip diameters of the driver gear, the vertical axis represents the tip diameters of the driven gear and the colour scale represents the centre distances.

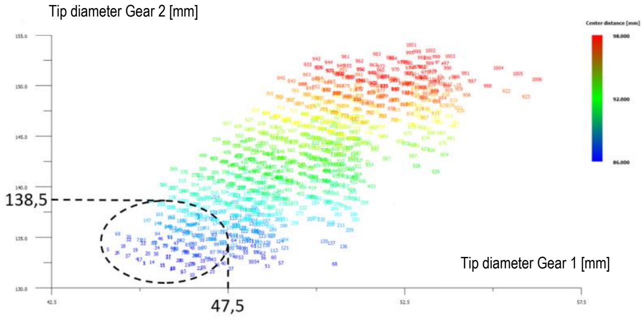


Fig. 15. Centre distance for the first speed gear pair

According to input constraints, the horizontal limit is 144 mm, and the vertical limit is 216 mm; thus, 186 mm is used due to assembly conditions. Additionally, 2 mm of space for each side of the gears was included; thus, the horizontal limit was considered to be 140 mm. Thus, the tip diameter of each gear should not be >140 mm. Additionally, the sum of the tip diameters of the gear pair should not be >186 mm. According to Fig. 15, the tip diameter of the driver gear can reach 47.5 mm, and the tip diameter of the driven gear can reach 138.5 mm when considering the two above-mentioned constraints. For these values, the solutions are shown in blue; thus, the centre distances between 86 mm and 90 mm are suitable for the first speed gear pair according to the colour scale.

Additionally, Fig. 16 was generated using the same solutions to optimise the centre distance for the case shown in Fig. 15. In the figure, the horizontal axis represents the centre distances, the vertical axis represents the flank safety of the gear pair and the colour scale represents the root safety of the gear pair. Increasing the centre distance is shown to increase the flank safety. Average flank safety must be approximately 1.1; however, the average flank safety is <1.1 for centre distances of 86 mm and 87 mm. Therefore, centre distances of 88 mm, 89 mm and 90 mm are suitable for the first speed gear pair based on Figs. 15 and 16.

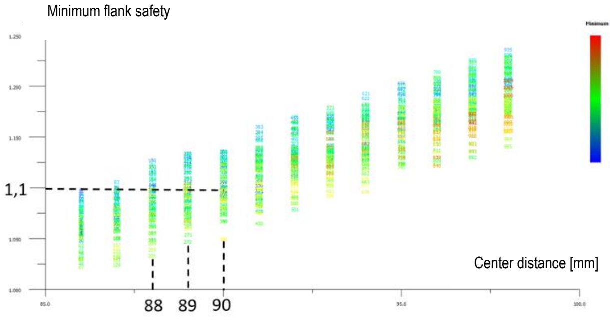


Fig. 16. Centre distance for the first speed gear pair

Then, optimisation was performed for the second speed gear pair, and 1,584 solutions were found according to the input constraints. Figs. 17 and 18 show that the tip diameter of the driver gear can reach 65.5 mm, and the tip diameter of the driven gear can reach 120.5 mm. Additionally, centre distances between 86 mm and 89 mm are shown to be suitable.

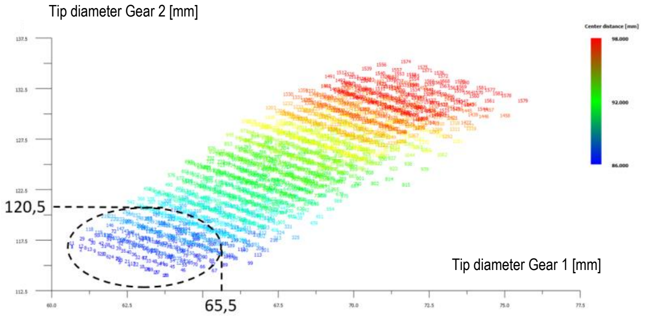


Fig. 17. Centre distance for the second gear pair

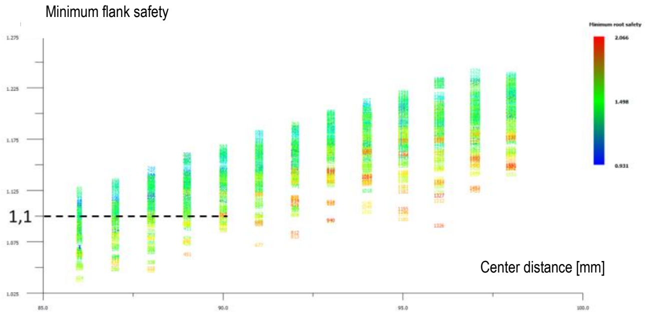


Fig. 18. Centre distance for the second gear pair

An optimisation of the third speed gear pair was then performed and yielded 2,265 solutions. As shown in Figs. 19 and 20, the tip diameter of the driver gear can reach 88 mm, and the tip diameter of the driven gear can reach 98 mm. A centre distance between 86 mm and 90 mm is shown to be suitable, and values >90 mm are overdesigned for this gear pair.

Finally, an optimisation was performed for the fourth speed gear pair, and 1,997 solutions were found. As shown in Figs. 21 and 22, the tip diameter of the driver gear can reach 107.5 mm, and the tip diameter of the driven gear can reach 78.5 mm. Additionally, all centre distances are suitable for the fourth speed gear pair, but other gear pairs must be considered during optimisation.

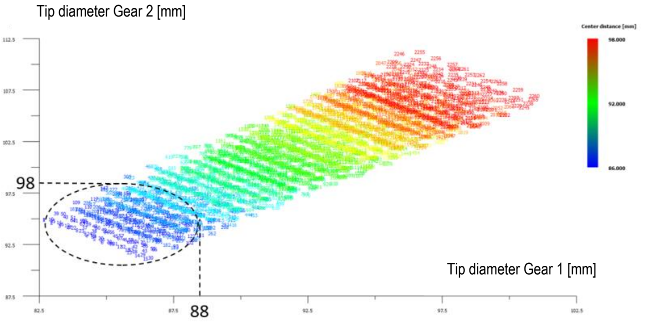


Fig. 19. Centre distance for the third gear pair

Results show that the centre distance can be between 88 mm and 90 mm when considering all gear pairs. Therefore, a centre distance of 89 mm was selected based on all of the optimisations performed.

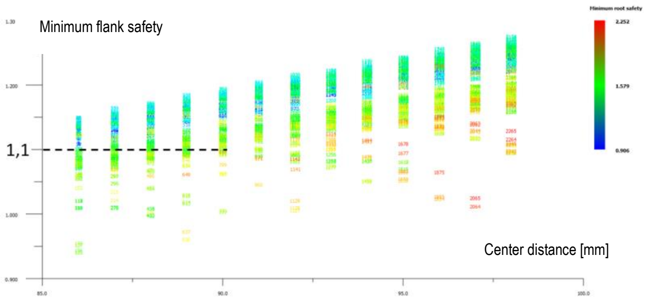


Fig. 20. Centre distance for the third gear pair

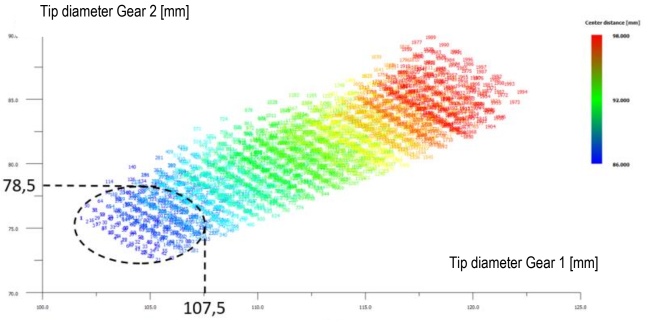


Fig. 21. Centre distance for the fourth speed gear pair

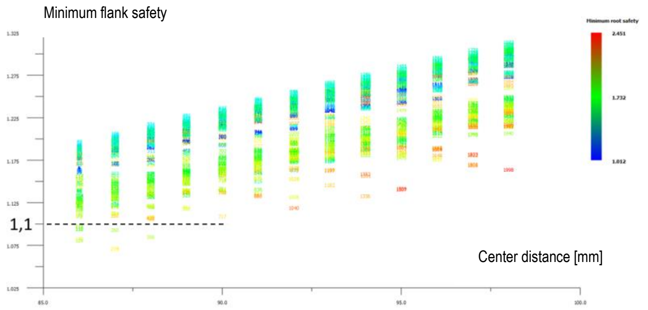


Fig. 22. Centre distance for the fourth speed gear pair

4.3. Last-level module optimisation

During the first-level module optimisation, the module was defined as a range for all speed gear pairs according to the results. A last-level module optimisation was then performed to determine certain modules.

This optimisation was performed for the first speed gear pair while considering the following input values: 50 kW power, 238 Nm torque, 3.1 gear ratio, 20° pressure angle, 17° helix angle, 89 mm centre distance, quality of 6, 40 mm gear width, and a module between 1.5 mm and 2.75 mm. With these inputs, 230 solutions were found. In Fig. 23, the horizontal axis represents the module, the vertical axis represents the root safety and the colour scale represents the flank safety.

The root safety must be between 1.3 and 1.5 for this last-level optimisation, and a 1.75 mm module yields a root safety between 1.3 and 1.5. However, according to previous experience from field tests, the durability of the first speed gear pair can be problematic due to the associated high gear ratio; therefore, a 2 mm module was used to increase safety. This situation is only valid for the first speed gear pair, and there is no need to overdesign the other speed gear pairs.

Then, an optimisation was performed for the second speed gear pair. According to calculations based on input constraints, 115 solutions were found and are shown in Fig. 24, where the root safety is between 1.3 and 1.5 with 2 mm modules.

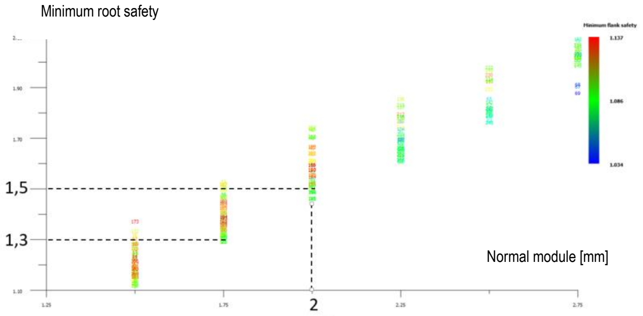


Fig. 23. Module optimisation for the first speed gear

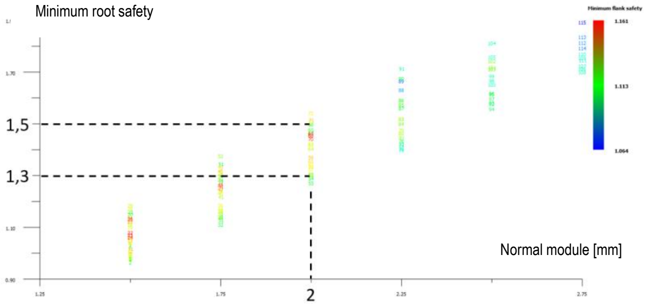


Fig. 24. Module optimisation for the second speed gear

Then, an optimisation was performed for the third speed gear pair, and 162 solutions were found, as indicated in Fig. 25. A 2 mm module is shown to be suitable for root safety, which is between 1.3 and 1.5.

Finally, an optimisation was performed for the fourth speed gear pair, and 162 solutions were found, as indicated in Fig. 26. A 1.75 mm module is shown to be satisfy root-safety requirements, which is between 1.3 and 1.5.

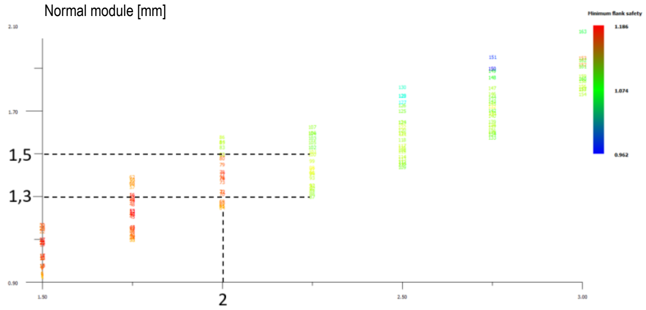


Fig. 25. Module optimisation for the third speed gear

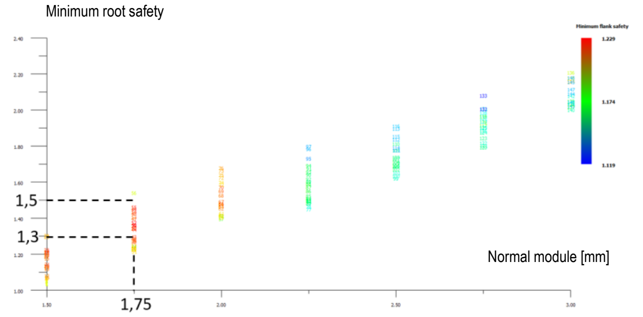


Fig. 26. Module optimisation for the fourth speed gear

DOI 10.2478/ama-2023-0016

4.4. Number of teeth optimisation

After optimising the gear width, gear quality, centre distance and module, the number of teeth of the gear pair that is suitable for the defined ratio must be determined. Therefore, an optimisation of the number of gear teeth was performed.

First, an optimisation was performed for the first speed gear pair with the following input values: 50 kW power, 238 Nm torque, 3.1 gear ratio, 20° pressure angle, 17° helix angle, 89 mm centre distance, gear quality of 6, 40 mm gear width and 2 mm module.

All gear pairs with defined ratios were calculated with KISSsoft. According to calculations, 14 different gear pairs were found, as shown in Tab. 7, which describes the number of teeth of driver gear (z_1), the number of teeth of driven gear (z_2), the profile shifts of driver gear (x_1), the profile shifts of driven gear (x_2), the total contact ratio (ε_γ), the maximum specific sliding ($\zeta_{max.}$) and the gear ratio (i).

In Fig. 27, the horizontal axis represents the maximum specific sliding, the vertical axis represents the number of teeth of driver gear and the colour scale represents the total contact ratio.

The maximum specific sliding should be between –1 and +1 to avoid wear on the gears. In the figure, solutions 2, 3, 6 and 9 appear to meet the requirements of specific sliding. Additionally, solution 9 has the highest contact ratio according to the colour scale. The gear pair that has a high contact ratio does not produce much noise during meshing. Therefore, a high contact ratio is desirable for gear systems. According to the results, solution 9 can be used for the first gear pair. However, as shown in Tab. 7, solution 9 has 21 teeth for the driver gear and 63 teeth for the driven gear. In this case, the ratio is 63/21 = 3, which causes wear on the teeth because the same teethes work during meshing. For this reason, solution 6 with a gear pair of 20–64 was preferred.

Tab. 7. Number of teeth optimised for the first speed gear pair

Sol. no.	z_1	z_2	x_1	x_2	ε_γ	$\zeta_{maks.}$	i
1	20	63	0.31984	0.87368	3.215	–1.268	3.15
2	20	63	0.41984	0.77368	3.203	–1.04	3.15
3	20	63	0.51984	0.67368	3.189	–0.847	3.15
4	20	64	0.20295	0.40344	3.319	–1.619	3.2
5	20	64	0.30295	0.30344	3.304	–1.294	3.2
6	20	64	0.40295	0.20344	3.287	–1.031	3.2
7	21	63	0.1933	0.41309	3.325	–1.529	3
8	21	63	0.2933	0.31309	3.311	–1.237	3
9	21	63	0.3933	0.21309	3.295	–0.996	3
10	21	64	0.09127	–0.03291	3.417	–2.003	3.048
11	21	64	0.19127	–0.13291	3.4	–1.561	3.048
12	21	64	0.29127	–0.23291	3.381	–1.218	3.048
13	21	65	0.01277	–0.46034	3.496	–2.794	3.095
14	21	65	0.11277	–0.56034	3.472	–2.023	3.095

Then, an optimisation was performed for the second speed gear pair, and 21 solutions were found, as indicated in Fig. 28. According to Fig. 28, solution 18 with a gear pair of 30–55 is suitable in terms of both the specific sliding and total contact ratio.

Then, an optimisation was performed for the third speed gear pair, and 24 solutions were found, as indicated in Fig. 29. Solution 13 with a gear pair of 40–45 is optimal.

Finally, an optimisation was performed for the fourth speed gear pair, and 27 solutions were found, as indicated in Fig. 30. According to the results, solution 20 with a gear pair of 57–41 is suitable in terms of specific sliding and the total contact ratio.

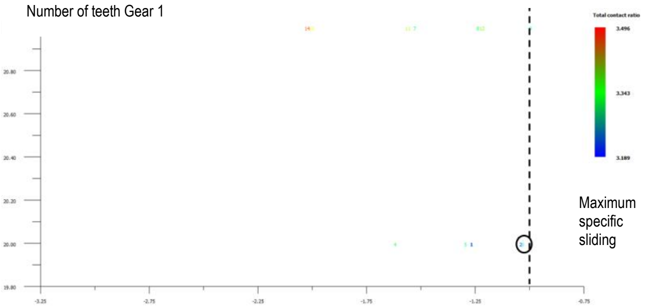


Fig. 27. Number of teeth optimised for the first speed gear pair

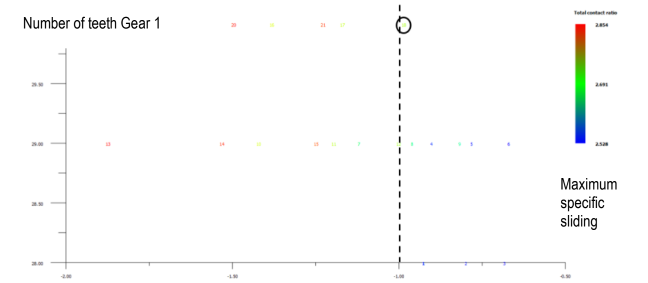


Fig. 28. Number of teeth optimised for the second speed gear pair

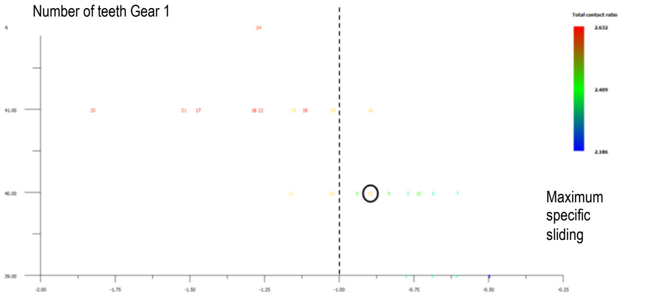


Fig. 29. Number of teeth optimised for the third speed gear pair

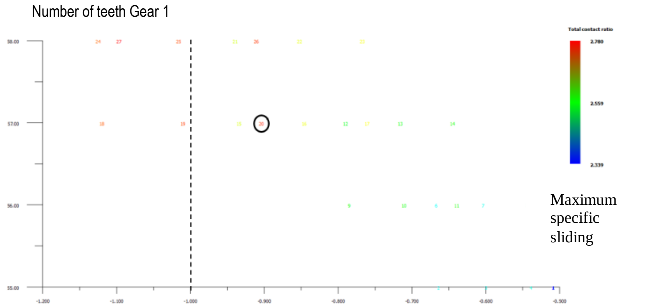


Fig. 30. Number of teeth optimisation for fourth speed gear pair

4.5. Helix angle optimisation

The helix angle of gear pairs was optimised after optimising the gear width, gear quality, module and number of teeth. Currently, the gear pair uses a helix angle so that it can operate quietly due to the high contact ratio. Although such a gear pair has an advantage in terms of the noise level, gear pairs with helix angles generate axial forces on the systems as shown in Fig.31. Therefore, the helix angle is important for the design of shafts and bearings.

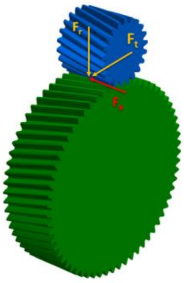


Fig. 31. Forces on gear

An optimisation of the helix angle for the first speed gear pair was performed while considering the following input values: 50 kW power, 238 Nm torque, 3.1 gear ratio, 20° pressure angle, 89 mm centre distance, quality of 6, 40 mm gear width, 2 mm module, and helix angles of 13°, 15°, 17° and 19°, which are geometrically satisfactory.

The calculated axial force and total contact ratio according to the helix angle are shown in Tab. 8. A higher contact ratio results in increased axial forces, and a high contact ratio tends to produce lower noise levels; however, a high axial force is not ideal for shaft and bearing systems.

For this system, a contact ratio of 2.5 appears minimal based on previous experience from field tests. According to Tab. 8, all helix angles have a contact ratio that is >2.5; thus, a 13° helix angle was selected for the first speed gear pair due to its minimal axial force.

Tab. 8. Helix angles for first speed gear pair

Helix angle β (°)	Axial force F_a (N)	Contact ratio (ε_y)
13	2,685.2	2.774
15	3,089.4	3.062
17	3,489.9	3.342
19	3,886.2	3.610

The results for the second speed gear pair are shown in Tab. 9, and a 15° helix angle was selected for the second speed gear pair due to its minimal helix angle, which yields a total contact ratio that is >2.5.

Tab. 9. Helix angle for second speed gear pair

Helix angle β (°)	Axial force F_a (N)	Contact ratio (ε_y)
13	1,790.1	2.364
15	2,059.6	2.566
17	2,326.6	2.759
19	2,590.8	2.934

The results for the third and fourth speed gear pairs are shown in Tabs. 10 and 11. A 17° helix angle for the third speed gear and a 15° helix angle for the fourth speed gear were selected based on the total contact ratio.

Tab. 10. Helix angles for third speed gear pair

Helix angle β (°)	Axial force F_a (N)	Contact ratio (ε_y)
13	1,342.6	2.170
15	1,544.7	2.358
17	1,745.0	2.536
19	1,943.1	2.694

Tab. 11. Helix angles for fourth gear pair

Helix angle β (°)	Axial force F_a (N)	Contact ratio (ε_y)
13	1,076.8	2.365
15	1,238.9	2.574
17	1,399.5	2.763

4.6. Addendum modification coefficient optimisation

An optimisation was then performed on the addendum modification coefficient, which can be determined based on the defined module, helix angle and centre distance.

First, an optimisation was performed for the first speed gear pair while considering the following inputs: 50 kW power, 238 Nm torque, 3.1 gear ratio, 20° pressure angle, 13° helix angle, 89 mm centre distance, quality of 6, 40 mm gear width, 2 mm module, and an addendum modification coefficient between 0 and +0.7 that is geometrically satisfactory. Results of this optimisation are shown in Tab. 12.

Tab. 12. Addendum modification coefficient for the first speed gear pair

x_1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
x_2	1.54	1.44	1.34	1.24	1.14	1.04	0.94	0.84
$\zeta_{1min.}$	-2.52	-2.07	-1.70	-1.41	-1.16	-0.95	-0.78	-0.63
$\zeta 1_{maks}$	0.26	0.30	0.33	0.36	0.39	0.41	0.43	0.46
$\zeta_{2min.}$	-0.36	-0.43	-0.50	-0.57	-0.64	-0.71	-0.78	-0.85
ζ_{2maks}	0.71	0.67	0.63	0.58	0.53	0.48	0.43	0.38
ε_y	2.77	2.77	2.76	2.75	2.74	2.73	2.72	2.70
S_{F1}	1.32	1.35	1.38	1.40	1.42	1.43	1.44	1.45
S_{F2}	1.70	1.63	1.56	1.51	1.46	1.42	1.39	1.36
S_H	1.09	1.09	1.09	1.08	1.08	1.07	1.07	1.06

In Tab. 12, increasing the addendum modification of the driver gear from 0 to +0.7 results in a decrease in the addendum modification of the driver gear from 1.54 to 0.84 due to the constant centre distance. Increasing the addendum modification of the driver gear from 0 to +0.7 results in an increase in root safety from 1.32 to 1.45 due to an increase in tooth thickness. Additionally, decreasing the addendum modification of the driven gear results in decreasing root safety from 1.70 to 1.36 due to thinning of tooth thickness.

Increasing the addendum modification of the driver gear results in a decrease in the total contact ratio from 2.77 to 2.70. Additionally, this change has some effect on flank safety.

Increasing the addendum modification of the driver gear from 0 to +0.7 results in decreasing the specific sliding of the driver gear from 2.52 to –0.63 of the driver gear, increasing the specific sliding of the driven gear from –0.3 to –0.85. According to the results, an addendum modification of +0.6 should be used for this system. Additionally, Fig. 32 shows the associated specific sliding according to the angle of rotation of the gear. In the figure, A-B-C-

D-E represents the contact point of the gear pair during meshing, the red curve represents the driver gear and the green curve represents the driven gear.

Then, an optimisation was performed for the second speed gear pair, and the results are shown in Tab. 13 and Fig. 33. According to the results, an addendum modification coefficient of +0.3 for the driver gear is suitable.

Then, an optimisation was performed for the third speed gear pair, and the results are shown in Tab. 14 and Fig. 34. According to the results, an addendum modification coefficient of +0.1 for the driver gear is suitable.

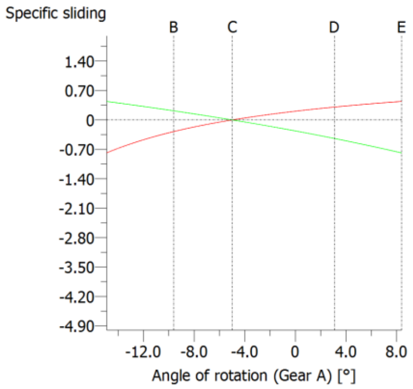


Fig. 32. Specific sliding of the first speed gear pair for +0.6 of the addendum modification coefficient of the driver gear

Tab. 13. Addendum modification coefficient for the second speed gear pair

x_1	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5
x_2	0.72	0.62	0.52	0.42	0.32	0.22	0.12	0.02
ζ_{1min}	-1.83	-1.59	-1.37	-1.18	-1.01	-0.86	-0.72	-0.60
ζ_{1maks}	0.30	0.33	0.37	0.40	0.42	0.45	0.48	0.50
Z_{2min}	-0.43	-0.50	-0.58	-0.66	-0.75	-0.83	-0.92	-1.01
Z_{2maks}	0.64	0.61	0.57	0.54	0.50	0.46	0.42	0.37
ε_y	2.56	2.56	2.56	2.56	2.55	2.54	2.53	2.52
S_{F1}	1.24	1.27	1.29	1.31	1.33	1.34	1.34	1.35
S_{F2}	1.36	1.34	1.33	1.32	1.31	1.29	1.28	1.26
S_H	1.14	1.14	1.14	1.14	1.13	1.13	1.13	1.12

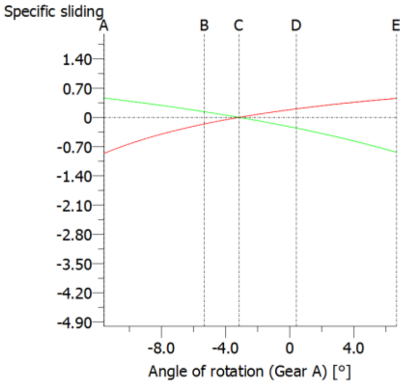


Fig. 33. Specific sliding of the second speed gear pair for +0.3 of addendum modification of the driver gear

Finally, an optimisation was performed for the fourth speed gear pair, and the results are shown in Tab. 15 and Fig. 35. According to the results, an addendum modification coefficient of 0 for the driver gear is suitable.

Tab. 14. Addendum modification coefficient for the third speed gear pair

x_1	-0.2	-0.1	0	0.1	0.2	0.3
x_2	0.25	0.15	0.05	-0.04	-0.14	-0.24
ζ_{1min}	-1.24	-1.09	-0.96	-0.84	-0.72	-0.61
ζ_{1maks}	0.38	0.42	0.45	0.48	0.51	0.54
Z_{2min}	-0.63	-0.73	-0.83	-0.94	-1.06	-1.19
Z_{2maks}	0.55	0.52	0.49	0.45	0.42	0.38
ε_y	2.532	2.535	2.536	2.534	2.529	2.522
S_{F1}	1.273	1.294	1.311	1.324	1.333	1.338
S_{F2}	1.338	1.329	1.318	1.303	1.284	1.262
S_H	1.128	1.129	1.129	1.128	1.127	1.125

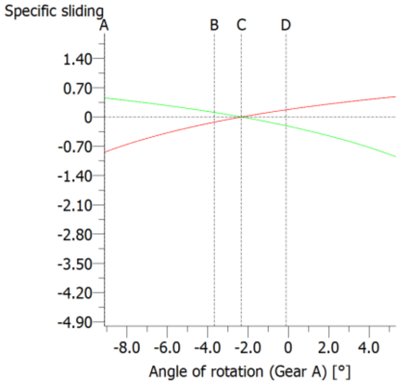


Fig. 34. Specific sliding of the third speed gear pair for +0.1 of addendum modification of the driver gear

Tab. 15. Addendum modification coefficient for the fourth speed gear pairs

x_1	–0.2	–0.1	0	0.1	0.2	0.3
x_2	0.32	0.22	0.12	0.02	–0.07	–0.17
ζ_{1min}	–0.96	–0.87	–0.78	–0.69	–0.61	–0.53
ζ_{1maks}	0.35	0.39	0.43	0.46	0.50	0.53
Z_{2min}	–0.56	–0.65	–0.76	–0.88	–1.00	–1.14
Z_{2maks}	0.49	0.46	0.43	0.41	0.38	0.34
ε_y	2.558	2.567	2.574	2.579	2.581	2.580
S_{F1}	1.284	1.300	1.315	1.327	1.339	1.350
S_{F2}	1.351	1.346	1.340	1.331	1.319	1.304
S_H	1.214	1.216	1.219	1.220	1.223	1.226

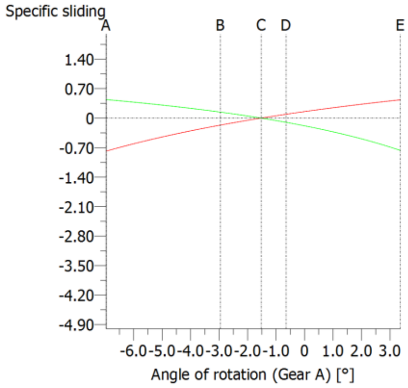


Fig. 35. Specific sliding of the fourth speed gear pair for +0.1 of addendum modification of the driver gear

4.7. Pressure-angle optimisation

After optimising other geometric parameters, pressure-angle optimisation was performed while considering pressure angles of 14°, 16°, 18°, 20° and 22°. Results are shown in Tab. 16.

Tab. 16. Pressure angles for the first speed gear pair

a_n	14º	16º	18º	20º	22º
x_1	0.6	0.6	0.6	0.6	0.6
x_2	1.0894	1.0236	0.9766	0.9419	0.9157
ζ_{1min}	-1.53	-1.19	-0.95	-0.78	-0.65
ζ_{1maks}	0.50	0.48	0.46	0.43	0.41
Z_{2min}	-1.00	-0.93	-0.86	-0.78	-0.71
Z_{2maks}	0.60	0.54	0.48	0.43	0.39
ε_y	2.784	2.772	2.750	2.724	2.697
S_{F1}	1.410	1.419	1.432	1.447	1.463
S_{F2}	1.367	1.363	1.373	1.391	1.412
S_H	1.017	1.038	1.055	1.070	1.082

Although the addendum modification of the driver gear was constant, the addendum modification of the driven gear changed due to changing the pressure angle from 14° to 22°. Increasing the pressure angle results in a decrease in the total contact ratio due to thinning of the top section of the gear.

Increasing the pressure angle results in increased root safety of the driver gear from 1,410 to 1,463 due to thickening of the tooth root.

Specific sliding was not limited at pressure angles of 14° and 16° but was at 18°, 20° and 22°. However, the difference among these values is small; therefore, a pressure angle of 20° was selected because it is a standard value that is used by most manufacturers. Additionally, the optimisation results are similar for other speed gear pairs, and a pressure angle of 20° was also selected for the other gear speed pairs.

4.8. Backlash optimisation

Backlash optimisation was performed after determining all geometric parameters of the gear pairs. Backlash occurs in gear pairs due to installation faults, gear quality and thermal expansion of the system. Therefore, some backlash is generated with tooth thickness tolerances and centre-distance tolerances. These two factors are optimised for suitable backlash.

After optimising the first speed gear pair, the tooth thicknesses of the gears are shown in Fig. 36.

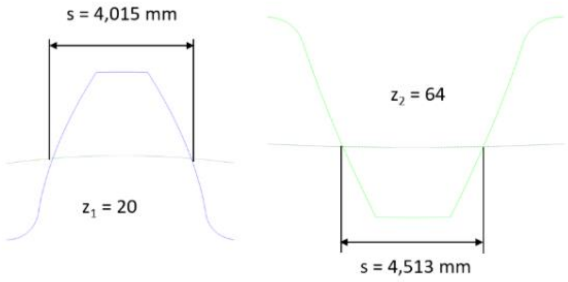


Fig. 36. Tooth thickness of first speed gear pair

There are different methods that are used by manufacturers to measure backlash. In this study, circumferential backlash is considered during optimisation as shown in Fig.37.

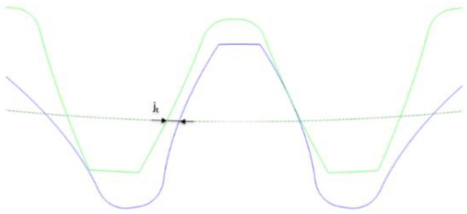


Fig. 37. Circumferential backlash

In this gear system, suitable backlash is considered to be between 0.1 mm and 0.3 mm based on practical experience. Therefore, tooth thickness tolerances and centre-distance tolerances are determined according to the considered backlash values. First, the backlash value is calculated while considering only the tooth thickness of the first speed gear pair; results are shown in Tab. 17. According to Tab. 17, backlash is suitable for tooth thickness tolerances between –0.05 mm and –0.14 mm.

Tab. 17. Backlash for first speed gear pair

Tolerance (mm)	Backlash (mm)
-0.03	0.064
-0.05	0.106
-0.08	0.170
-0.13	0.275
-0.14	0.297
-0.15	0.318

In addition to tooth thickness tolerances, centre-distance tolerances affect the backlash values. If gear pairs are near each other, backlash decreases, and if the gear pair's axes are more distant, backlash increases as shown in Fig.38. Centre-distance tolerances consist of the dimensional deviation of the shaft, bearing and casting.

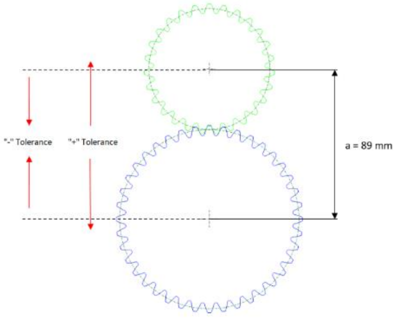


Fig. 38. Centre-distance tolerances

As shown in Tab. 18, backlash values were calculated considering the ISO 286 standard of centre-distance tolerances and tooth thickness tolerances between –0.05 mm and –0.14 mm. As mentioned before, suitable backlash is between 0.1 mm and 0.3 mm. According to Tab. 17, the backlash values are near the target for J6, J7 and J8 centre-distance tolerances. Additionally, the backlash values are outside of the target for J9 and J10. Although the backlash is near the target for J6 and J7, these tolerance groups are not suitable in terms of manufacturing tractor parts. Therefore, the tolerances of J8 are selected for this system.

Tab. 18. Backlash for first speed gear pair

Standards	Positive (+) tolerance (mm)	Negative (–) tolerance (mm)	Min. backlash (mm)	Max. backlash (mm)
-	0	0	0.106	0.297
ISO 286 J6	+0.011	-0.011	0.096	0.307
ISO 286 J7	+0.0175	-0.0175	0.090	0.313
ISO 286 J8	+0.027	-0.027	0.081	0.322
ISO 286 J9	+0.0435	-0.0435	0.066	0.337
ISO 286 J10	+0.07	-0.07	0.041	0.362

Then, tooth thickness tolerances were calculated again considering ISO 286 J8 so that the backlash values are between 0.1 mm and 0.3 mm. Results are shown in Tab. 19. According to the results, –0.06 mm and –0.13 mm tolerances of tooth thickness are suitable for the first speed gear pairs.

The tolerances of other gear pairs and dimensions for inspection of all the gear pairs were calculated similarly and are shown in Tab. 19.

4.9. Optimisation results

The optimisation results for the four speed gear groups are shown in Tab. 19. Additionally, the dimensions and tolerances for inspection of all the gears are shown in Tab. 19.

The optimal four speed gears were placed in a volume, as shown in Fig. 39, which confirms that the design meets the volume constraint; all gears can be assembled properly in the volume.

Tab. 19. Optimisation results

Parameters	1. Speed	2. Speed	3. Speed	4. Speed
Number of teeth z_1	20	30	40	57
Number of teeth z_2	64	55	45	41
Module (mm)	2	2	2	1.75
Pressure angle (°)	20	20	20	20
Helix angle (°)	13	15	17	15
Addendum modification coefficient x_1	0.6	0.3	0.1	0
Addendum modification coefficient x_2	0.9419	0.2201	-0.0416	0.1297
Face width (mm)	40	25	20	20
Quality	6	7	8	8
Total weight (kg)	5.262	2.861	2.133	2.155
Tooth thickness of z_1 (mm)	3.885–3.955	3.448–3.518	3.157–3.227	2.619–2,689
Tooth thickness of z_2 (mm)	4.383–4.453	3.332–3.402	2.951–3.021	2.784–2.854
Base tangent length between k teeth of z_1	21.967–22.033 k: 4 teeth	27.785–27.851 k: 5 teeth	33.761–33.827 k: 6 teeth	35.001–35.066 k: 7 teeth
Base tangent length between k teeth of z_2	59.188–59.253 k: 10 teeth	46.162–46.227 k: 8 teeth	33.726–33.792 k: 6 teeth	29.557–29.623 k: 6 teeth
Dimension over balls/ball diameter of z_1	49.800–49.925 D: 4.250	68.655–68.807 D: 3.750	88.706–88.877 D: 3.500	107.068–107.249 D: 3.000
Dimension over balls/ball diameter of z_2	140.226–140.381 D: 3.750	119.340–119.512 D: 3.500	98.594–98.772 D: 3.500	78.472–78.644 D: 3.000
Centre distance (mm)	88.973–89.027			
Backlash (mm)	0.102–0.300	0.102–0.300	0.105–0.293	0.105–0.293

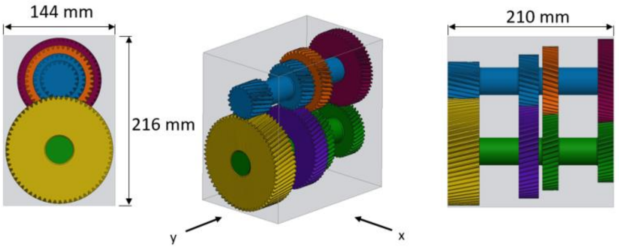


Fig. 39. Speed gears in a volume

5. RESULTS AND DISCUSSION

Many results and relations between geometric parameters of the gear pairs were observed during optimisation. These relations are important when designing efficient gear systems.

Changing the module was found to affect the root safety by changing the gear profiles, as shown in Fig. 40. Increasing the module has an effect on the tooth thickness and seriously affects root safety.

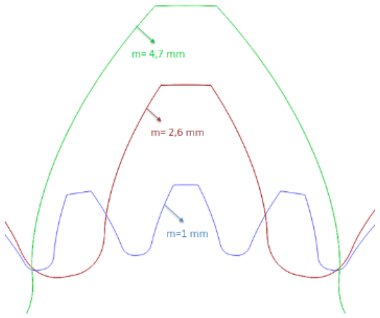


Fig. 40. Gear profiles based on the module

Apart from gear profiles, the module affects the overlap ratio, which also affects both the durability and noise level of the gear pairs. In Figs. 41–43, the contact pattern of the gear pairs is shown on the face width. Decreasing the module thus increases the overlap ratio due to the decreased size of the gear profiles.

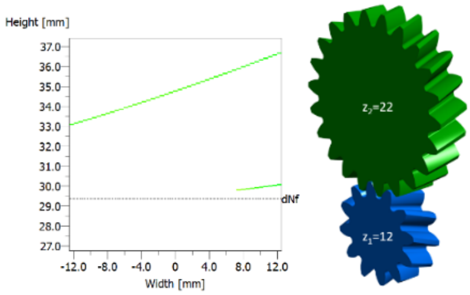


Fig. 41. Contact pattern (m = 5 mm, i = 1.8)

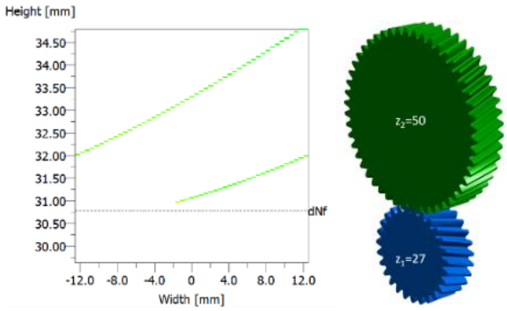


Fig. 42. Contact pattern (m = 2.3 mm, i = 1.8)

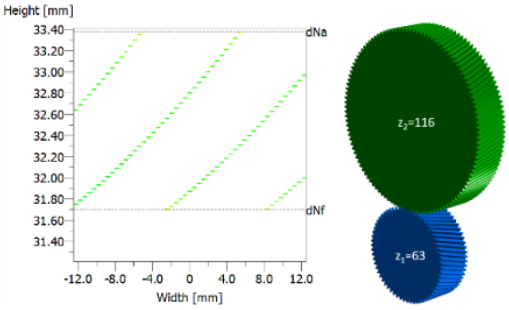


Fig. 43. Contact pattern (m = 1 mm, i = 1.8)

Finally, the module affects the weights of the gear pair, which is an important constraint for the systems. As shown in Fig. 44, increasing the module increases the weight for the same ratio of gear pairs.

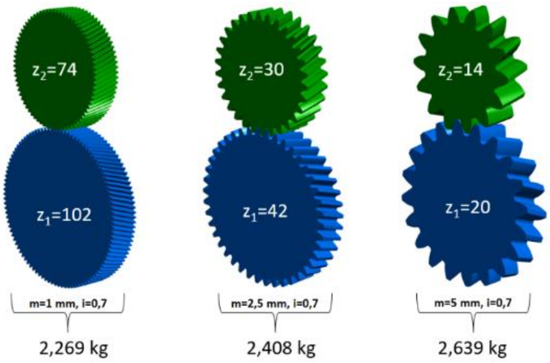


Fig. 44. Gear pair weight according to module

The tooth bending stress, tooth contact stress and safety factor of pinion gears for each speed, which were optimised via KISSsoft, were also calculated according to the mathematical model in ISO 6336. The characteristics of the pinion gears from both KISSsoft and the mathematical model are shown in Tab. 20.

Tab. 20. Optimisation results

Optimisation results	1. Pinion	2. Pinion	3. Pinion	4. Pinion
Tooth-root stress σ_F from KISSsoft (N/mm^2)	543.96	569.46	576.47	570.03
Tooth-root stress σ_F from mathematical model (N/mm^2)	546.97	607.52	625.93	561.37
Safety factor for bending stress S_F from KISSsoft	1.41	1.34	1.32	1.34
Safety factor for bending stress S_F from mathematical model	1.46	1.32	1.28	1.43
Contact stress σ_H from KISSsoft (N/mm^2)	1,363.60	1,286.05	1,292.8	1,197.23
Contact stress ϬH from mathematical model (N/mm2)	1,345.09	1,264.66	1,239.2	1,226.15
Safety factor for contact stress S_H from KISSsoft	1.10	1.13	1.13	1.22
Safety factor for contact stress S_H from mathematical model	1.12	1.19	1.21	1.22

In Tab. 20, the tooth-root stress of pinion 1 is 546.97 N/mm^2 according to the mathematical model, which is 0.5% larger than the KISSsoft result (543.96 N/mm^2). For the root-safety factor, the results of the mathematical model are 1.46% and 3.5% larger than the results of KISSsoft (1.41). The tooth contact stress of pinion 1 is 1345.09 N/mm^2 according to the mathematical model, which is 1.4% smaller than the results of KISSsoft (1363.60 N/mm^2). For the flank safety factor, the mathematical model result is 1.12 (1.8% larger) than the KISSsoft result (1.10).

For pinion 2, the mathematical model result is 6.7% larger than the KISSsoft result for tooth-root stress. The tooth contact stress according to the mathematical model is 1.7% smaller than the KISSsoft result. For the root-safety factor, the mathematical model result is 1.5% smaller than the KISSsoft result, and the flank safety factor calculated by the mathematical model is 5.3% larger than the KISSsoft result.

For pinion 3, the tooth-root stress according to the mathematical model is 8.5% larger, and the tooth contact stress according to the mathematical model is 4.1% smaller. For the root-safety factor, the result of the mathematical model is 3% smaller, and the flank safety factor for the KISSsoft result is 7% smaller than that of the mathematical model.

For pinion 4, the tooth-root stress according to the mathematical model is 1.5% smaller, and the tooth contact stress is 2.4% larger than the KISSsoft result. The root-safety factor of the mathematical model is 6.7% larger than the KISSsoft result, and the flank safety factors are the same.

According to the results, there is a maximum 8.5% difference between the KISSsoft and mathematical model results (σ_F = 625.93 N/mm^2 by mathematical model and σ_F = 576.47 N/mm^2 by KISSsoft) since KISSsoft considers the tolerances and deformation of gears, and KISSsoft specifies correction coefficients based on user inputs. Although there are some differences between the results, the safety factors calculated using both methods are suitable based on the target values. Therefore, KISSsoft can be used reliably to calculate the strength of gears.

This study is not an improvement study of any design. It is an original work, and therefore there are no pre-existing values with which to compare the results obtained.

6. CONCLUSION

Four speed gears in a tractor transmission were optimised using KISSsoft software. During optimisation, the input power torque, ratios and maximum volume were considered constraints, and the face width, centre distance, module, quality, number of teeth, helix angle, addendum modification coefficient and pressure angle of gear pairs were specified as inputs for the optimisation according to a flow chart. Then, the tooth-root stress, tooth contact stress and safety factors were calculated according to the mathematical model described in ISO 6336. Results were then compared. Regarding tooth-root stresses, the maximum difference was 8.5% for pinion 3. Regarding tooth contact stresses, the maximum difference was 4.1% for pinion 3. Regarding root-safety factors, the maximum difference was 6.7% for pinion 4. Regarding the flank safety factors, the maximum difference was 7% for pinion 3.

According to this study, both the KISSsoft software's results and the mathematical model's results are within the range of the target value. Additionally, the following results were determined via optimisation:

1. Increasing the module increases the root-safety factor and decreases the flank safety factor.
2. Increasing the face width of gears increases the flank safety factor.
3. Increasing the gear quality results increases the flank safety factor.
4. Decreasing the module results increases the number of teeth and overlap ratio.
5. Decreasing the module results increases the sensitivity of the ratio, which can be chosen for the gear pair.
6. Increasing the centre distance decreases the tooth contact stress.
7. Increasing the helix angle increases the contact ratio and axial forces.
8. Increasing the addendum modification coefficient increases the root-safety factor by increasing the tooth thickness.
9. Increasing the addendum modification coefficient decreases the contact ratio.
10. Increasing the module increases the weight of a gear pair.
11. Increasing the pressure angle increases the tooth thickness at the root zone and decreases the tooth thickness at the tip zone.

Optimised gears with other drivetrain components like shafts, bearings, washers, circlip, synchromeshes are presented as a concept design in Fig. 45-46.

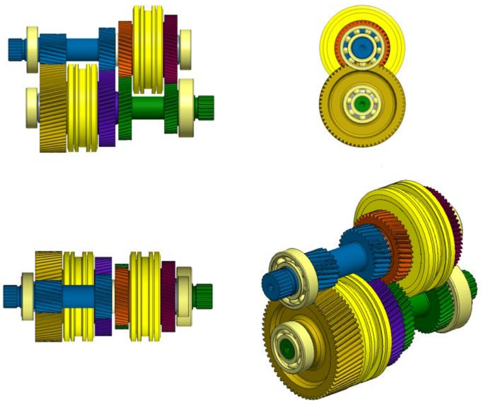


Fig. 45. Concept design

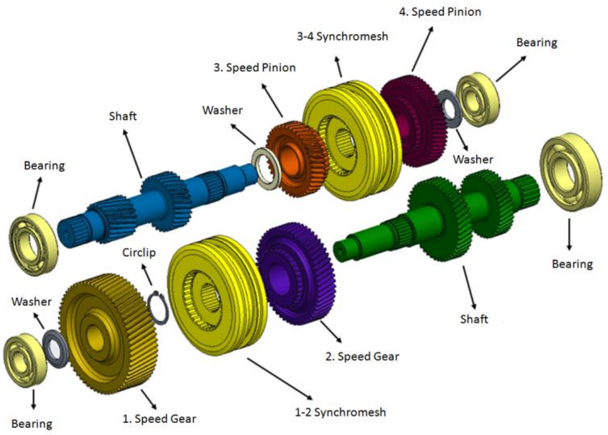


Fig. 46. Disassembly of concept design

The results obtained from this study will be especially beneficial to engineers working in the industry. This is because, in this study, it has been revealed that optimisation serves as a useful tool in the design phase, in which the capability is available to simultaneously and accurately optimise several parameters.

REFERENCES

1. Hlwan HHS, Htay HW, Myint T. Design and contact stress analysis of helical gear for light-weight car. Proceedings of 105th The IIER International Conference. Bangkok. Thailand. 5th – 6th June 2017.
2. Wenkatesh B. Prabhakar SV, Deva PS, Investigate the combined effect of gear ratio, helix angle, facewidth and module on bending and compressive stress of steel alloy helical gear. 3rd International Conference on Materials Processing and Characterisation. 2014;6: 1865-1870.
3. James K, John K. Effect of gear design variables on the dynamic stress of multistage gears. Innovative Systems Design and Engineering. 2012;2:30-42.
4. Murali MV, Ajit SL. Influence of module and pressure angle on contact stresses in spur gears. International Journal of Mechanical Engineering and Robotics Research. 2016;3:224-228.
5. Dyaneshwar S, Mangrulkar KS. Effect of backlash on bending stresses in spur Gears. International Journal of Scientific Development and Research. 2016;7:349-354.

6. Zvonko S, Mileta R, Bozidar R, Dragan R, Zivoslav A. The ınfluence of gear parameters on the surface durability of gear flanks. Strojarstvo. 2011;53(5):383-387.
7. Nijazi I, Sadullah A. The influence of sliding speed and specific sliding of the interval meshing gears. 10th International Research/ Expert Conference Trends in the Development of Machinery and Associated Technology. Barcelona, Spain, 11-15 September 2006.
8. Karadere G, Yilmaz I. Investigation of the effects of profile shift in helical gear mechanisms with analytical and numerical methods. World Journal of Mechanics. 2018;8:200-209.
9. Mujammil A, Mushtaq AC. Optimization of addendum modification for bending strength of involute spur gear. International Engineering Research Journal. 2015;1093-1097.
10. Vishal S. Finite element analysis of helical gear pair for bending and contact stresses. International Journal of Computer Engineering in Research Trends. 2018;5:136-140.
11. Bozca M, Dikmen F. Optimisation of geometric parameters of ears under variable loading condition. Advanced Materials Research. 2012;1005-1010.
12. Bozca M. Optimisation of effective design parameters for an automotive transmission gearbox to reduce tooth bending stress. Modern Mechanical Engineering. 2017;7:35-36.
13. Bozca M, Fietkau P. Empirical model based optimization of gearbox geometric design parameters to reduce rattle noise in an automotive transmission. Mechanism and Machine Theory. 2010;1599-1612.
14. Bozca M. Torsional vibration model based optimization of gearbox geometric design parameters to reduce rattle noise in an automotive transmission. Mechanism and Machine Theory. 2010;1583-1598.
15. Bozca M. Transmission error model-based optimisation of the geometric design parameters of an automotive transmission gearbox to reduce gear-rattle noise. Applied Acoustics. 2017;247-259.
16. ISO 6336–3: calculation of load capacity of spur and helical gears, part 3: calculation of tooth bending strength.
17. ISO 6336–2: calculation of load capacity of spur and helical gears, part 2: calculation of surface durability (pitting).

Acknowledgements: This work was supported by Research Fund of the Yildiz Technical Üniversity (BAP) Project Number: FYL-2018-3506.

Emre Can: https://orcid.org/0000-0002-9298-2616

Mehmet Bozca: https://orcid.org/0000-0002-2620-6053

DOI 10.2478/ama-2023-0017

GEOLOGICAL AND GEOTECHNICAL ASSESSMENT OF AGGREGATES USED IN NAGAR PARKER DISTRICT THARPARKAR SINDH PAKISTAN

Muzafar A. KALWAR*

*Institute of Hydro-Engineering, Polish Academy of Sciences, ul. Kościerska 7, 80-328 Gdańsk, Poland

muzafar.kalwar@pg.edu.pl

received 20 August 2022, revised 18 November 2022, accepted 20 November 2022

Abstract: This study aims to determine the geology of granite and evaluate the engineering properties of the samples to make recommendations for the construction industry. The study area is situated in the Nagar Parker complex in Pakistan, which is located in the extreme south-east of the Thar District and the desert of the Sindh Province, near the Run of Kutch (24° 15′–35 30′ N, 70° 40′–58 07′ E), and it covers ca. 500–1,000 km2. In this region, several Quaternary deposits, subordinate and dispersed Jurassic–Tertiary sandstones and clays are overlying the Nagar Igneous Complex basement. According to international standards, there are various possible aggregate sources. However, only a few of them have been reviewed for suitability reasons. Six quarries in Nagar Parker, Pakistan, were selected for evaluation as coarse aggregate in concrete construction and civil engineering works in this research. Although the aggregates from the six quarries are specified and already widely used in the Sindh Province, there is a lack of studies on their geological properties. The results of the presented research revealed that samples from Dhedvero, Karai, Nagarparkar, Mokrio, Dinsi and Wadlai meet all of the international standard requirements for aggregates. Geotechnical, petrographic and geochemistry laboratory tests were conducted in this research and included bulk density, water absorption, specific gravity test, index of flakiness and elongation, soundness aggregate test, crushing value aggregate, impact value aggregate and abrasion value of Los Angeles. Furthermore, chemical alkali-silica reaction potential test and petrographic examination were tested. As a result, we evaluated the properties of granite, which is a crystalline igneous rock with a visibly crystalline structure and texture, made up of feldspar, i.e., potash feldspar and oligoclase. The evaluated minerals are compatible with the standards of civil engineering works and can be used as a concrete aggregate. The evaluated three types of minerals included Dhedvero simple intrusion, Nagar pink granite and grey granite.

Key words: geotechnical engineering, granite, aggregate, construction, Nagar Parker

1. INTRODUCTION

To Natural building aggregate is one of the most available and commonly used common resources. The majority of roads, bridges, dams, houses and other infrastructure projects are made up of construction aggregate, crushed and sized rock material used in concrete and asphalt. Construction aggregate accounted for >90% of asphalt paving and 80% of concrete. The rest is made up of a binder like asphalt or cement. Crushed stone accounts for about 52% of all building aggregate, with sand and gravel accounting for the remaining 48% [1]. There are many important factors to be considered when deciding if a rock is suitable for use as aggregates in road construction. The aggregate used in the road's surface course (running surface) must be resistant to the polishing action from vehicle tyres, or the road will become slick, particularly when wet. Aggregates used in construction must be durable.

On the other hand, concrete is a construction substance made from coarse cement paste and fine aggregates. Gravel, sand and crushed stone aggregates are examples of coarse to medium-grained particulate material referred to as 'construction aggregate' or 'aggregate'. Aggregates may be either natural or synthetic, depending on their origin, ACI, E1 [2]. Aggregates are the most quarried materials on the planet, and they are used to add strength to composite materials. Besides that, since aggregates make up more than three-quarters of concrete volume—their physic-mechanical chemical—they have bulk volume and reduce concrete cost. The consistency of the finished product is directly influenced by mineral properties [3]. Destructive manifestations, such as alkali-silica/alkali carbonate reactions/physical effects, may occur if aggregate properties are not up to standard, compromising concrete strength and, as a result, lowering construction quality [4].

Aggregates are valuable raw materials used in various industries, with the cement and building industries being the primary consumers. Physical, mechanical, chemical and mineralogical properties of regional or local aggregate resources must all be considered in a strategic assessment. It is the most common source of coarse aggregate in the construction industry. This rock can be found in large quantities in Pakistan. Pakistan has investigated their suitability for use in various industries, including construction [5–7]. Previous studies [8] examined the engineering characteristics of 35 quarries in the Khyber Pakhtunkhwa (KPK) area of Nowshera. The study identifies locally available aggregates in Nowshera, reducing the reliance on the Margalla crush, widely used due to its high quality and suitability for structure [9, 20]. Aggregates from the Allai region were investigated for use in the restoration and renovation of buildings damaged by the October 2005 earthquake [11]. Earlier studies [12] investigated the effects of freezing and thawing on the hardness of coarse aggregate concrete. Little analysis has been done so far to determine the quality of the resources because the aggregate material is already being extracted in the study area. The classification of aggregate is significant, particularly in the construction of megastructures, since the aggregate material determines the quality of concrete [13–14].

At the confluence of the Thar Desert and Rann of Kutch, Pakistan's Nagar Parkar area occupies about 500 km2. A recently constructed metalled road connects it to Mithi and Badin; the Rann and Indian Territories surround it on three sides. This area is characterised by granitic rock mounds and bold hills in an otherwise levelled landscape of sandy, salty and salty plains, except for the Thar Desert, which is mostly covered by dunes. Karunjhar hill, a prominent desert geomorphic landmark in southeastern Pakistan, rises to 356 m a.s.l. At Karai, Dhedvero and the central part of the district, coarse-grained gabbro dykes and medium- to fine-grained dolerites cut through igneous and metamorphic rocks [15]. Several authors have provided detailed information about geology, mineral formations, petrology and the complex's geochemistry [16–20]. However, we believe that more research into the engineering properties of Dhedvero, Karai, Nagarparkar, Mokrio, Dinsi and Wadlai is required.

1.1. Geological setup

The Nagar Parker area is located in the extreme southeast region of Pakistan, on the border with India (Fig. 1).

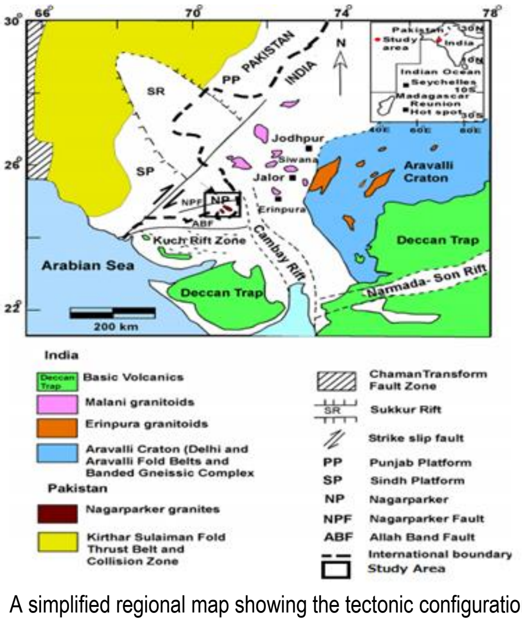


Fig. 1. A simplified regional map showing the tectonic configuration and geographical position of Nagar Parker in the Indian shield [20, 37–38]

It is considered the western continuation of the Precambrian Indian shield [20–23]. Existing works [16] presented a geological map of the district, noting Quaternary deposits, subordinate Jurassic-Tertiary sandstone, clay and the overlying basement, and classifying the region as the Dhedvero simple intrusions, the Nagar Parker pink granite and the Karunjhar grey granite (Fig. 3). Small exposures of schists, paragneisses, migmatites and quartzites can be found to the southeast of the Dhedvero and west of Berana, especially near Walo jo Wandio (Fig. 2) and Moti jo Wandio (not shown on the map). At Dhedvero (24°24'27" N, 70°57'14" E), a 75-m-long and 10-m-wide quartzite body was discovered as xenoblock within coarse-grained pink granite.

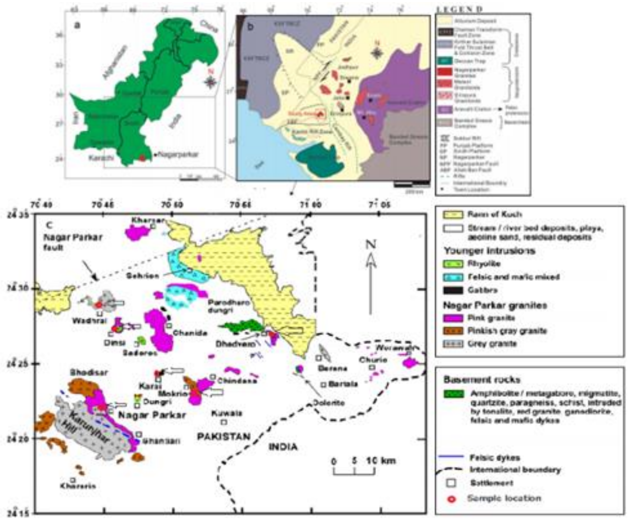


Fig. 2. Regional geological map and geological map of NPIC area, location of samples in the study area (a) the map of Pakistan and location of the study area (b) a simplified regional geological map, showing the tectonic configuration and geographical position of the Nagar Parker area of the Indian shield (modified after the literature [37, 38]) (c) geological map of Nagar Parker, Sindh province of Pakistan [17, 18, 20, 39]

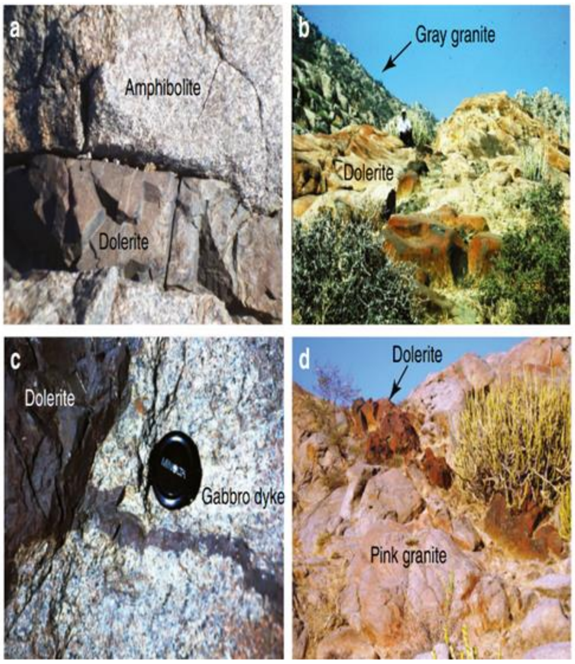


Fig. 3. (a) Dolerite intruding amphibolites at Dhedvero, (b) Dolerite intruding gabbro dyke at Karai, (c) Dolerite intruding gray granite at Nagarparker and (d) Dolerite intruding pink granite at Mokrio [20]

The paragneisses have an N50°W striking angle and a 55° SW dip. Pink granites, mafic and acidic dykes, and amphibolites are intruded by these metasedimentary rock assemblages. Shearing and deformation have occurred in metamorphic rocks, intruded granites, and mafic dykes in various forms. Existing studies [20] classified the Nagar Parker granites into five rock units based on field observations and mineral and geochemical compositions: pink, reddish pink, pinkish grey, grey and greyish white granites. Pink granites have minor differences in mineralogy and geochemistry, but they are distinct from grey and greyish white granites [20].

2. METHODOLOGY

2.1. Field work

This research was limited to the aggregates collected from quarries in District Tharparkar: Dhedvero, Karai, Nagarparkar, Mokrio, Dinsi and Wadlai. We performed the examination and evaluation of the coarse aggregate source material which was performed in the study area. Fig. 2 demonstrates the locations of the quarries, which were sampled for the comparison of the minerals.

2.2. Sampling

The device ASTM D-75 was used to collect representative samples from the quarries when necessary [24]. The conveyor belts and flowing barrels were used as stockpiles to achieve the minimum mass recommended by ASTM D-75 in equal increments. The available quarries for this research study are described in the table below along with their selection status.

The samples were then transported to the following laboratories for processing and examination under the ASTM and BS standards: We performed the geotechnical experiments in the Mehran University of Engineering and Technology Jamshoro's Geotechnical and Concrete Laboratory, as well as the Center for Pure and Applied Geology Jamshoro's Petrographic and Geochemistry Laboratory. The following tests were performed: bulk density, water absorption and specific gravity test, index of flakiness and elongation, soundness aggregate test, crushing value aggregate, impact value aggregate and abrasion value of Los Angeles, and furthermore, chemical alkali-silica reaction potential test and petrographic examination were conducted. In the end, the concrete cylinders were observed by compressive strength test.

3. RESULTS AND DISCUSSIONS

3.1. Bulk density

According to literature [25, 26], all of the examined specimens are standard weight aggregates, with bulk densities ranging from 80 pcf to 1,130 pcf (1,130–1,920 kg/m3).

3.2. Specific gravity and water absorption

The aggregates and specific gravity are measured under various moisture conditions: (a) when the water has been poured through all of the aggregate's apertures, the surface still remains dry. This is referred to as a surface dry saturated state SSD. (b) The aggregate is seen to be air dry (AD) when the aperture water is removed by drying it in the air, and bone dry or oven dry (OD) when all the moisture is removed by drying the aggregate in an oven.

The test was conceded in compliance with existing standards [27]. The standard weight aggregates have a relative bulk gravity of 2.4–2.7 (ASTM C 125-03), so the specimens from all quarries are standard weight aggregates. The quarry aggregates have a water absorption value of <2%, which is suitable for construction purposes.

3.3. Flakiness index (FI) and elongation index (EI)

The coarse aggregate's regular geometric texture is essential because they disturb compaction, workability and stress resistance. The workability of concrete and its bonding properties are improved using smooth and rounded aggregates [28]. The workability of the rough, flaky, angular or elongated aggregate, on the other hand, is reduced. This test was conducted following the existing methodology [29], by mass approach. The FI and EI of samples from most quarries are within the acceptable limits of international standards (Tab. 1).

3.4. Soundness

The aggregates refer to their capability to withstand possible weather circumstances, including freeze and thaw cycles, as well as wetting and drying cycles. Temperature changes cause volume expansion, which compromises the resilience of the concrete. This test was conducted according to the existing methodology [30], and all of the aggregate samples were considered sound because their values were within the ASTM acceptable range (Tab. 1).

3.5. Aggregate crushing value

This test was carried out following the standard BS 812: Part 110 [31]. According to the IS: 2,386 requirements, the following are the acceptable limits:

- Concrete Works = 45%
- Pavement wearing surfaces = 30%

We found the aggregate crushing values for specimens from Dhedvero, Karai, Nagarparkar, Mokrio, Dinsi and Wadlai to be <30%, indicating that they can be used on wear surfaces and in other concrete works. Different samples, on the other hand, can be safely used for both concrete and pavement (see Fig. 4 and Tab. 1).

3.6. Aggregate impact value (AIV)

The studied area as Dhedvero, Karai, and Nagarparkar aggregates are appropriate for heavy-duty concrete floorings or rigid asphalts, according to the guidelines outlined in BS 882 [32]. They can also be used in standard concrete projects (see Tab. 1).

3.7. Loss angeles abrasion value

This Loss Angeles test was carried out according to the standard ASTM C131. For the use in concrete, the appropriate range of abrasion is up to 50% (ASTM C-131 and AASHTO T-96) [33], but the specified limits may differ, depending on the type of use (Tab. 1).

Tab. 1. Summary of the test results

Sr. No.	Quarries	Water Absorption (%)	FI (%)	EI (%)	Soundness (%)	ACV (%)	AIV (%)	LAA (%)	Petrography	ASR	Compressive strength (psi)
01	Dhedvero	0.4	25.8	18.2	2.3	25.1	21.8	16.1	Innocuous	Deleterious	4,809
02	Karai	0.5	22.4	15	4.1	29.4	23.4	17.7	Deleterious	Innocuous	4,910
03	Nagar parkar	0.43	16.2	13.4	1.3	18.3	21.3	15.8	Innocuous	Deleterious	5,128
04	Mokrio	0.46	17.1	32.1	2.1	27.2	24.7	15.9	Deleterious	Deleterious	4,712
05	Dinsi	0.54	14.9	14.6	5.7	22.6	25	18.3	Deleterious	Innocuous	4,739
06	Wadlai	0.64	14.8	14	2.3	21.4	25.9	18.1	Innocuous	Deleterious	4,612

ACV, aggregate compression value; AIV, aggregate impact value; ASR, alkai silica reactivity; EI, elongation index; FI, flakiness index; LAA, Loss Angeles abrasion

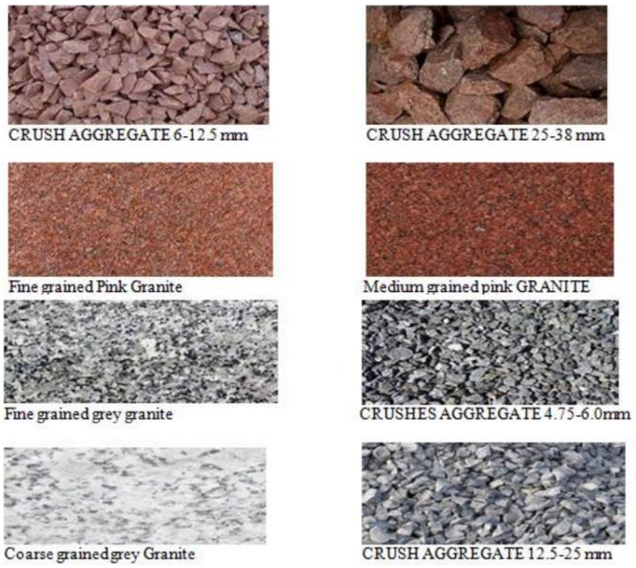

CRUSH AGGREGATE 6-12.5 mm
CRUSH AGGREGATE 25-38 mm
Fine grained Pink Granite
Medium grained pink GRANITE
Fine grained grey granite
CRUSHES AGGREGATE 4.75-6.0mm
Coarse grained grey Granite
CRUSH AGGREGATE 12.5-25 mm

Fig. 4. Study area samples and varieties used in the construction

3.8. Compressive strength

The compressive strength of the aggregates varies from 12,000 psi to 42,000 psi, while tensile strength is 400–2,400 psi [8]. All of the concrete cylinder samples have identical compressive strength values (Tab. 1). Before being tested for strength, the cylinders were cast according to the requirements and cured for 7 days and 28 days. The ratio of the cylinder casting was 1:2:4, and the water/cement ratio was 0.49.

The results can be found in the remaining information in Tab. 1. Although mortar crushed the majority of the cylinders, a few aggregate particles were found split.

3.9. Alkali silica

A chemical reaction between alkalis in the cement (sodium oxide or potassium) and certain siliceous minerals in the aggregate occurs under certain moisture conditions, resulting in the formation of an alkali-silica gel, which causes concrete expansion/swelling and cracking. We used the ASTM C-289 [34], which is the method to determine alkali-silica reactivity. The aggregates from the respite of the quarries were initiated to be stable.

3.10. Petrographic examination

The existing standard ASTM C 295 [35] specifies a petrographic test. The petrography of the samples revealed that Dhedvero and Nagar Parker samples contain harmful materials. Tab. 1 summarizes the petrographic results of all of the test samples.

There are a few coarse-grained gabbro dykes among the medium- to fine-grained dolerites. Biotite is a mineral that surrounds olivine and clinopyroxene grains in minor quantities (Fig. 5a). Clinopyroxene is found as an oikocryst containing plagioclase in sample 1 (Fig. 5b). Dolerites with a medium to fine grain are found all over the world. The dolerites at Karai are about 25 m wide and 52 m long, with amphibolites and gabbros intruding in the N40° W and N30° E directions sample 2 (Figs. 3a,b). Plagioclase, titaniferous augite, hornblende, biotite and titanite are all contained in these dolerites. Accessories include apatite and zircon. The opaques are magnetite, hematite, titanomagnetite, ilmenite, pyrite and chalcopyrite, and dolerites with medium to fine grain, such as brown, pink and greyish-white granites dating back from 1,100 Ma to 750 Ma (Figs. 3c,d). Pink granite at Nagar Parker, Mokrio, Kharsar, and Wadali, and amphibolite and meta-sedimentary rocks at Dhedvero are situated [20].

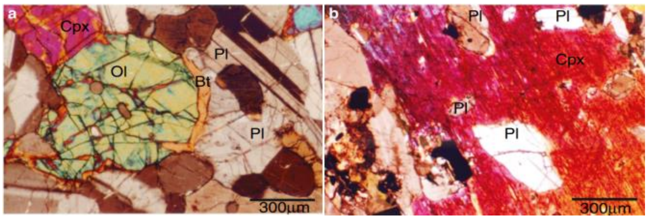


Fig. 5. (a) Photomicrograph of gabbro dyke (sample 1) illustrating olivine, clinopyroxene, and plagioclase. Olivine contains plagioclase as inclusion. Biotite is developed rimming the olivine and clinopyroxene grains. All photomicrographs are taken in crossed polars. (b) Photomicrograph of gabbro dyke (sample 2) showing clinopyroxene oikocryst containing plagioclase

Plagioclase, hornblende and augite are the main minerals in dolerites. They range from subophitic to ophitic in thin sections. Plagioclase, on the one hand, is transformed into sericite, calcite and epidote in part. On the other hand, hornblende is transformed into biotite and chlorite in part. Some of the dolerites that have intruded through the deformed pink granites have been deformed (Fig. 6a). Quartz grains appear as interstitial masses and megacrysts in one of the dyke samples, sample 3. Stresses have fractured the quartz megacryst in sample 3 (Fig. 6b). At the Nagar Parker [36] site, this dyke intrudes 750-Ma reddish-pink granite.

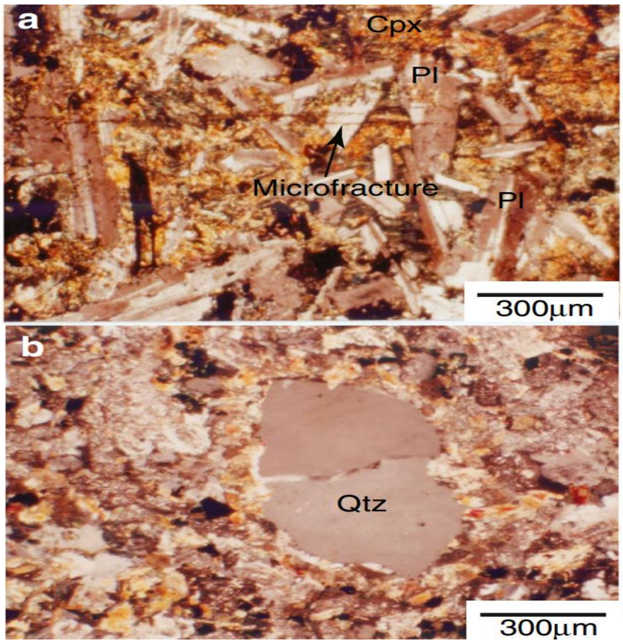


Fig. 6. (a) Photomicrograph of dolerite showing deformation as depicted by microfracture in the rock, the bulk of the rock comprises plagioclase, clinopyroxene, hornblende, biotite and opaques. (b) Photomicrograph of dolerite (sample 3) depicting deformed quartz phenocryst in altered ground mass

The current study revealed that some localities in the Tharparkar and Nagar Parker districts have a substantial supply of coarse aggregate material which can be used as an important material. The crush from the studied area is of high quality, as demonstrated by numerous tests performed on these samples, which revealed that they follow ASTM or BS standards. Next in quality to Nagar Parker and Karai, this has strong physical and chemical properties making it suitable for aggregate use. The Dhedvero crush, on the other hand, had excellent physical results but failed the chemical examination.

4. CONCLUSIONS

The Pakistani Geological Survey estimates over 297 billion tonnes of granite reserves in the Nagar Parkar region of Sindh, Pakistan, with over 25 different granite types in various colours and varieties. We conducted the study to provide information to prospective investors to plan a complete business plan for the quarry of their selection. Granite is a crystalline igneous rock with a visibly crystalline structure and texture. Granite is made up of feldspar, i.e., mostly potash feldspar and oligoclase, primarily whitish or grey, with a speckled appearance caused by the darker crystals. Granite is generally used for outdoor applications, such as flooring tiles, highway design, lodge and house cladding, and so on. Granite contains a specific gravity ranging from 2.63 to 2.7. It has become a common building stone, with applications including external flooring and facing and internal flooring.

In comparison to marble, it is more robust, more durable, less porous and needs less maintenance. Granite does not need re-polishing after it has been polished and fixed in the desired place, while marble needs polishing at least once every two years because of the rigid compacted granular shape. The low porosity of granite does not absorb water and is thus commonly used in kitchens, bathrooms, research laboratories and other offices. Aggregates are used in all civil projects, such as houses, bridges, dams, canals, highways and railway tracks. Therefore, the construction industry is dependent on the supply of granite worldwide. Nagar Parker granite has all the engineering characteristics for road construction in compliance with existing standards AASHTO, ASTM and BS due to its negligible low porosity and strong crushing strength. Besides, the increased road construction in Thar District's surrounding areas and quality development to supply aggregates to the local market and reduce long-distance transportation costs. To classify the mineralogy grain, we used both petrologic and petrographic methods in the study which include size, shape, fabric and weathering states.

5. RECOMMENDATIONS

Based on the findings of this study, the following recommendations have been made:

- The aggregate deposit is highly recommended for use in all types of concrete and highway construction.
- Nagar Parker aggregates can be used safely on regular roads and concrete projects.
- The aggregates Dedhvero and Karai are more appropriate for the building.
- The Mokrio, Dinsi and Wadlai aggregates performed well in physical tests to be used in everyday construction.
- More studies should be conducted in the adjacent areas of Nagar Parkar and Tharparkar in light of the present study.

REFERENCES

1. Egesi N, Akaha C. Engineering-Geological Evaluation of Rock Materials from Bansara, Bamenda Massif Southeastern Nigeria, as Aggregates for Pavement Construction. Geosciences. 2012;2(5):107–11.
2. Aggregates for Concrete De-veloped by ACI Committee E-701. ACI Education Bulletin. 2007;
3. Neville AM. Properties of Concrete. 4 th ed. Pearson Education Asia Pte. Ltd; 2000.
4. Lopez-Buendia AM, Climent V, Verdu P. Lithologi-cal Influence of Aggregate in the Alkali- Carbonate Reactio. Cement and Concrete Research, Elsevier Ltd Valencia, Spain. 2006;36.
5. Bilqees R, Shah MT. Industrial Applications of Limestone Deposits of Kohat, NWFP-a research towards the sustainability of the deposit. Pakistan Journal of Scientific and Industrial Research. 2007;50(5):293–8.
6. Jan N, Bilqees R, Riaz M, Noor S, Younas M. Study of Limestone of Nizampur area for Industrial utilization. Jour Chemical Society of Pakistan. 2009;(1):16–20.
7. Bilqees R, Khan T, Pirzada N, Abbas SM. Indus-trial applications of Abbottabad limestone: Utilizing its chemi-cal and engineering properties. Journal of Himalayan Earth Sciences. 2012;45(1):91–6.
8. Haq IU. To Investigate the Suitability of Coarse Aggregate Sources Available in Districts Nowshera for use in ordinary structural concrete. Peshawar; 2012.

9. Gondal MI, Ahsan N, Javaid AZ. Engineering properties of potential aggregate resources from eastern and central Salt Range, Pakistan. Pakistan Geological Bulletin of Punjab University. 2009;44:97–103.
10. Ghaffar A, Siddiqi ZA, Ahmed K. Assessing Suita-bility of Margalla Crush for Ultra High Strength Concrete. Pa-kistan Journal of Engineering and Applied Science. 2010;7:38–46.
11. Iqbal MM, Chaudhry MN, Khan ZK. Allai Aggregate f or Rehibilitation And Reconstruc-tion of October 8. Bulletin Punjab University. 2005;44:43–54
12. Kevern JT, Wang K, Schaefer VR. Effect of coarse aggregate on the freeze-thaw durability of pervious concrete. J Mater Civ Eng [Internet]. 2010;22(5):469–75. Available from: http://dx.doi.org/10.1061/(asce)mt.1943-5533.0000049
13. Fookes PG, Gourley CS, Ohikere C. Rock weathering in engineering time. q j eng geol [Internet]. 1988;21(1):33–57. Available from: http://dx.doi.org/10.1144/gsl.qjeg.1988.021.01.03
14. Ahsan N, Baloch IH, Chaudhry MN, Majid CM. Strength Evaluation of Blends of Lawrencepur, Chenab and Ravi Sands with Lockhart and Margala Hill Limestones for use in Concrete, Special Issue Pakistan Museum of Natural History. Pakistan Science Foundation. 2000; 213–40.
15. Jan MQ, Agheem MH, Laghari A, Anjum S. Geology and petrography of the Nagar Parkar igneous complex, southeastern Sindh: the Dinsi body. J Himal Earth Sci. 2014;47:1–14.
16. Kazmi AH, Khan RA. The report on the geology, min-erals and water resources of Nagar Parkar, Pakistan. Pakistan Geol Surv Pak Inf Release. 1973;64:1–32.
17. Jan MQ, Laghari A, Khan MA. Petrography of Nagar Parkar igneous complex, Tharparkar, southeast Sindh. Geol Bull Univ Peshawar. 1997;30:227–49.
18. Muslim M, Akhtar T, Khan ZM, Khan T. Geology of Nagar Parkar area, Thar Parkar district, Sindh, Pakistan. Geol Surv Pak Inf Release. 1997;605:1–21.
19. Ahmad SM, Chaudhry MN. A-type granites from the Nagarparker complex, Pakistan: geochemistry and origin. Geol Bull Punjab Univ. 2008;43:69–81.
20. Khan T, Murata M, Rehman HU, Zafar M, Ozawa H. Nagarparker granites showing Rodinia remnants in the southeastern part of Pakistan. J Asian Earth Sci [Internet]. 2012;59:39–51. Available from: http://dx.doi.org/10.1016/j.jseaes.2012.05.028
21. Pathan MT, Rais A. Preliminary report of the investiga-tion of Nagarparkar igneous complex. Sindh Univ J Sci. 1975;1:93–7.
22. Ahmad SM, Chaudhry MN. Geochemical characteriza-tion and origin of the Karai-gabbro from the Neoproterozoic Nagarparker complex, Pakistan. Pakistan Geol Bull Punjab Univ. 2007;42:1–14.
23. Ahsan SN, Firdous R, Mastoi AS, Ghuryani S. Dhedvero iron oxide-gold±copper prospect, a preliminary evaluation, Nagar Parker Taluka, Thar Parker District, Sindh, Pakistan. Geol Surv Pak Inf Release. 2008;872:1–29.
24. Standard Practice for Sampling Aggregates. ASTM. 2014;
25. Standard Test Method for Bulk Density (unit weight) and voids of aggregates. 2016.
26. Standard Terminology Relating to Con-crete and Concrete Aggregates. 2016.
27. Standard Test Method for Relative Density (Specific Gravity), and Absorption of Coarse Aggregates. ASTM. 2015;127.
28. Shetty MS, Jain AK. Concrete Technology (Theory and Practice), 8e. S. Chand Publishing; 2019.
29. Methods for Determination of Flakiness & Elongation Index. 1990.
30. Standard Test Method for Soundness of Aggregates by using Sodium Sulphate or Magnesium Sul-phate. ASTM C. 2013;88.
31. Testing aggregates - Part 110: Methods for determination of aggregate crushing value (ACV). BSI British Standards Institution BS. 1990;812.
32. British Standard Institute BS 812-112: 1990. Methods for determination of aggregate impact value. AIV;
33. Standard Test Method for Resistance to Degradation of Small Size Coarse Aggregates by Abrasion and Impact in the Los Angeles Machine. 2013.
34. Standard Test Method for Potential Alkali Silica Reactivity (Chemical Method). 2006.
35. Standard Test Method for Petrographic Examination of Aggregates for Concrete. 2012.
36. Khan T, Murata M, Zafar M, Hafiz Ur Rehman &. Origin of the mafic dykes in Na-garparker area of Pakistan. Arab Journal Geosciences. 2015;8:6095–104.
37. Raza HA, Ahmed R, Ali SM, Sheikh AM, Shafique NA. Exploration performance in sedimentary zones of Pakistan. Pakistan Journal of Hydrocarbon Research. 1989;1:1–7.
38. Biswas SK. A review of structure and tectonics of Kutch basin, western India with special reference to earth-quakes. Current Science. 2005;88:1592–600.
39. Muhammad A. Preliminary econom-ic evaluation of granite deposits of Nagarparkar, district Tharparkar. Vol. 861. Sindh, Pakistan; 2007.

Acknowledgments: We acknowledge the Centre for Pure and Applied Geology, University of Sindh, Jamshoro and the Mehran University of Engineering Science and Technology Jamshoro, for the use of the laboratory support. The researchers also acknowledge Mr. Ali Akbar Kalwar (Assistant Executive Engineer at small dams Nagar Parker) Irrigation Department, Government of Sindh Pakistan, for providing vehicles and facilities during the fieldwork.

Muzafar Ali Kalwar: https://orcid.org/0000-0001-5837-6156

DOI 10.2478/ama-2023-0018

MICROBIOLOGY OF METALWORKING FLUIDS: WHAT WE KNOW AND LESSONS TO BE LEARNT

Thomas KOCH*

*Dr. Thomas Koch Industrie Beratung, Isarstraße 95, Bremen 28199, Germany

tkoch@uni-bremen.de

received 12 September 2022, revised 17 October 2022, accepted 24 October 2022

Abstract: Water-miscible metalworking fluids are an essential component of many manufacturing processes. During their lifetime they are subject to permanent changes in their physical and chemical characteristics. Due to their high content of water and their chemical composition in use, metalworking fluids (MWF) are prone to microbial life, i.e. the proliferation of bacteria and fungi. The microbial activity leads to significant changes in the chemical composition of the MWF, which can result in the loss of their technical properties. This paper briefly discusses the influences of microbial contamination on the technical quality of MWF and presents common monitoring systems for the detection of microorganisms. Finally, measures are described that can be taken to protect MWF from damage caused by high microbial loads in daily practice. In a short outlook, alternative research approaches are mentioned that aim at sustainable use of MWF.

Key words: Metalworking fluids, microbiology, biofilms, monitoring

1. INTRODUCTION: METALWORKING FLUIDS (MWF)

The use of MWF in metal cutting and forming processes has been state of the art for decades. In machining processes, a large proportion of the machine power input is converted into heat, which can have a detrimental effect on the machined workpiece and the tools. MWF counteract this process by dissipating the heat energy directly at the point of machining and/or reducing the generation of heat by reducing friction (s. Fig. 1). The use of MWF generally leads to better machining results and ensures cost-saving production.

The elementary role that machining and MWF play in every-day life was described by Hans Ernst as early as 1951 as follows: 'Directly or indirectly, it (metal cutting) affects every aspect of our civilization. Every product we use, wear, or eat is related to metal cutting, either directly, in its own manufacture, or indirectly, through the manufacture of the machine that makes it' [1].

For the user, the longest possible service life of the MWF with constant performance is of great relevance. Drag-out losses due to chips and scooping parts require subsequent redosing with MWF. A decrease in technical quality due to physical-chemical processes and in particular microbial activity in water-mixed MWF leads to changes in MWF chemistry and a decline in the working result. Therefore, proper monitoring and surveillance of the MWF are necessary. An optimally coordinated use of MWF can lead to cost savings irrespective of the size of the plant and equipment. In addition, extended change intervals and reduced waste quantities also contribute to environmentally conscious production. The use of water-mixed lubricants thus contributes significantly to the cost structure in a production chain [2, 3].

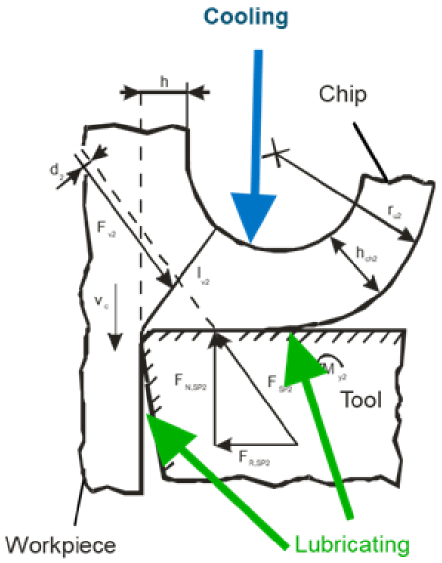


Fig. 1. Primary functions of cooling and lubricating of a MWF. The arrows indicate the areas in which the two functions are effective

2. MWF' TASKS AND CHEMISTRY

According to DIN51385, coolant lubricants are divided into three groups: non-water-miscible, water-miscible and water-mixed coolant lubricants. The water-mixed MWF designates the application state of the water-miscible MWF after mixing a MWF concentrate with 90–95% water. Mineral oils, hydrocrack oils, polyolefins and synthetic esters of various origins with different degrees of refinement are used as base oils in these technical fluids. In their application state, MWFs are divided into emulsions and solutions. To convert the two phases of water and oil from the concentrate into a stable emulsion, surface-active substances known as emulsifiers or surfactants are added to the concentrate. These have an amphiphilic molecular structure, i.e. they consist of a nonpolar, hydrophobic part and a polar, hydrophilic part. They reduce the interfacial tension between the two phases by dissolving with their non-polar part in the oil and their polar part in the water. As a result, they form finely dispersed droplets, known as micelles, with the oil. The stability of the emulsion depends on the size of the micelles that form. They represent the degree of dispersity of the emulsion. The average droplet size of a MWF emulsion is between 0.1 μm and 10 μm [4]. Finely dispersed emulsions have smaller micelles and are more stable. MWF emulsions can be demulsified by lowering the pH value, salinisation, evaporation, or energy input using ultrasound. MWF solutions are true solutions in the chemical sense, in which the ingredients are uniformly dissolved in water. They are homogeneous mixtures of inorganic and/or organic substances with water [5–7].

High demands are placed on the machining process when cutting metals. Here, MWF has the task of reducing friction and wear at the contact points of the tool and workpiece, as well as dissipating the developing heat. Due to the use of MWF, changes in friction mechanisms and reduced wear occur, and the quality of the newly created surface in the machining process can be significantly improved. The occurrence of lower cutting forces enables the increase of process performance. Fig. 1 displays the areas of cooling and lubricating in a metal-cutting process. Gottwein et al. [8] wrote a fundamental work on whose content almost all subsequent work was based. A detailed description of the MWF ingredients will not be given here; this has already been done in detail in [9–11]. The turnover rate of in-use MWFs in Germany is about 600.000 tons p.a. [12].

In addition to the primary functions of cooling, lubrication and transport of chips and swarf there are further relevant functions of MWF. Depending on the machining process, there are secondary requirements, such as corrosion protection, oxidation stability and a low tendency to foam formation. Due to the high demands placed on the machining process in manufacturing today, the fields that MWF must cover have expanded significantly. In addition to other literature, in Germany, the rule DGUV R-109-003 of the Employer's Liability Insurance Association defines the requirements that an effective MWF should meet today. The following is a selection of MWF secondary requirements [6, 11, 13]:

- Cooling capacity
- Lubricity
- Flushing capacity
- Long service lifetime
- Corrosion protection
- Anti-foaming properties
- Skin compatibility
- Low hazardous potential
- Environmental friendliness
- Wetting capacity
- Compatibility with other machine tool components
- Low disposal costs

Due to these requirements, a MWF can be formulated with up to 30 different ingredients, and >300 raw materials are available for the formulation of MWFs. The degree of purity of the raw materials used is of technical quality, so in any case admixtures of unknown origin and concentration must be expected [9, 10, 14].

3. MICROORGANISMS AND MWF

Microorganisms are ubiquitously present. In MWF bacteria and fungi are the main microbial inhabitants. Microorganisms have special abilities. They represent the oldest form of life, they are involved in all natural cycles, e.g. C-, N-, and they are adapted to live in extreme environments e.g. temperatures >100°C, pressures >100 bars, pH-values 0-14. Bacterial cells usually have a size between about 0.5 μm and about 5 μm, and the single cells of fungi can reach sizes of up to 20 μm. The existence of bacteria and fungi in MWF however, has been described since the early 1920s. The first studies on the interactions between microbial contamination, hygienic aspects and the quality of the MWF date back to the 1920s [15–18]. The significance of the microbial infestation of MWF derives from the microorganism's capability to use the ingredients of a MWF for their metabolism. The degradation led to a decline in the components and consequently altered the technical quality of the MWF [19, 20].

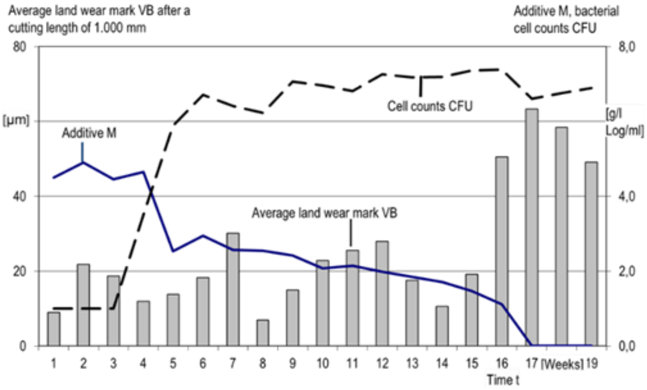


Fig. 2. Course of average tool wear, bacterial cell counts and a selected additive in a drilling process over a run time of 19 weeks

All technical fluids containing water are susceptible to microbial infestation. Due to their water content of >90% MWF are prone to microbial infection. Microorganisms find very favourable conditions for their growth and cell reproduction in MWF. They utilise all biological usable components of the MWF as a source of food and energy. The temperatures in the cooling system ranging from 20°C to 45°C also ensure optimum conditions for mesophilic microorganisms. In the course of the degradation processes, changes occur in the chemical composition of the MWF and the structure of ingredients. This results in a loss of the desired technical properties of the MWF. The microbial utilisation of MWF ingredients is accompanied by the formation of new, unknown intermediates and end products and a drop in the pH value. The microbial degradation of MWF ingredients as emulsifiers, corrosion inhibitors and/or performance additives causes increased wear and corrosion of tools and workpieces as well as the formation of foul odours and discolouration of the MWF. The increase in average land wear mark in a drilling process due to microbial degradation of ingredients (e.g. additive M) and the increase in bacterial cell counts in a MWF is given in Fig. 2. The process parameters were a drilling process of 16MnCr5, tool had a diameter of 10.0 mm, land wear mark VB was determined after a cutting length of 1.000 mm, MWF was a 5% mineral oil-based emulsion, bacterial cell counts detected by colony forming units

(CFU). Microbial infestation of MWF leads to increased consumption of concentrate, biocides and additives, as well as a shift in the chemical composition. The microbial degradation of a MWF is complex and does not affect a single ingredient or property alone. Therefore, the initial condition of the MWF cannot be restored by the addition of additives. The damaged MWF must be replaced and disposed of prematurely. This results in costs for the user that can be avoided. Careful selection, proper care and maintenance, and proper personal protection during use can significantly improve productivity and worker safety [9, 14, 19, 20].

In summary, the microbial degradation processes can have the following effects on the MWF system [13, 21]:

1. Effects on the MWF
   - Degradation of the ingredients
   - Decrease in the pH-value
   - Decrease in corrosion protection
2. Effects on the circulation system
   - Oil separation, emulsion splitting
   - Clogging of lines by biofilms
   - Occurrence of foam
   - Hydraulic difficulties during filtration
3. Effects on the environment
   - Risk of infection for employees
   - Accumulation of skin irritation
   - Occurrence of odours
4. Effects on downstream systems
   - Interference with ultrafiltration during disposal of cooling lubricants
   - Carryover into downstream systems
   - Increase in dissolved organic carbon/chemical oxygen demand (DOC/CSB) in the wastewater disposal stream

In the production process, the aim should be to achieve the lowest possible microbial load in MWF to keep the technical quality of the MWF at a high level.

A particular problem of contamination of MWF is the development of biofilms present in the system. The term biofilms refers to the extensive growth of bacteria and fungi on interfaces. Biofilms are found on almost all interfaces that are in contact with water. On clean surfaces in contact with an aqueous phase, a conditioning film of macromolecules forms spontaneously within a few minutes, on which bacteria grow to form a biofilm. According to Costerton et al. [23] bacterial biofilms represent the most successful form of life regarding the entire biomass as well as the diversity, type and expansion of populated habitats.

Life in biofilms offers microorganisms a series of advantages over existence as free-floating organisms in the water zone. The biofilm has to be understood as a three-dimensional habitat that is neither shaped evenly in a temporal nor in a spatial sense. It presents its inhabitants with protection from hydraulic loads, fluctuation in pH, osmotic stress, dehydration and pollutants such as biocides. At the same time, the biofilm is in a way decoupled from the water body. That makes it impossible to read from data from the water body where biofilms are, what extent they have and which organisms live within them. The extracellular polymeric substance (EPS) is the base frame for biofilms; it is traversed by water-filled channels in which signalling and messenger substances and nutrients are transported. It serves as a storage space for nutrients and is the compartment for chemical-physical reaction processes. The coexistence of different species of microorganisms forms a pool of genetic information in a narrowly defined space leading to the development of micro-consortia. This results in the possibility of horizontal gene transfer and the ability to degrade molecules in co-metabolic processes, which cannot be degraded by single microbes. With increasing thickness of the biofilm and decreasing oxygen content, more ecological niches emerge for anaerobic organisms. These local habitats do not emerge in the free water zone, as anaerobic bacteria could not be active physiologically there. Fig. 3 displays a schematic of a biofilm structure [23–27].

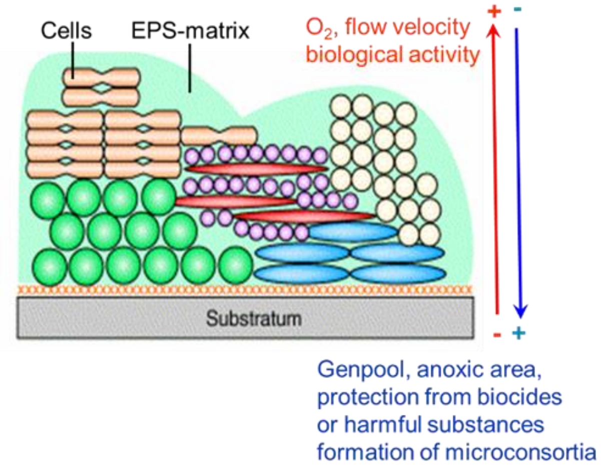


Fig. 3. Schematic representation of a biofilm and selected gradients over the depth of the biofilm. Natural biofilms usually consist of >90% water and in the EPS matrix, the proportion of microorganisms is very low [acc. [28] modified]

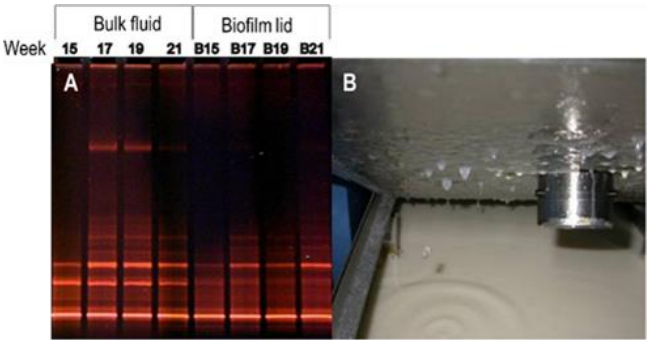


Fig. 4. DGGE-analysis of the bacterial community in Biofilms in a MWF system, A = Plot of the DGGE-analysis, B = Backside of the lid containing a visible biofilm dripping in the bulk fluid

Biofilms ranging in thickness from a few micrometres to a few centimetres are regularly found in MWF systems. In a MWF system, the microbial community of a biofilm is in exchange with the bulk liquid in terms of the predominant species spectrum. Fig. 4 shows the analysis of denaturing gradient gel electrophoresis (DGGE) of the bulk fluid and the biofilm on the backside of a lid in a MWF system, and the testing interval was six weeks. Every single bar in the lanes of the gel plot represents a specific bacterial strain. It is clearly visible that the species composition is nearly identical in the bulk fluid and the biofilm. In pipelines and filter systems, the reduction of the cross section or the complete blocking of the lines can lead to hydraulic difficulties in the MWF system. Due to their good adhesion and persistence to chemicals, biofilms can only be removed from a MWF system by a combination of system cleaning and mechanical post-treatment of the surfaces [25].

4. MONITORING SYSTEMS

In Germany, the monitoring of MWF is regulated in DGUV-R-109-003. Once a week the noticeable alterations in the MWF, pH value, MWF concentration, and nitrite content of the MWF have to be measured and documented. The measurement of the microbial load in MWF is not provided. In case of an increased microbial occurrence in the MWF, reference is made to the rule DGUV-I 051 (BGI762) and the documents available in the appendix of the rule [13, 31].

Due to the high relevance of microbial loads in MWF, this chapter will give an overview of the common techniques for the detection of bacteria and fungi in MWF displaying their advantages and disadvantages. It should be noted that all methods described below evaluate the cell counts in the bulk fluid. Cell count determination in microbial biofilms is more complex and is not performed in MWF monitoring practice.

4.1. Dip-Slides

Dip-Slides are still the most widely used method in operational practice to monitor microbial cell counts in MWF. They are plastic strips coated with a growth media for the identification of bacteria and fungi/yeasts on each side respectively. For cell enumeration, the strips are dipped into the MWF to be wetted with the fluid. Afterwards the strips are sealed in a tube. After an incubation time of mostly 48 h for bacterial and 72 h for fungi/yeast growth respectively at a given temperature, the slides can be evaluated. The addition of a dye to the medium serves the colouring of the bacteria colonies and is supposed to simplify the reading. The classification of the results is conducted with a colour scale supplied by the manufacturer, which displays the cell number estimation in intervals of orders of magnitude (s. Fig. 5).

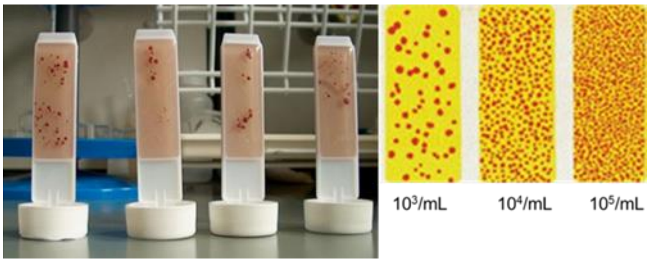


Fig. 5. Left: Dip-Slides after 48h incubation time. MWF were taken at different sample sites of the machine tool. The photograph shows the side with the media for bacterial growth. The red dots are coloured colonies that have arisen from a single cell. Right: Example of a colour scale for the cell count enumeration of the Dip-Slides

Unfortunately, the results of the cell count estimation via Dip-Slides are overrated in operational practice, and a nonexistent growth is interpreted as sterility. The results are more an estimation of cell counts than a reliable determination due to the following inaccuracies:

- The volume of MWF remaining on the Dip-Slide after dipping is dependent on viscosity and temperature and is therefore not constant
- By dipping the slide directly into the MWF, the oily surface film is brought onto the Dip-Slide
- Bacterial strains that are not colourised by the dye are often overlooked in the reading
- Bacterial strains with a low rate of cell division do not form colonies large enough to be identified by eye inspection
- The method is time-consuming (48—72 h)

The main advantages of Dip-Slides are their low price and their easy handling, which does not require a laboratory [32].

4.2. CFU/plate counts

The culturing of microorganisms on solid media has been a proven technique to determine microbial cell counts for decades. The reproductive rate of microorganisms is used to make them visible to the human eye and therefore accessible for quantification. A defined volume (normally 0.1 mL) of fluid/MWF out of a dilution series is applied to a solid medium and evenly spread to separate the organisms. A numerical evaluation of the forming colonies takes place after a predetermined incubation period of normally between 24 h and 48 h. Fig. 6 shows as an example a section of a CFU plate with bacterial cultures of different sizes. The result is represented in cell count per mL. The determination of the CFU offers the possibility to cultivate bacterial strains selectively by using specific media and thus differentiate species within bacteria populations. Single strains or bacteria groups from one mixed culture can be isolated and verified selectively. The determination of the cell counts employing CFU requires laboratory equipment and requires qualified personnel [33, 34]. Therefore, this technique is mainly used in research on MWF microbiology. The advantages and disadvantages of using CFU for the determination of microbiology in MWF can be summarised as follows:

- Lab required, therefore costs are higher
- Reliable results when using the same cultivation media in routine control
- Selective growth conditions can be set up
- Routine required for reliable assessment
- Time consuming (>48 h)

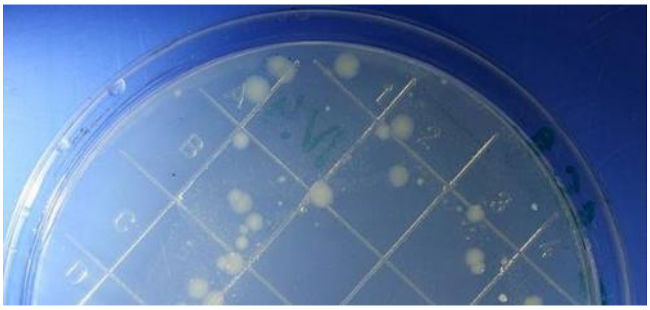


Fig. 6. Section of a CFU plate incubated with MWF. The dots are bacterial colonies that have grown from a single cell. The different sizes of the colonies are due to the different growth rates of the respective bacterial strains and indicates the presence of different species in the sample

4.3. Adenosine triphosphate (ATP)-measurement

ATP is the overall energy supply used in cells, which is traceable in all living beings. The measurement of the ATP content is therefore an appropriate method for the detection and monitoring of microbial loads in MWF. The basis of the ATP determination is the reaction of luciferin with the enzyme luciferase. In the presence of ATP and oxygen, the luciferin-luciferase-complex is oxidised by the emission of light (s. Fig. 7).

$LL + ATP + O_2\ (Mg^{2+}) \rightarrow$
$LL\ (oxid.) + AMP + CO_2 + PP_i + Light$
LL = luciferin-luciferase-complex; AMP = adenosine monophosphate; PP_i = inorganic phosphate (acc. [35]).

In the detection process according to ASTM E2694 [36] (American Society for Testing and Materials) at first, the microbial cells were disrupted, and the released ATP is bound to a filter. After a washing step, the ATP is resolved and can react with the added luciferin-luciferase complex. The emitted light is detected and from these data, microorganism equivalents can be calculated. Depending on the ATP measuring device in use, the time needed for an ATP content measurement is between 5 min and 20 min. Thus, it offers fast results and allows intervention in the MWF system if necessary. It can be carried out without any special laboratory equipment. Furthermore, the method is well suited to show changes in the microbial load in monitoring over time. If changes in the measured values occur the process can be flanked by further methods (e.g. determination CFU). Passman et al. [37] confirm this statement and emphasise the high reproducibility of the measurement results. They conclude that ATP measurement is a powerful tool to improve microbial monitoring in MWF.

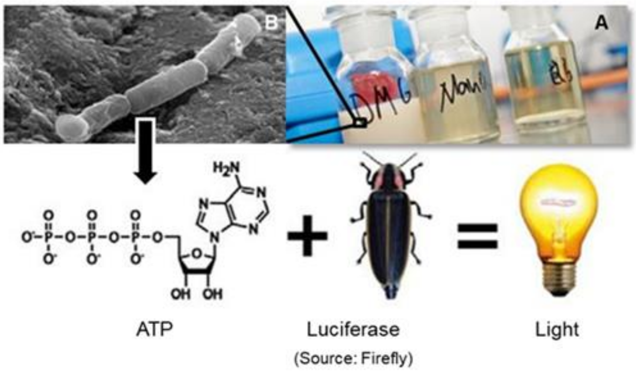
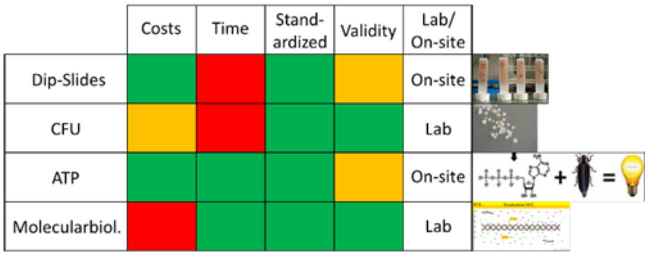


Fig. 7. Scheme for the measurement of adenosine triphosphate (ATP). Upper part: A = Microbial contaminated MWF samples. B = Bacterial cells containing ATP. Below: Reaction of the extracted ATP with the luciferin-luciferase-complex from firefly lead a light emission

4.4. Molecular biological techniques

The permanent development of chemical-analytical methods has allowed for new procedures for the fast determination and exact identification of microorganisms. Molecular biological methods, such as the amplification of DNA with the polymerase chain reaction (PCR), are standard in daily work in microbiological laboratories. Based on this, further methods have been established in the most scientific reflection on MWFs microbiology. The PCR technique requires laboratory and qualified personnel and therefore have currently no role in daily MWF surveillance. But they have been contributing significantly to the scientific understanding of the microbiology of MWF. Developments such as Lab on Chip, which are already state of the art in other fields of application, will certainly also find their way into MWF monitoring shortly and contribute to fast and reliable test results.

Fig. 8 compares the methods presented here for their characteristics in terms of cost, the time required between sampling and result presentation, routine, validity and laboratory requirements. The methods Dip-Slides and ATP measurement are, due to the easy handling and the possibility to be used on-site, the methods that have the widest spread in practice.



	Costs	Time	Standardized	Validity	Lab/ On-site
Dip-Slides	🟩	🟥	🟩	🟨	On-site
CFU	🟨	🟥	🟩	🟩	Lab
ATP	🟩	🟩	🟩	🟨	On-site
Molecularbiol.	🟥	🟩	🟩	🟩	Lab

Fig. 8. Comparison of the presented methods for the determination of microbial cell counts in MWF. The evaluation of the characteristics costs, the time between sampling and result presentation, standardisation, validation and necessity of a laboratory is done by a traffic light system

5. METHODS FOR COUNTERACTING

This section briefly presents physical methods and chemical counteracting that exert a biocidal effect in MWF and have achieved relevance in practice.

Physical methods:

UV-Radiation in the wavelength 254 nm destroys chemical bonds in the microbial DNA. This method can only be used in bypass, as the penetration depth is low and in addition, shadowing effects occur.

Ultrasonic treatment is a method that is mainly used in the food industry to get liquids free of microbes. When used in MWF, there is a risk that the energy input will destroy emulsion droplets and macromolecules.

Contact catalysts as e.g. AGXX® disrupt the electric field of a cell, causing damage to the cell membrane that can lead to cell death. It has a limited range and produces the effect only in passing liquids.

The physical methods mentioned here, only act in the bulk fluid and have no biocidal effect on biofilms.

Chemical counteracting:

Common substance classes of chemical agents are formaldehyde releasing agents, isothiazolinone, chlorocresol and iodcarbamate. The main acting site of chemical agents is also within the bulk fluid. In biofilms, the impact of chemical biocides is decreased due to reactions of the substances with the EPS matrix and a significantly lower flow velocity. With the entry into force of the Globally Harmonised System GHS regulation, the number of chemicals available on the market that may be used as biostatics or biocides in the EU has been significantly reduced. Currently, 11 substances are approved for storage preservation and 6 substances as protective agents for machining fluids [24, 29, 30].

6. RESEARCH RESULTS

Microbial life in MWF is dynamic regarding the cell counts and the diversity of the microorganisms. Fig. 9 displays the development of microbial cell counts in a MWF system detected over a period of 27 weeks. CFU plate counts, ATP measurement according to ASTM E2694, and Dip-Slides were used as techniques for the enumeration of the bacterial cell counts. Dip-Slides lead to an underevaluation in the range of one to more than two orders of magnitude compared to CFU and ATP measurement. Therefore Dip-Slides could give only a hint of bacterial life in MWF and should not be used for a reliable enumeration of cell counts. For the description of the changes in bacterial diversity patterns over time, the dominating strains on CFU plates were identified by molecular biological techniques. The identified bacterial strains are displayed in the text below Fig. 5. Pseudomonas pseudoalcaligenes was the strain which could be detected in the longest time range in the MWF, starting two weeks after the start until the end of the investigation time of 27 weeks. The changes in diversity are due to the varying rate of degradation of the MWF ingredients and thus the varying nutrient supply for the bacteria.

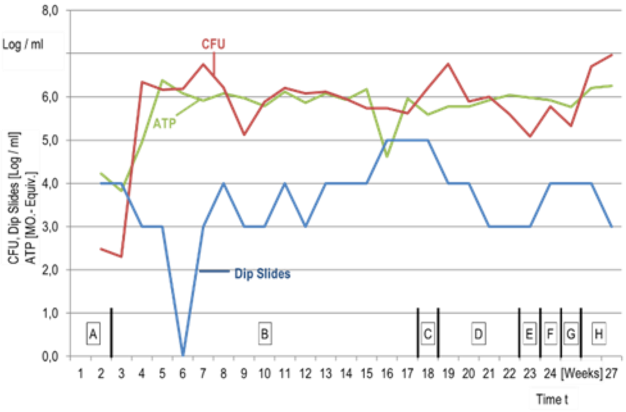
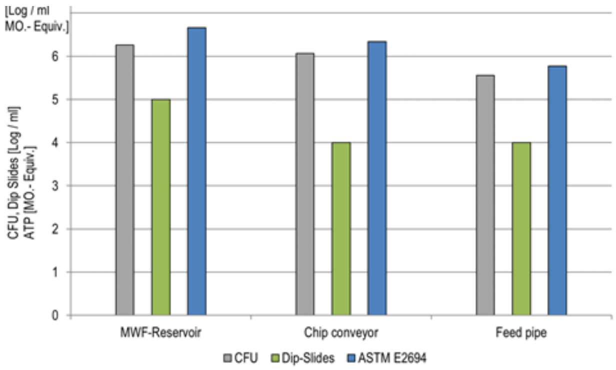


Fig. 9. The graph displays microbial cell counts in a MWF over a time period of 27 weeks, detected by CFU, Dip-Slides and ATP measurement according to ASTM E2694. The underestimation of cell counts by Dip-Slides is evident. The dynamic in diversity is given by the names of the predominant bacterial species identified per measurement date from the CFU plates.
A = Brevundimonas diminuta, Ochrobactrum rhizospherum, unidentified strain; B = Pseudomonas pseudoalcaligenes; C = Arthroabcter sulfureus, Ps. pseudoalcaligenes, unid. strain; D = Sphingomonas yanoikuyae, Ps. pseudoalcaligenes, unid. strain; E = Acinetobacter sp., Sp. yanoikuyae, Ps. pseudoalcaligenes, Brev. diminuta; F = Sp. yanoikuyae, Ps. pseudoalcaligenes; G = Sp. yanoikuyae, Ps. pseudoalcaligenes, unid. strain; H = Sp. yanoikuyae, Ps. pseudoalcaligenes. (acc. [38] modified)

The bacterial diversity also varies in one MWF at the same time in one machine tool at different places. In Fig. 10 the results of measuring the microbial cell counts using the methods CFU, Dip-Slides and ATP-measurement according to ASTM E2694 are compared in the graph. In addition, the dominating strains on CFU plates were identified by molecular biological techniques. These results are given in the table above the graph. The samples were taken at the same time from different sample sites in the machine tool. From these results, the underestimation using Dip-Slides compared to the other techniques can be seen. In addition, the cell counts as well as the abundance of the bacterial species are different, depending on the site samples were taken. For MWF surveillance in daily practice, this means that routines concerning the sampling site are necessary to achieve valid results.



	Abundance		
Bacterial species	MWF-Reservoir	Chip conveyor	Feed pipe
Sphingomonas yanoikuyae	+++	++	++
Pseudomonas pseudoalcaligenes	+	++	++
Brevundimonas diminuta	J.	J.	+
Stenotrophomonas maltophilia	J.	J.	+

Fig. 10. Comparison of the results of measurement of microbial cell counts by CFU, Dip-Slides and ATP measurement according to ASTM E2694. The samples were taken from one machine tool at different sample sites. Cell counts differ, depending on the sample site and method used. The bacterial diversity differs concerning the sample site. (acc. [38] modified)

7. CONCLUSIONS

The life and proliferation of microorganisms are an integral part of almost all aqueous technical fluids. The degradation of fluid ingredients by microbial activity led to a loss of the required technical properties of the fluid, hygienic problems and, in the case of biofilm formation, disturbances concerning fluidic and thermal processes.

The microbiology of MWF has a history of decades, with only little evolution regarding monitoring and prevention techniques. Nevertheless, it still represents a relevant problem in manufacturing technology today, which should be brought back into focus. Due to legal regulations, fewer biocidal agents are available for use as preservation of MWF. This results in improved health protection for employees, but also in the problem of an increasing spread of microorganisms in MWF.

The following points should be realised and implemented in practice and research to enable an extended service life and thus a resource-saving use of MWF in the future:

- Microbial growth in MWF and other water containing technical fluids cannot be avoided.
- Besides the regular examinations of the physical-chemical parameters, the determination of microbial cell counts is necessary for successful monitoring and surveillance of the MWF.
- Molecular biological surveillance techniques should be state of the art in MWF surveillance.
- The result of an individual measurement of the microbial cell counts is not meaningful, a time course of measurements must always be recorded.
- Only suitable methods for cell count determination should be used.
- Compliance with industrial hygiene and machine cleaning standards must be maintained.
- The development of new, alternative biocidal agents or technical processes should be the focus of scientific research.

- The knowledge about biofilms in MWF systems is only in its infancy and also requires more intensive research.
- Legislation such as Biocidal Products Regulation (BPR) and REACh/GHS can be a driver for new developments regarding new biocidal products with a lower hazardous potential.
- Research is needed for the development of alternative and renewable ingredients, e.g. plant-based emulsifiers, antioxidants, performance additives or biocidal substances.

REFERENCES

1. Ernst H. Fundamental Aspects of Metal cutting and Cutting Fluid Action. Annals of the New York Academy of Sciences. 1951, 53: 936-961.
2. Kiechle A. Kostenanalyse beim Einsatz von Kühlschmierstoffen. 11th International Colloquium Industrial and Automotive Lubrication. 1998 January 13th-15th, Stuttgart Ostfildern.
3. Koch T. Kühlschmierstoffe als gleichrangige Systemkomponente. In: Brinksmeier E, Werner K, Klocke F. (eds.) 2nd European Conference on Grinding, 2006 November 16th-17th: 14-1-14-33.
4. Assemheimer C; Domingos A.S, Glasse B, Fritsching U, Guardani R. Long-term monitoring of metalworking fluid emulsion aging using a spectroscopic sensor. Can. J. Chem. Eng. 2017, 95:2341–2349. DOI 10.1002/cjce.22931.
5. DIN 51385:2013-12: Schmierstoffe - Bearbeitungsmedien für die Umformung und Zerspanung von Werkstoffen – Begriffe. Beuth-Verlag, Berlin, 2013.
6. Byers, J.P.: Metalworking Fluids. Third Edition, Boca Raton, Taylor & Francis, CRC Press, 2017.
7. VDI 3035:2008-05 Gestaltung von Werkzeugmaschinen, Fertigungsanlagen und peripheren Einrichtungen für den Einsatz von Kühlschmierstoffen. Beuth-Verlag, Berlin, 2008.
8. Gottwein K, Reichel, W. Kühlschmieren bei wirtschaftlicher Metallbearbeitung einschließlich Kühlmittelrückgewinnung und Werkstückentfettung. München: Carl Hauser Verlag, 1953.
9. Möller U.J, Nasser, J. Schmierstoffe im Betrieb. Springer Verlag Berlin Heidelberg 2. Auflage, 2002.
10. Torbacke M, Rudolphi A.K, Kassfeldt, E. Lubricants Introduction to Properties and Performance, John Wiley & Sons Ltd., 2014.
11. Czichos H, Habig K.H. Tribologie-Handbuch, Springer Vieweg, Wiesbaden, 4. Auflage, 2015.
12. bafa Bundesamt für Wirtschaft und Ausfuhrkontrolle: Amtliche Mineralölstatistiken. https://www.bafa.de/DE/Energie/Rohstoffe/Mineraloelstatistik/mineraloel_node.html (26.01.2022).
13. DGUV Regel 109-003. Tätigkeiten mit Kühlschmierstoffen. Deutsche Gesetzliche Unfallversicherung. 2009.
14. Häusser M., Dicke F, Ippen H. Kühlschmiermittel-Bestandteile und ihre gesundheitliche Wirkung. Zentralblatt Arbeitsmedizin. 1985, 35(6):176-181.
15. Houghton E.F. Causes of Skin Sores and Boils Among Metal Workers: An Investigation by the Houghton Research Staff. Edgar Vaughan & Company, Limited, 1920.
16. Lee M, Chandler A.C. A Study of the Nature, Growth and Control of Bacteria in Cutting Compounds. J. Bacteriol. 1941,41:373-386.
17. McConnell W.J. Dermatosis following the use of cutting oils and lubricating compounds. U.S. Pub. Health Reports. 1922, 37: 1773-1794.
18. Tower A. Dermatitis from cutting oils and compounds. Industrial Medicine. 1938, 7:515-516.
19. Koch T. Auswirkungen des mikrobiellen Befalls von wassergemischten Kühlschmierstoffen auf das Zerspanergebnis (Teil 2). Härterei-Technische Mitteilungen HTM. 2008, 63(1):50-59.
20. Koch T.: Auswirkungen des mikrobiellen Befalls von wassergemischten Kühlschmierstoffen auf das Zerspanergebnis (Teil 3). Härterei-Technische Mitteilungen HTM. 2008, 63(2):115-132.
21. Passman F.J, Küenzi P. Microbiology in water-miscible metalworking fluids. Tribology Transactions. 2020, 63(6):1147-1171. DOI: 10.1080/10402004.2020.1764684 2020
22. Weyandt R.G. Kontrolle des mikrobiellen Befalls wassergemischter Kühlschmierstoffe. Abschlussbericht, Projektinitiator: Süddt. Metall-BG, Mainz, Projektnehmer: Institut Fresenius GmbH, Abt. Umweltbiologie, Taunusstein, 1996
23. Costerton J.W, Lewandowski Z, Caldwell D.E, Korber D.R, Lappin-Scott H.M. Microbial biofilms. Annual Review of Microbiology. 1995, 49:711-745.
24. Cloete T.E. Resistance mechanisms of bacteria to antimicrobial compounds. International Biodeterioration & Biodegradation. 2003 51(4):277-282.
25. Koch T. Biofilms. In: Mang T. (ed.) Encyclopedia of Lubricants and Lubrication, Springer Verlag, Berlin, Heidelberg. 2014:163-168. ISBN 978-3-642-22646-5.
26. Kumar C.G, Anand, S.K. Significance of microbial biofilms in food industry: a review. International Journal of Food Microbiology. 1998, 42(1-2):9-27.
27. Szewzyk U, Szewzyk R. Biofilme – die etwas andere Lebensweise. BIOspektrum. 2003, 9(3):253-255.
28. Rickard A.H, Gilbert P, High N.J, Kolenbrander P.E, Handley P.S. Bacterial coaggregation: an integral process in the development of multi-species biofilms. Trends in Microbiology. 2003, 11(2):94-100.
29. Koch T. Microbial Loads: Counteracting. In: T. Mang (ed.) Encyclopedia of Lubricants and Lubrication, Springer Verlag, Berlin, Heidelberg. 2014:1160-1164. ISBN 978-3-642-22646-5.
30. www.reach-clp-biozidhelpdesk.de/DE/Biozide/Wirkstoffe/Genehmigte-Wirkstoffe/Genehmigte-Wirkstoffe-0.html (30.04.2022)
31. DGUV Information 209-051 (BGI 762). Keimbelastung wassergemischter Kühlschmierstoffe. Deutsche Gesetzliche Unfallversicherung. 2016.
32. Koch T.: Microbial Monitoring. In: T. Mang (ed.) Encyclopedia of Lubricants and Lubrication, Springer Verlag, Berlin, Heidelberg. 2014: 1164-1169. ISBN 978-3-642-22646-5.
33. Bast E. Mikrobiologische Methoden. 2. Auflage Heidelberg: Spektrum Akademischer Verlag 2001.
34. Schlegel H.G. Allgemeine Mikrobiologie. 6. Auflage Georg Thieme Verlag Stuttgart New York. 1985.
35. Webster A.R, Lee J.Y, Deininger R.A. Rapid Assessment of Microbial Hazards in Metalworking Fluids. Journal of Occupational and Environmental Hygiene. 2005, 2(4).213-218.
36. ASTM 26944e11. Standard Test Method for Measurement of Adenosine Triphosphate in Water-Miscible Metalworking Fluids. ASTM International, West Conshohocken. http://dx.doi.org/10.1520/E2694e11. online at www.astm.org
37. Passman F.J, Egger G.L, Hallahan S, Skinner B.W, Deschepper M. Real-Time Testing of Bioburdens in Metalworking Fluids Using Adenosine Triphosphate as a Biomass Indicator. Tribology Transactions. 2009, 52(6):788-792.
38. Koch T, Passman F, Rabenstein A. Comparative Study of Microbiological Monitoring of Water-Miscible Metalworking Fluids. International Biodeterioration & Biodegradation. 2015, 98:19-25.

Thomas Koch: https://orcid.org/0000-0002-5649-9328

sciendo

DOI 10.2478/ama-2023-0019

acta mechanica et automatica, vol.17 no.2 (2023)

Special Issue "Modeling and experimental investigations of thermo-hydraulic, hydrogen, manufacturing and mechanical systems"

INFLUENCE OF VALVE-SEAT ANGLES TO OPERATION VALUES AND EMISSIONS OF MEDIUM-SPEED DIESEL ENGINES

Leander MARQUARDT*, Heiner-Joachim KATKE*
Andreas REINKE*, Niklas KOCKSKÄMPER*

*Fakultät Maschinenbau, Hochschule Stralsund, Zur Schwedenschanze 15, D-18435 Stralsund, Germany

leander.marquardt@hochschule-stralsund.de, heiner.katke@hochschule-stralsund.de
andreas.reinke@hochschule-stralsund.de, niklas.kockskaemper@hochschule-stralsund.de

received 14 September 2022, revised 7 December 2022, accepted 11 December 2022

Abstract: For the development of gas exchange for large diesel engines, a compromise has to be found between efficient valve-flow and the time between overhauls. On the one hand, large effective flow areas, especially during valve-overlap, are demanded. On the other hand, there are limitations of cylinder bore regarding the maximum diameter of inlet and outlet valves and the minimum distance (dead space) between valves and piston, as well as wear-related smaller seat angles. For large medium-speed diesel engines, a valve-seat angle of $\beta = \mathbf{30°}$ for inlet and outlet valves is a standard application. For engine-operation with clean fuels, a valve-seat lubrication (gasoil) or smaller seat angles (natural gas) need to be applied. With this presentation, the basic influence of different valve-seat angles on the operation values and emissions will be considered for the example of the single-cylinder research engine FM16/24. Using a self-developed testbed, experimental investigations into effective flow areas as a function of valve-lift at inlet and outlet valves have to be executed. With this input, different cycle calculations including T/C have to be carried out to determine deviances in specific fuel-oil consumption, exhaust-gas temperatures, NOx emissions and air/fuel ratio. The results will be discussed critically.

Key words: medium-speed engine, gas exchange, valve-seat angle, flow coefficient, thermal load

1. PROJECT MOTIVATION

More than 95 % of the world's merchant ship fleets are powered by diesel engines. Currently about 71 % are burning heavy fuel-oil [1]. The market share of ≈1 % for gas-engines is mainly induced by special boil-off requirements of LNG-carriers. Directly driving 2-stroke engines with speeds up to 104 rev/min dominate the propulsion market for container ships, tankers and bulkers. Cruise liners, ferries, container-feeders and special-purpose vessels are mainly propelled with reduction gears by medium-speed four-stroke engines with engine speeds between 333 rpm and 1,000 rpm. Gensets are mostly medium-speed powered. Intensive competition in the propulsion market led to extreme concentration of manufacturing capacities.

As a result of aggressive competition, the remaining companies increased the specific output of their engines to ensure the resultant cost-attenuation that would enable them to retain themselves in this market. The basic requirement is a sufficient charge-air pressure, limited by the maximum circumferential speed of T/C wheels. Large four-stroke diesel-engines with optimised volumetric efficiency generate with every bar charge-air pressure ≈6 bar mean effective pressure, which are paid by customers, while parts of available charge-air pressure must be used for emission reduction by Miller-timing. Optimised gas exchange is a usable way for breaking up this trade-off between vendable power and limited T/C capacity.

For the development of gas exchange of large diesel engines, a compromise must be found between efficient valve-flow and sufficient time between overhauls of at least 5,000 h for valves. At one side, large effective flow areas, especially during valve-overlap, are demanded for low exhaust-gas temperatures. Otherwise, there are limitations of cylinder bore regarding the maximum diameter of inlet and outlet valves and the minimum distance (dead space) between valves and piston, as well as wear-related smaller seat angles. For large medium-speed diesel engines, a valve-seat angle of β_V = 30° (Fig. 1) for inlet and outlet valves is a standard application.

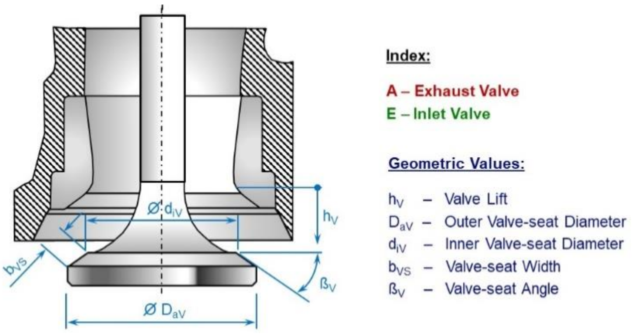


Fig. 1. Flow-relevant geometry of engine-valves

For engine-operation with clean fuels (i.e. all fuels without lubricating effects of combustion deposits, such as distillate fuels or fuel gases), a valve-seat lubrication (gasoil-operated MAN medium-speed) or smaller seat angles (natural-gas fired MaK engines) would be needed.

DOI 10.2478/ama-2023-0019

Based on our study of the literature, we have gauged that there are no publications pertaining to the influence of valve-seat angles on operational values and emissions. By interchanges with development engineers of MaK and MAN, this result was confirmed. According to the authors' experience from industry, no advantages of CFD calculations versus experimental investigations regarding correctness of results and economic issues were found. With special consideration for financial limits and staff-potentials of a small university, it was decided to carry out only experimental investigations and cycle calculations related to this issue.

2. EXPERIMENTAL INVESTIGATIONS

2.1. Investigations at fluid-dynamic testbed

For experimental investigations of the behaviour of gas exchange valves in cylinder heads, a special testbed was designed and assembled (Fig. 2). The supply of compressed air was realised by a dry screw-compressor (for charging of test engines), calmed in a 400 L volume reservoir and measured by a thermal mass flow meter. As long as possible a constant difference-pressure of 50 mbar over the valves was used for measurements at smaller valve-lifts. For a valve-lift larger than $h_V \approx 6$ mm, the flow-capacity of the screw-compressor was not sufficient to keep this difference-pressure. The compressor was operated at maximum speed, and the resultant difference-pressure was considered for these measurements. At the maximum valve-lift of $h_V = 12$ mm, a pressure difference of ≈20 mbar was obtained. According to the measurement described in the literature [2, 3], this pressure difference used does not exercise any influence on the flow coefficients.

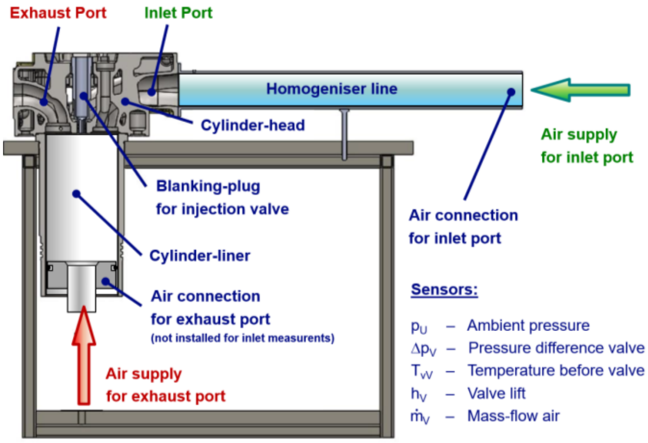


Fig. 2. Testbed for fluid-flow experiments

Every discrete valve-lift was adjusted and measured by an ordinary dial indicator. Pressure difference over valves was detected by a digital manometer. With a pressure calibrator, the ambient pressure was quantified.

This testbed represents valve-flow under steady-state conditions. Conception and design of this testbed for fluid-flow experiments orientated to experiences according to the testbed concepts at MAN Augsburg and MaK Kiel as well as the same model for cycle-calculation based on [4] is used. The results of these measurements had to be comparable to the flow-measurements at serial models to ensure representative results for different families of medium-speed engines. Up to now, no dynamic methods as described in the study of Szpica [5] are used at German manufactures of commercial medium-speed engines.

The jacket of a truncated cone normal to valve-seat was chosen as a reference surface for the geometric flow area (Fig. 3a) [4, 6 and source code of 7 without references to other authors]. Additionally, the significant effective flow area is calculated by multiplication of geometric area and flow coefficient, and the definition of the geometrical value is even. Published in the handbook for mechanical engineering of Beitz and Grote [8] is the alternative use of a cylinder's jacket, erected at the inner valve-seat diameter up to the seat ring. In the study of Heywood [9], a more complicated section-related version with three different cases in the order of valve-lift is published. All three variants are compared in the study of Swiderski [2]. Maximum deviation between the easier description as a truncated cone's jacket with just one equation as used for these investigations [4] and the complex formulas according to Heywood [9] is with 3.7 % at maximum valve lift for the test engine used quite low. A different formula for the same modelling of the geometric flow area as an area of a truncated cone's jacket is given in the study of Tanaka [10, p. 293].

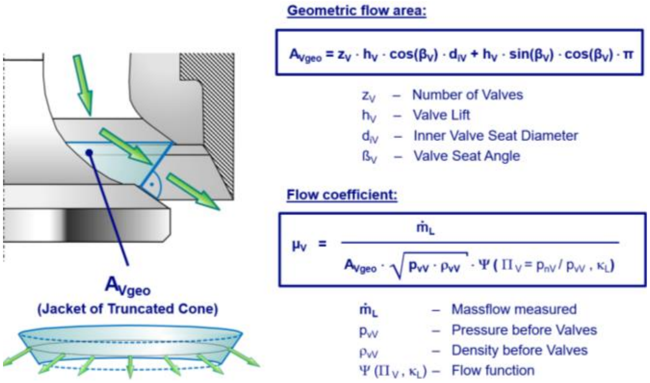

Geometric flow area:
$A_{Vgeo} = z_V \cdot h_V \cdot \cos(\beta_V) \cdot d_{iV} + h_V \cdot \sin(\beta_V) \cdot \cos(\beta_V) \cdot \pi$
z_V – Number of Valves
h_V – Valve Lift
d_{iV} – Inner Valve Seat Diameter
β_V – Valve Seat Angle
Flow coefficient:
$\mu_V = \frac{\dot{m}_L}{A_{Vgeo} \cdot \sqrt{p_{vV} \cdot \rho_{vV}} \cdot \Psi(\Pi_V = p_{nV} / p_{vV}, \kappa_L)}$
A_{Vgeo}
(Jacket of Truncated Cone)
$\dot{m}_L$ – Massflow measured
p_{vV} – Pressure before Valves
ρ_{vV} – Density before Valves
$\Psi(\Pi_V, \kappa_L)$ – Flow function

Fig. 3a. Calculation of flow coefficients

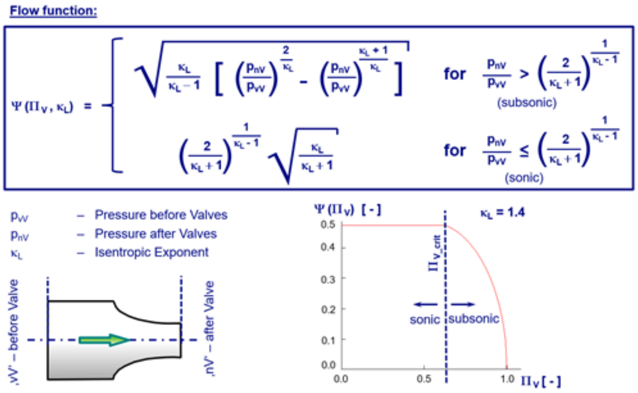


Fig. 3b. Definition of Flow-Function

Pressures before and after valves were calculated from ambient pressure and pressure difference according to the position of pressure sensors and flow direction at valves. Properties of air were calculated with formulas according to Zellbeck [4]. Definition of flow-function (Fig. 3b) was used according to Zellbeck [4] and Urlaub [11] with the fluid-dynamic model of a not-expanded orifice.

A different formulation for flow-function is known from the Japanese manufacturer Niigata.

To ensure realistic values in engine field operation for all related measurements, a used cylinder head, grabbed out after ≈15,000 h at the main engines of SY 'Sea Cloud' (Ex 'Hussar'), was used instead of a new one. Reference measurements for inlet and exhaust were carried out several times to quantify their repeatability.

Diagrammed in Fig. 4 are the results of three measurements for inlet and two for outlet port versus valve-lift up to maximum value in the serial engine. Flow coefficients for zero lift are extrapolated with measured ones for 1 mm and 2 mm. In spite of the difficulties in correct adjusting small lifts at both valves and the resultant relative error in geometric flow area even at small valve lifts the repeatability is with a tolerance of max. 4.5 % at 1 mm lift of inlet valve quite well.

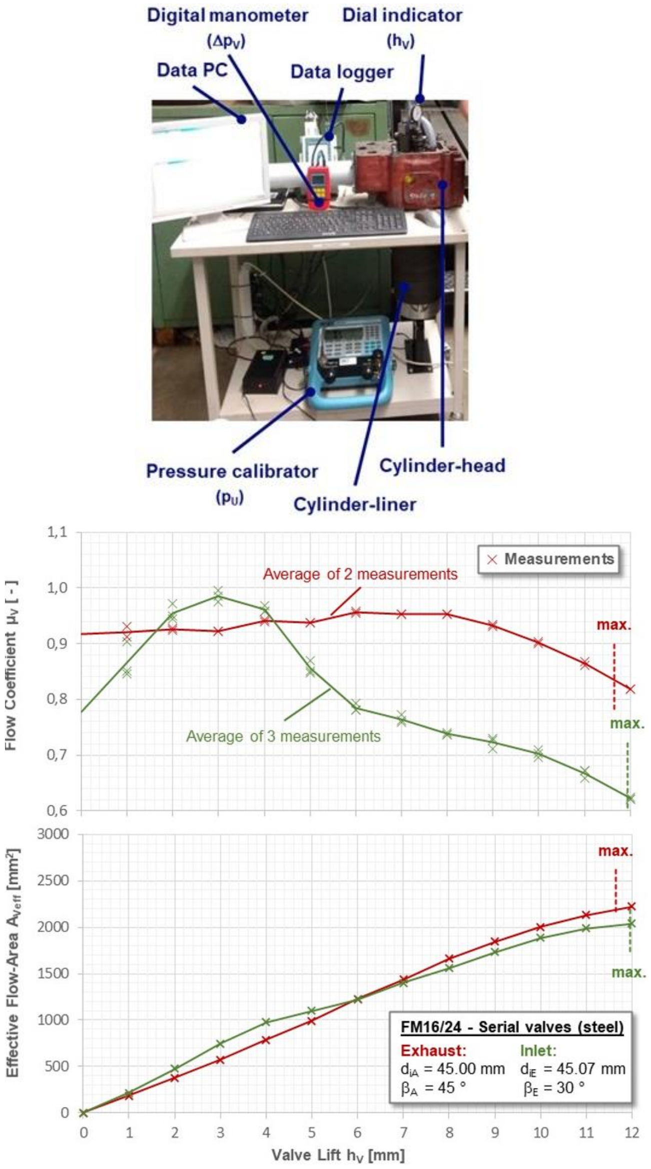


Fig. 4. Repeatability of fluid-dynamic measurements (serial valves)

As known from other medium-speed engines, the flow coefficients for the exhaust valves at higher valve-lifts are larger than the ones at inlet side due to diffuser effects in the exhaust-port. While at higher valve-lifts rising geometric flow areas are more compensated by falling flow coefficients, the increase of effective flow area becomes smaller. With maximum possible valve-lift inside the engine, the maxima of effective flow area were not reached. The valve-lift of the engine is not designed too large.

Several prototype-valves with different seat angles were designed, based on the geometry of standard valve for serial application (Fig. 5). For all prototype-valves, the diameter of valve-desk was kept constant for a sufficient mounting space inside the cylinder head. Equal heights of valve-desks ensured constant dead space between valves and piston inside an engine, as well as a constant compression ratio.

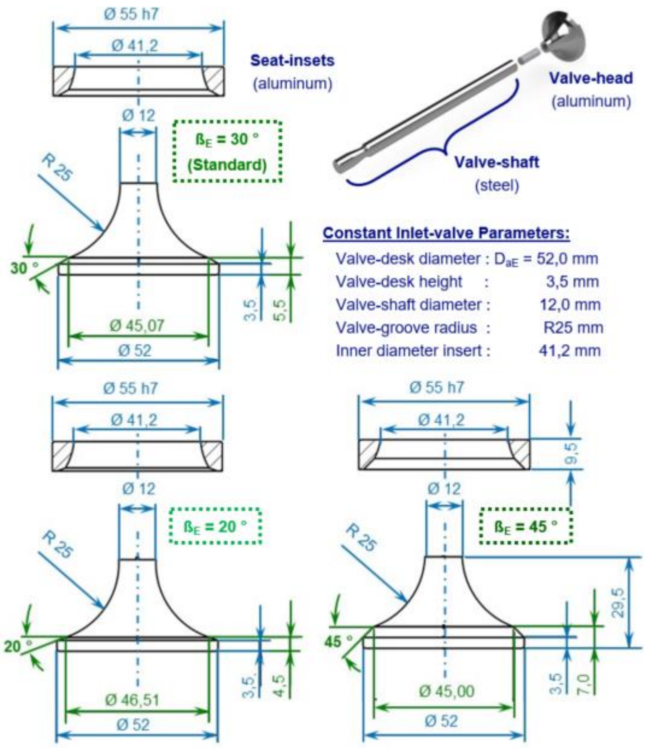


Fig. 5. Split-design inlet valves and valve-seats for fluid-dynamic tests

With a 25 mm radius of damageable valve-groove, it was ensured that sufficient safety would be available regarding mechanical stresses. For economic reasons, all prototype-valves for fluid-dynamic testbed were manufactured in split-design with steel valve-shafts and replaceable valve-heads made of aluminium. Seat-inserts for cylinder head (i.e. 'Seat-rings') were made of aluminium, too. Furthermore, prototypes with standard geometry were made to determine the influence of the manufacturing process. For flow-dynamic tests, the same cylinder head was used as for measurements with serial valves. Serial seat-inserts inside this cylinder head were ground out and replaced by the aluminium ones.

Fig. 6 shows the results of the fluid-dynamic test for the variation of valve-seat angle at inlet. With the split-designed valve in serial geometry, the influence of manufacturing process for prototype could be quantified.

Compared with flow coefficients of serial valves (dotted line and green line in Fig. 4), it is to be seen that the split-designed prototype valve with the same geometry as serial valves delivers slightly lower volume flows at higher valve-lifts and in that way larger flow-velocities. That can be explained by the manufacturing tolerance around the joint between valve-heads and valve-shaft. This deviation is about 7 % in volume-flow. All interpretations of influence of valve-seat angle are referred to prototype valve in serial geometry. Diagrams with flow coefficient in relation to the quotient of valve-lift and inner seat diameter (h_V/d_{Vi}) were not used, and comparisons to other engines were not considered.

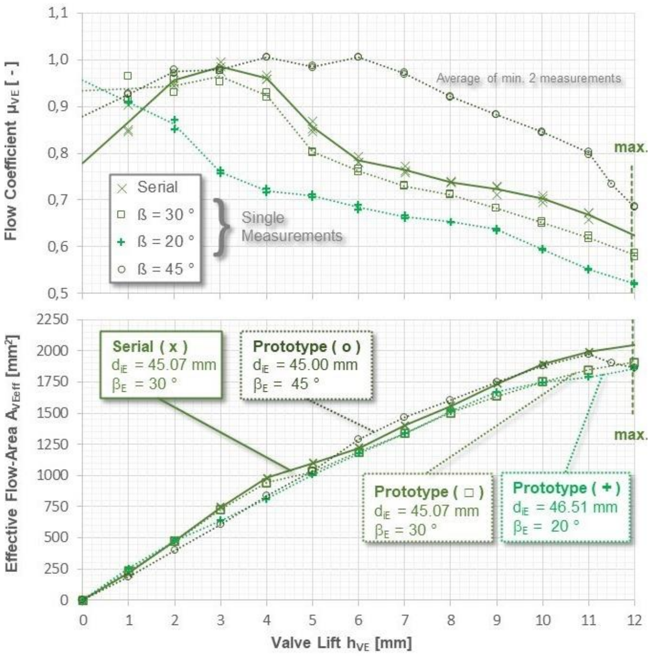


Fig. 6. Results of fluid-dynamic tests for variation of inlet valve-seat angle

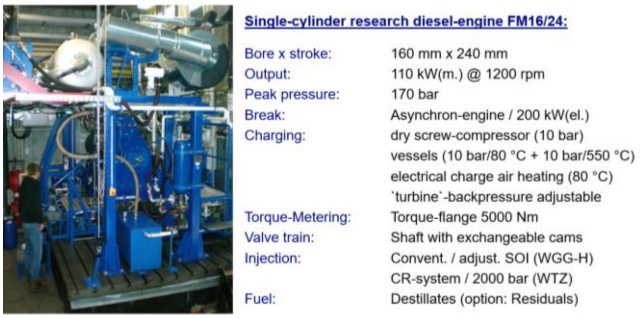
As expected, the larger valve-seat angles deliver slightly larger effective flow areas. But the higher flow coefficients are partly compensated by smaller geometric flow areas, so that the effective flow area at higher valve-lifts is only slightly enlarged for a valve-seat angle of 45°. Up to a valve-lift of 5 mm, which is relevant for scavenging during valve-overlap, the valve-seat with 45° delivers worse flow characteristics compared to a standard geometry. The trend of both effects on the thermal load of the engine are opposing, so that an estimation seems to be difficult. As a next step, this impact had to be considered by real cycle calculations including the T/C.

2.2. Test-run at single-cylinder engine and calibration of calculations

Tests at constant-speed operation were carried out at a single-cylinder test engine, shown with its main technical data in Fig. 7. Charge-air pressure was controlled by an electronic frequency-converter of the dry screw-compressor. Back-pressure of T/C's turbine is simulated by throttle-valve after exhaust-gas vessel. Charge-air temperature remains at 45°C before the cylinder for all operating points. To keep comparableness to serial multi-cylinder engines in field operation, the same indicated power (not the effective output at crankshaft) was adjusted. Three attached pumps for lubrication oil and cooling water (fresh and sea water) were taken into consideration.

For this test-run, a 60-Hz constant-speed operation at 1,200 rpm between 100 % and 10 % load was chosen.

The commencement of delivery of the injection pump was set constant at all loads. This operation represents a typical medium-speed propulsion plant with controllable-pitch propeller (CPP operation) and shaft-generator and gensets for auxiliary use on board ships or stationary electric power-generation. Medium-speed engines driving fixed pitch propellers (FPP operation) are rare in commercial ships.


Single-cylinder research diesel-engine FM16/24:

Bore x stroke:	160 mm x 240 mm
Output:	110 kW(m.) @ 1200 rpm
Peak pressure:	170 bar
Break:	Asynchron-engine / 200 kW(el.)
Charging:	dry screw-compressor (10 bar) vessels (10 bar/80 °C + 10 bar/550 °C) electrical charge air heating (80 °C) `turbine`-backpressure adjustable
Torque-Metering:	Torque-flange 5000 Nm
Valve train:	Shaft with exchangeable cams
Injection:	Convent. / adjust. SOI (WGG-H) CR-system / 2000 bar (WTZ)
Fuel:	Destillates (option: Residuals)

Fig. 7. Test engine for calibration of calculation model

Fuel consumption was measured by a mass flowmeter (Coriolis). All relevant exhaust-gas components were detected by classical physical methods (NDIR, MPA, FIA, CLA). The engine is full-indicated by pressure probes before, inside and after the cylinder as well as in the injection-line. Operational values and emissions according to ISO8178, start of injection, injection rate and heat release in cylinder are calculated by own-developed computer programs.

These measurements carried out were used for calibration and calculation of the cycle as well as for estimation of NOx emissions, as describe below. With additional experimental variations of charge-air pressure, charge-air temperature and start of delivery, the used cycle calculation models were proven regarding right estimation of operational values and NOx emissions.

3. CYCLE CALCULATIONS FOR ESTIMATION OF OPERATIONAL BEHAVIOR AT DIFFERENT SEAT ANGLES

As a next step, cycle calculations were carried out for the measured operating points described in Section 2.2 to validate the calculation models used for these considerations.

For the related cycle calculations the FORTRAN-code 'DYN', developed in a FVV working group [4] for all German medium-speed manufactures, was applied. This source code is open for members and therefore it was developed further and adapted for special needs of the users. For this consideration, the version V36.2 of MAN Augsburg was used [12].

Basis is a real cycle calculation with a single-zone model. Up to 20 cylinders with up to two turbochargers can be connected with up to 12 containers, so that even two-stage charging can be simulated. Calculation of gas exchange is done by simple quasi-static 'fill-and-relief'-method, without any considerations of gas-dynamic effects inside inlet manifold and exhaust-gas line.

All layouts for large 4-stroke MAN diesel engines were carried out with that program. Several comparisons with commercial software (e.g. GT power) showed its reliability even for predictions of back-flow into charge-air receivers at 5L engines. An example for such derated layouts is published in the study of Marquardt [13].

For calculation of the measured operational values, a model with one cylinder and two containers (before and after cylinder) was used (Fig. 8a). Both containers were set to the pressure values for charge-air and exhaust-gas as measured. From pressure indication at full load, the measured heat release inside the cylinder was replaced by a Vibe-function [14]. Vibe-parameters were converted for the other loads following the Woschni/Anisits-rule [15] (Fig. 8b). Therefore, the cylinder mass exponent [16] was set to X = 0.5. Heat transfers to cylinder walls were calculated by Woschni law, as modified by Gerstle/Eilts [6]. For NOx-calculation of medium-speed engines, the two-zone-model by Heider [17] has been well proofed. Its weakness regarding presumptions for variations of air/fuel ratio is known [16] and presented at test-engine FM16/24, too.

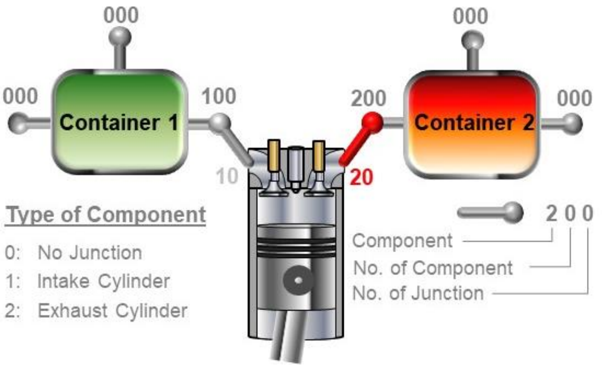
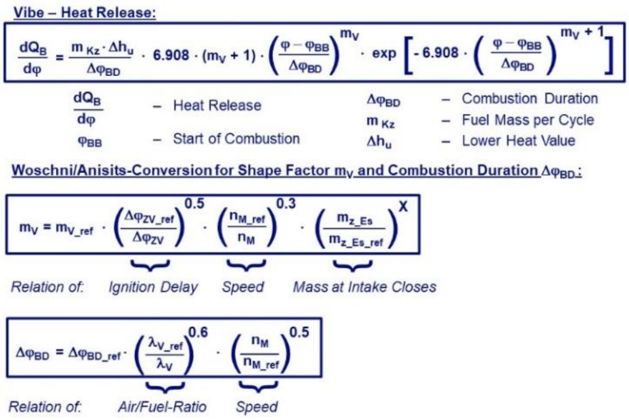


Fig. 8a. Model with two containers for comparison of measurement and cycle calculation


Vibe – Heat Release:
Heat Release
Start of Combustion
Combustion Duration
Fuel Mass per Cycle
Lower Heat Value
Woschni/Anisits-Conversion for Shape Factor m_V and Combustion Duration Δφ_BD:
Relation of:
Ignition Delay
Speed
Mass at Intake Closes
Relation of:
Air/Fuel-Ratio
Speed

$$\frac{dQ_B}{d\varphi} = \frac{m_{Kz} \cdot \Delta h_u}{\Delta\varphi_{BD}} \cdot 6.908 \cdot (m_V + 1) \cdot \left(\frac{\varphi - \varphi_{BB}}{\Delta\varphi_{BD}}\right)^{m_V} \cdot \exp\left[-6.908 \cdot \left(\frac{\varphi - \varphi_{BB}}{\Delta\varphi_{BD}}\right)^{m_V + 1}\right]$$

$\frac{dQ_B}{d\varphi}$ – Heat Release; $\Delta\varphi_{BD}$ – Combustion Duration; m_{Kz} – Fuel Mass per Cycle; φ_{BB} – Start of Combustion; Δh_u – Lower Heat Value

$$m_V = m_{V\_ref} \cdot \underbrace{\left(\frac{\Delta\varphi_{ZV\_ref}}{\Delta\varphi_{ZV}}\right)^{0.5}}_{\text{Ignition Delay}} \cdot \underbrace{\left(\frac{n_{M\_ref}}{n_M}\right)^{0.3}}_{\text{Speed}} \cdot \underbrace{\left(\frac{m_{z\_Es}}{m_{z\_Es\_ref}}\right)^{X}}_{\text{Mass at Intake Closes}}$$

$$\Delta\varphi_{BD} = \Delta\varphi_{BD\_ref} \cdot \underbrace{\left(\frac{\lambda_{V\_ref}}{\lambda_V}\right)^{0.6}}_{\text{Air/Fuel-Ratio}} \cdot \underbrace{\left(\frac{n_M}{n_{M\_ref}}\right)^{0.5}}_{\text{Speed}}$$

Fig. 8b. Modelling and conversion of heat release by Vibe-function [3, 16]

As mentioned before, the test-engine FM16/24 was operated at the same cylinder-indicated power (not with the same effective output at crankshaft) to ensure a comparison to a 9L serial engine. The mechanical efficiency of a MAN-Holeby 9L16/24 inclusive of three attached pumps was known by the authors. The mechanical losses of the crank drive of the single-cylinder FM16/24 test engine were determined before by indicator-method. The conversion of both effective outputs at crankshaft were rendered possible this way. All calculation results are diagrammed in relation to the equivalent output of the serial multi-cylinder engine.

Fig. 8c shows the satisfying precision of cycle calculation for loads between 100 % and 25 % at constant-speed operation. Fuel-oil consumption, peak-pressure in cylinder and air consumption are estimated well at loads above 50 %. Experience shows that differences become larger at low loads. Down to 25 % load the exhaust-gas temperature before turbine has a maximum failure of 10 K, which is quite well. A two-zone calculation model for NOx emissions was calibrated at full load and gives calculated values with a maximum error of 0.25 g/kWh at 25 % load. A valid calibration of engine-model as a precondition for estimations of the influence of valve-seat angle by cycle calculations is documented with this comparison of measurement and cycle calculation.

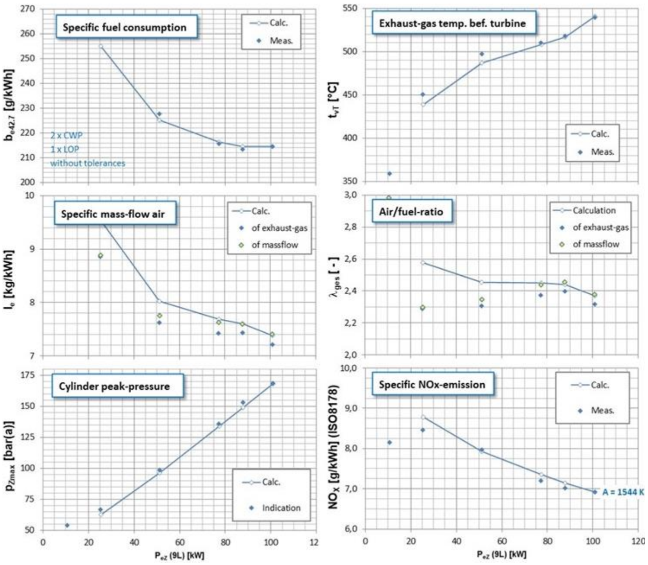


Fig. 8c. Comparison of measurement and cycle calculation at constant-speed operation

For estimation of operational values and emissions, with inlet-valve–seat angle dissenting from standard 30°, cycle calculations for 9L engines with T/C (MAN TCR14-41xxx) were carried out with a six-containers-model (Fig. 9). As is usual for modern engines, a single-exhaust pipe was chosen. The diameter of the exhaust pipe was set to the value of the cylinder bore to ensure constant-pressure charging as applied for field operations of this engine-type. Cooling-water pumps (CWP) for both temperature levels (HT and LT) were attached, as well as an engine-driven lub-oil pump (LOP). The power required for engine-driven pumps was estimated at 19.8 kW. For the sufficient trade-off between fuel-oil consumption and NOx emission for engines with restricted peak-pressure, an injection timing without relevant increase of cylinder-pressure after commencement of combustion was set as a constant for all calculations.

The engine is equipped with conventional valve-timing according to inlet-closing for maximum volumetric efficiency and an overlap of 105° crank-angle. With constant-pressure combustion, a charge-air pressure could be layout to 3.7 bar(a) without exceeding the peak-pressure limits (Table 1). According to the T/C performance of TCR14-41xxx, a maximum turbine-inlet temperature of ≈530°C could be achieved, which is quite high at an ambient pressure of 1 bar. Air/fuel ratio >2 at full load ensures acceptable thermal loads of combustion chamber. NOx emissions according to cycle E2 (ISO8178) for CPP operation ensures IMO-certification for Tier 3 outside Emission-controlled Areas (ECA).

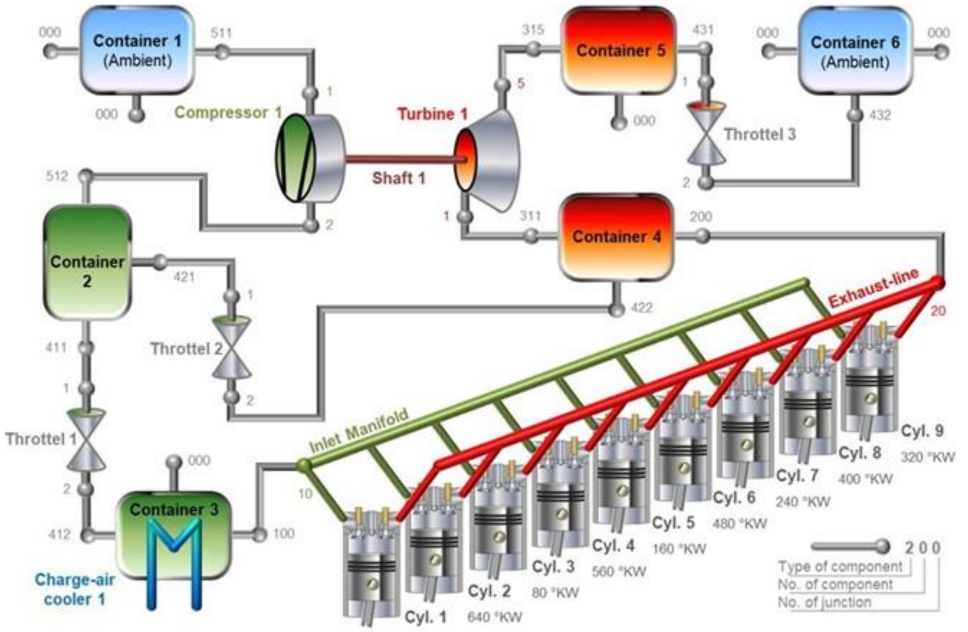
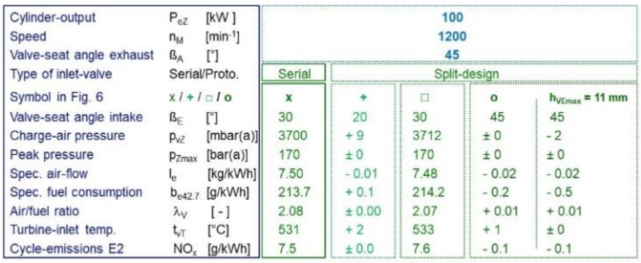


Fig. 9. Calculation model of 9L-engine with constant-pressure T/C and single-stage charge-air cooler

Tab. 1. Results of cycle calculation



Cylinder-output	P_{eZ}	[kW]	100				
Speed	n_M	[min^{-1}]	1200				
Valve-seat angle exhaust	$ß_A$	[°]	45				
Type of inlet-valve	Serial/Proto.		Serial	Split-design			
Symbol in Fig. 6	x / + / □ / o		x	+	□	o	h_{VEmax} = 11 mm
Valve-seat angle intake	$ß_E$	[°]	30	20	30	45	45
Charge-air pressure	p_{vZ}	[mbar(a)]	3700	+ 9	3712	± 0	- 2
Peak pressure	p_{Zmax}	[bar(a)]	170	± 0	170	± 0	± 0
Spec. air-flow	l_e	[kg/kWh]	7.50	- 0.01	7.48	- 0.02	- 0.02
Spec. fuel consumption	$b_{e42.7}$	[g/kWh]	213.7	+ 0.1	214.2	- 0.2	- 0.5
Air/fuel ratio	λ_V	[-]	2.08	± 0.00	2.07	+ 0.01	+ 0.01
Turbine-inlet temp.	t_{vT}	[°C]	531	+ 2	533	+ 1	± 0
Cycle-emissions E2	NO_x	[g/kWh]	7.5	± 0.0	7.6	- 0.1	- 0.1

Deviations to the calculated values with same valve-geometry in split-design are quite small. To compensate the slightly enlarged flow-resistance, the charge-air pressure had to be raised by 12 mbar. The resulting fuel consumption is enlarged by 0.2 % according to the higher losses in gas exchange. Exhaust-gas temperature is nearly the same. The chosen strategy of using split-designed valves for fluid-dynamic tests seemed to be useful.

From that layout, the deviations in operational values and emissions were estimated by further cycle calculations for the unconventional inlet valve-seat angles of 20° and 45°. In every case, charge-air pressure at full load was adjusted to achieve peak-pressures of 170 bar for all layouts. With valve-set angle of 20°, the exhaust-gas temperature at turbine inlet raises by barely 2 K, which seems to be a very small increase in thermal load.

By using the larger seat angle 45°, the thermal load of the engine inside the cylinder, mainly determined by its air/fuel ratio, is nearly the same as with standard geometry. Temperature at turbine inlet is slightly increased by 1 K, due to the lower effective flow area at full lift of inlet valve. No significant influence of valve-seat angles on NOx emissions is identified. This result is in conformity with the fact that usually no IMO-numbers are marked at valves, so that they are not classified as NOx-relevant engine components.

According to Fig. 6, for a valve-seat angle of 45°, the largest effective flow area is not reached at maximum valve-lift. The maximum effective flow area was reached 1 mm before maximum valve-lift, so that the valve-lift at inlet could be reduced with marginal thermodynamic advantages and a slightly improved fuel consumption. Higher dead-spaces between valves and piston or reduced deepness of cut-outs in piston-crown (i.e. 'valve pockets') could also be chosen. To this end, a newly designed inlet cam, inclusive of new NOx certification from IMO, would be necessary.

4. CONCLUSIONS

The influence of unconventional valve-seat angles at inlet of large medium-speed diesel engines was considered with experimental investigations at fluid-dynamic testbed. Before commencing this small research project, related literature studies were undertaken, and no comparable investigations were found.
The impacts of the resulting effective flow areas at the engine's inlet on the operational behaviour and NOx emissions were estimated by cycle calculation for turbo-charged multi-cylinder serial engines. For this use, calculation models were calibrated by experimental data, generated at the single-cylinder test-engine FM16/24.

By wear optimised seat-angles at inlet of β_E = 20° specific fuel consumption raised by 0.1 g/kWh compared to standard β_E = 30°, temperature deviation of 2 K at turbine inlet does not influence the thermal load of the T/C significantly. The air/fuel ratio during combustion is not changed, so that the temperatures at valves will not be higher than the standard design. No negative influences were identified for reduced valve angles at inlet. Given the expected larger valve-seat wear while operating medium-speed engines on fuels without lubricating combustion deposits, the reduction of valve-seat angles at inlet seems to be useful.

Fuel savings with enlarged inlet valve-seat angles β_E = 45° are with b_e = –0.2 g/kWh lower than the tolerances of testbeds for factory acceptance tests. Even with an optimised cam-geometry no thermal derating of the turbine could be detected by these investigations. Enlarged seat angles at the inlet seem to be unproductive.

Mechanical impacts to the wear of valve-seats and head-inserts must be considered by tests in field operation. With valve-seat lubrication, an easy and economical countermeasure was suggested, which can be applied if needed.

Investigations are continued by analogue experiments at exhaust-side as well as with studies of the influence of serial tolerances of casting processes for the cylinder heads [18].

REFERENCES

1. Wimmer A. The Future Role of IC Engines for Sustainable Ship Propulsion. 2nd LEC Sustainable Shipping Technologies Forum. Graz 27.04.2021.
2. Swiderski E. Experimentelle Bestimmung der Ventil-Durchflussbeiwerte eines Großdieselmotors. Studienarbeit im Masterstudiengang „Maschinenbau“ an der Universität Rostock. durchgeführt an der FH Stralsund, Rostock 2017.
3. Klimpel T. Untersuchungen zum Luftdurchsatz und zur Energiebilanz mittelschnellaufender Dieselmotoren, Diplomarbeit TU Berlin/MAN Augsburg, Berlin 1995.
4. Zellbeck H. Ermittlung des dynamischen Betriebsverhaltens von abgasturboaufgeladenen Dieselmotoren. Abschlussbericht FVV-Vorhaben 234 + 286, Heft 304. Frankfurt/M. 1982.
5. Szpica D. The assessment of the influence of temperature differences in individual ducts of an intake manifold on the unevenness of air filling in a cylinder of a combustion engine. Combustion Engines. 2008; 2(133):44ff.
6. Gerstle M. Simulation des instationären Betriebsverhaltens hochaufgeladener Vier- und Zweitakt-Dieselmotoren. Diss. Universität Hannover. 1999.
7. Flenker H, Woschni, G. Programmiertes Berechnungsverfahren zur Bestimmung der Prozessdaten aufgeladener Vier- und Zweitaktdieselmotoren bei geänderten Betriebsbedingungen. Bericht des Inst. für Verbrennungskraftmaschinen TU Braunschweig Nr. 74/2 / CIMAC Working Group „Supercharging“, Frankfurt/M. 1974
8. Beitz W, Grote KH. DUBBEL – Taschenbuch für den Maschinenbau. 19. Auflage, Springer-Verlag, Berlin 1997, S. P52.
9. Heywood JB. Internal Combustion Engine Fundamentals. McGraw-Hill Inc., New York, 1988, S. 220ff.
10. Tanaka K. Air Flow Through Suction Valves of Conical Seat (Part I. Experimental Research), Aeronautical Research Institute. Tokio Emperial University. Report No. 50, October 1929 .
11. Urlaub A. Verbrennungsmotoren. 2. Auflage. Springer-Verlag, Berlin 1995, S. 171ff.
12. Heider K. Thermodynamische Berechnungen mit dem Programm DYN V36.2. Arbeitspapier MAN B&W Diesel, Augsburg 2004.
13. Marquardt L. Leistungsbemessung an 5-Zylinder-Gensets. HANSA – International Maritime Journal, Schiffahrts-Verlag „Hansa“ Hamburg, Heft 09/2022, S.42f.
14. Vibe II. Brennverlauf und Kreisprozeß von Verbrennungsmotoren. VEB Verlag Technik. Berlin 1970.
15. Woschni G. Anisits, F. Eine Methode zur Vorausberechnung der Änderungen des Brennverlaufs mittelschnellaufender Dieselmotoren bei geänderten Betriebsbedingungen. MTZ 34 (1973), S. 106 ff.
16. Marquardt L. Theoretische und experimentelle Untersuchungen zur innermotorischen Stickoxidreduzierung bei mittelschnellen Großdieselmotoren im Schwerölbetrieb. Diss. TUHH, Shaker-Verlag. Aachen 2007, ISBN 978-3-8322-5912-9.
17. Heider G. Rechenmodell zur Vorausrechnung der NO-Emissionen von Dieselmotoren. Diss. TU München. 1996.
18. Marquardt L, Katke HJ, Reinke A, Kockskämper N. Möglichkeiten und Grenzen der konstruktiven Ladungswechseloptimierung für mittelschnelle Großmotoren. 29. REGWA Energie-Symposium, Stralsund 2022.

Acknowledgements: This work has been accomplished under research project No. 1-2022, financed by the Internal Research Funding of Hochschule Stralsund.

Authors wish to express their thanks for all support to

DFG:
INST 295/44-1 FUGG „Einzylinder-Forschungsmotor“
INST 295/54-1 „Abgasanalyse“

NORIS Automation Rostock:
Modification of testbed automation system

EFRE:
GHS16-0027 „Messkette Rest-O2“
GHS16-0028 „ScopeCorder“
GHS16-0033 „Messsystem für Druckschwingungen“
GHS17-0024 „Ladungswechseloptimierung“

DMR-Kloska:
7 x cylinder heads SKL VD24/16-AL

This paper pre-reviewed by L. Grünsteudel (Serial design and licencees) and P. Holand (Thermodynamics) at MAN Augsburg.

HOST DFG Deutsche Forschungsgemeinschaft NORIS

EUROPÄISCHE UNION Europäischer Fonds für Regionale Entwicklung Kloska Group www.kloska.com MAN

Leander Marquardt: https://orcid.org/0000-0003-0600-6166

Heiner-Joachim Katke: https://orcid.org/0000-0001-7137-8244

Andreas Reinke: https://orcid.org/0000-0002-3011-2762

Niklas Kockskämper: https://orcid.org/0000-0003-3454-6706

DOI 10.2478/ama-2023-0020

FRACTIONAL ORDER, STATE SPACE MODEL OF THE TEMPERATURE FIELD IN THE PCB PLATE

Krzysztof OPRZĘDKIEWICZ*, Wojciech MITKOWSKI*, Maciej ROSÓŁ*

*Faculty of Electrical Engineering Automatic Control Informatics and Biomedical Engineering,
Department of Automatic Control and Robotics, AGH University of Science and Technology,
al. A Mickiewicza 30, 30-059 Kraków, Poland

kop@agh.edu.pl, wojciech.mitkowski@agh.edu.pl, mr@agh.edu.pl

received 9 July 2022, revised 26 October 2022, accepted 6 November 2022

Abstract: In the paper the fractional order, state space model of a temperature field in a two-dimensional metallic surface is addressed. The proposed model is the two dimensional generalization of the one dimensional, fractional order, state space of model of the heat transfer process. It uses fractional derivatives along time and length. The proposed model assures better accuracy with lower order than models using integer order derivatives. Elementary properties of the proposed model are analysed. Theoretical results are experimentally verifed using data from industrial thermal camera.

Key words: fractional order systems, fractional order state equation, temperature field, heat transfer, thermal camera

Abbreviations: MIMO – Multiple Input Multiple Output, IO – Integer Order, FO – Fractional Order, CFE – Continuous Fraction Expansion, IIR – Infinite Impulse Response, FIR – Finite Impulse Response, PSE – Power Series Expansion.

1. INTRODUCTION

The modelling of processes and plants hard to describe using another mathematical tools is one of the main areas of application of the FO calculus.

Fractional models of different physical phenomena have been presented by many Authors for years. Fundamental results are presented e.g. by (1), (2) (the heating of an one dimensional beam), (3) (FO models of chaotic systems and Ionic Polymer Metal Composites), (4). Fractional Order diffusion processes are considered u.a. in (5), (6), (7). A collection of results using new Atangana-Baleanu operator can be found in (8). This book presents i.e. the FO blood alcohol model, the Christov diffusion equation and fractional advection-dispersion equation for groundwater transport processes.

Recently FO models are used u.a. to describe a spread of diseases. This issue is considered e.g. in the papers: (9) (the modelling of the dynamics of COVID using Caputo-Fabrizio operator), (10) (the modelling of a transmission of Zika virus using Atangana-Baleanu operator).

The state space FO models of the one dimensional heat transfer have been proposed in many previous papers of authors, e.g. (11), (12), (13), (14), (15), (16), (17), (18). These models employed different FO operators: Grünwald-Letnikov, Caputo, Caputo-Fabrizio and Atangana-Baleanu as well as discrete operators CFE and PSE. Each proposed model has been thoroughly theoretically and experimentally verified. In addition, each of them assures better accuracy in the sense of square cost function than its IO analogue.

Models of temperature fields obtained with the use of thermal cameras are presented e.g. by (19), (20). Analytical solution of the two-dimensional, IO heat transfer equation is presented in the paper (21). Numerical methods of solution of PDE-s can be found, e.g., in (22). Fractional Fourier integral operators are analyzed u.a. by (23). It is important to note that a significant part of known investigations deals only with a steady-state temperature elds with omitting their dynamics.

The paper (24) presents the generalization of FO models mentioned above to a two dimensional surface. It is important to note that in this paper the FO derivation only along the time is considered. The derivation along both space coordinates is described by the 2'nd order operator.

In this paper we propose and analyze a new, FO, state space model of heat transfer in a flat metallic surface. The model uses FO derivatives along time and space coordinates. Such an approach allows to obtain the better accuracy in the sense of a square cost function than model proposed previously. In addition, it is expected that satisfying accuracy will be achieved for relatively low order. Our knowledge shows that such a model has not been proposed yet. The proposed approach can be employed e.g. to modelling and reconstruction images from thermal cameras.

The paper is organized as follows. Preliminaries recall some basic ideas from fractional calculus, necessary to present results. Next the model using FO operator along the time is remembered. Furthermore its generalization applying FO operators along both time and space coordinates is presented and analyzed. Finally theoretical results are verified with the use of experimental data.

2. PRELIMINARIES

At the beginning the non-integer order, integro-differential operator is presented (see e.g. (1), (4), (25), (26)).

Definition 1 *(The elementary non integer order operator)* The non-integer order integro-differential operator is defined as follows:

$$ {}_aD_t^{\alpha} g(t) = \begin{cases} \frac{d^{\alpha} g(t)}{dt^{\alpha}} \ \alpha > 0 \\ g(t) \ \alpha = 0 \\ \int_a^t g(\tau)(d\tau)^{-\alpha} \ \alpha < 0 \end{cases} \quad (1) $$

where a and t denote time limits for operator calculation, $\alpha \in \mathbb{R}$ denotes the non integer order of the operation.

Next an idea of complete Gamma Euler function is recalled (see for example (26)):

Definition 2 *(The complete Gamma function)*

$$ \Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad (2) $$

An idea of Mittag-Leffler function needs to be given next. It is a non-integer order generalization of exponential function et and it plays crucial role in the solution of a FO state equation. The two parameter Mittag-Leffler function is defined as follows:

Definition 3 *(The two parameter Mittag-Leffler function)*

$$ E_{\alpha,\beta}(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(k\alpha+\beta)} \quad (3) $$

For $\beta = 1$ we obtain the one parameter Mittag-Leffler function:

Definition 4 *(The one parameter Mittag-Leffler function)*

$$ E_{\alpha}(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(k\alpha+1)} \quad (4) $$

The fractional order, integro-differential operator (1) is described by different definitions, given by Grünwald and Letnikov (GL definition), Riemann and Liouville (RL definition) and Caputo (C definition). Relations between Caputo and Riemann-Liouville, between Riemann-Liouville and Grünwald-Letnikov operators are given e.g. in (27), (4). Discrete versions of these operators are analysed with details in (28). The C definition has a simple interpretation of an initial condition (it is analogical as in integer order case) and intuitive Laplace transform. Additionally its value from a constant equals to zero, in contrast to e.g. RL definition. That's why in the further consideration the C definition will be used. It is recalled beneath.

Definition 5 *(The Caputo definition of the FO operator)*

$$ {}_0^CD_t^{\alpha} f(t) = \frac{1}{\Gamma(V-\alpha)} \int_0^{\infty} \frac{f^{(V)}(\tau)}{(t-\tau)^{\alpha+1-V}} d\tau \quad (5) $$

In (5) V is a limiter of the non-integer order: $V-1 \le \alpha < V$. If $V = 1$ then consequently $0 \le \alpha < 1$ is considered and the definition (5) takes the form:

$$ {}_0^CD_t^{\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^{\infty} \frac{\dot{f}(\tau)}{(t-\tau)^{\alpha}} d\tau \quad (6) $$

Finally a fractional linear state equation using Caputo definition should be recalled. It is as follows:

$$ \begin{aligned} {}_0^CD_t^{\alpha} x(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \quad (7) $$

where $\alpha \in (0,1)$ is the fractional order of the state equation, $x(t) \in \mathbb{R}^R$, $u(t) \in \mathbb{R}^L$, $y(t) \in \mathbb{R}^P$ are the state, control and output vectors respectively, *A*, *B*, *C* are the state, control, and output matrices.

3. THE EXPERIMENTAL SYSTEM AND ITS FO MODEL

The Figure 1 shows the simplified scheme of the considered heat system. This is the PCB plate of size *X* × *Y* pixels. The values of *X* and *Y* are determined by the resolution of a sensor of a camera. The plate is heated by at heater denoted by *H*. Its coordinates are denoted by x_{h1}, x_{h2}, y_{h1} and y_{h2} respectively. The surface area S_H of the heater is equal:

$$ S_H = d_{xh} d_{yh} \quad (8) $$

where:

$$ \begin{aligned} d_{xh} &= x_{h2} - x_{h1} \\ d_{yh} &= y_{h2} - y_{h1} \end{aligned} \quad (9) $$

The temperature is measured using thermal camera, the area of measurement is configurable and denoted by *S*. Its coordinates are equal x_{s1}, x_{s2}, y_{s1} and y_{s2}. The surface area S_S of the measurement area is equal:

$$ S_S = d_{xs} d_{ys} \quad (10) $$

where:

$$ \begin{aligned} d_{xs} &= x_{s2} - x_{s1} \\ d_{ys} &= y_{s2} - y_{s1} \end{aligned} \quad (11) $$

More details about the construction of this laboratory system are given in the section "Experimental Results". The heat transfer in the surface is described by the Partial Differential Equation (PDE) of the parabolic type. All the side surfaces of plate are much smaller than its frontal surface. This allows to assume the homogeneous Neumann boundary conditions at all edges of the plate as well as the heat exchange on the surface needs to be also considered. It is expressed by coefficient R_a. The control and observation are distributed due to the size of heater and size of temperature field read by camera. The heat conduction coefficient aw along both directions x and y is the same.

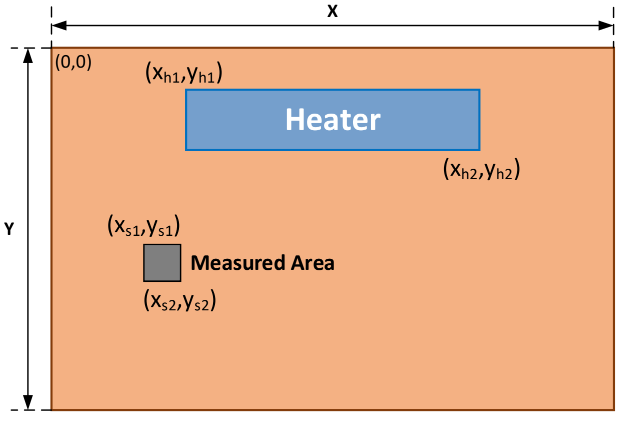


Fig. 1. The simplified scheme of the experimental system. Origin of the coordinate system is located in the left upper corner

DOI 10.2478/ama-2023-0020

The two dimensional, IO heat transfer equation has been considered in many papers (e.g. (29), (30), (31)). The fractional version with fractional derivative along the time and 2'nd order integer derivative along the length is presented with details in the paper (24).

This paper presents the model employing the fractional derivatives along both coordinates. This allows to obtain better accuracy with its smaller size.

The proposed model is as follows:

$$\begin{cases} {}_0^C D_t^\alpha Q(x,y,t) = a_w\left(\frac{\partial^\beta Q(x,y,t)}{\partial x^\beta} + \frac{\partial^\beta Q(x,y,t)}{\partial y^\beta}\right), \\ -R_a Q(x,y,t) + b(x,y)u(t), \\ \frac{\partial Q(0,y,t)}{\partial x} = 0, t \geq 0, \\ \frac{\partial Q(X,y,t)}{\partial x} = 0, t \geq 0, \\ \frac{\partial Q(x,0,t)}{\partial y} = 0, t \geq 0, \\ \frac{\partial Q(x,Y,t)}{\partial y} = 0, t \geq 0, \\ Q(x,y,0) = Q_0, 0 \leq x \leq X, 0 \leq y \leq Y, \\ y(t) = k_0 \int_0^X \int_0^Y Q(x,y,0)c(x,y)dxdy. \end{cases} \tag{12}$$

In (12) α and β are non-integer orders of the system, $a_w > 0$, $R_a \in \mathbb{R}$ are coefficients of heat conduction and heat exchange, k_0 is a steady-state gain of the model, $b(x, y)$ and $c(x, y)$ are heater and sensor functions described as follows:

$$b(x,y) = \begin{cases} 1, x,y \in H \\ 0, x,y \notin H \end{cases} \tag{13}$$

$$c(x,y) = \begin{cases} 1, x,y \in S \\ 0, x,y \notin S \end{cases} \tag{14}$$

Let $\Omega \in \mathbb{R}^N$ it be an appropriate restricted area. The fractional Laplace operator $\Delta = \frac{\partial^\beta(..)}{\partial x^\beta} + \frac{\partial^\beta(..)}{\partial y^\beta}$ in $L^\beta(\Omega)$ with Dirichlet or Neumann boundary conditions is a discrete operator. The discrete operator has only a point spectrum (see e.g. (32), pp. 204, 460). Without going into details it is generally known from the spectral theorem for compact and self-adjoint operators that all eigenvalues $\lambda_{m,n}$ of the Laplace operator Δ in $L^2(\Omega)$ (with Dirichlet or Neumann boundary conditions) are non-negative, with finite multiplicities and $\lambda_n \to \infty$ for $n \to \infty$. Additionally, there is in $L^2(\Omega)$ an orthonormal basis (complete system) composed of eigenfunctions of the appropriate operator Δ. In special cases of the area $\Omega \in \mathbb{R}^N$ (e.g. for a rectangle on a plane) analytical formulae for eigenvalues and eigenfunctions of the appropriate Laplace operator Δ can be given (see e.g. (33) pp. 21, 26 for the Dirichlet problem or (34) p 133, 138, 301, 305).

The Laplace operator Δ is self-adjointed, it has compact resolvent a it builds the Hilbert base $L^\beta(\Omega) = \{v: \int v^\beta(\xi)d\xi < \infty\}$ with standard scalar product $\langle v,u \rangle = \int_u u(\xi)y(\xi)d\xi$.

The eigenfunctions and eigenvalues for the Laplace operator Δ and Dirichlet boundary conditions are given u.a. by (34), pp.133, 138, 301, 305 or in book (32), pp.253-255.

The construction of the experimental system requires to assume the homogenous Neumann boundary conditions. This yields the following form of eigenfunctions and eigenvalues:

$$w_{m,n}(x,y) = \begin{cases} 1, m,n = 0, \\ \frac{2Y}{\pi n}\cos\frac{n\pi y}{Y}, m = 0, n = 1,2, \dots \\ \frac{2X}{\pi m}\cos\frac{m\pi x}{X}, n = 0, m = 1,2, \dots \\ \frac{2}{\pi}\frac{1}{\sqrt[\beta]{\frac{m^\beta}{X^\beta}+\frac{n^\beta}{Y^\beta}}}\cos\frac{m\pi x}{X}\cos\frac{n\pi y}{Y}, m,n = 1,2, \dots \end{cases} \tag{15}$$

$$\lambda_{m,n}(x,y) = -a_w\left[\frac{m^\beta}{X^\beta} + \frac{n^\beta}{Y^\beta}\right]\pi^\beta - R_a,\ m,n = 1,2, \dots \tag{16}$$

Consequently the considered two dimensional heat equation (12) can be expressed as an infinite dimensional state equation:

$$\begin{aligned} {}_0^C D_t^\alpha Q(t) &= AQ(t) + Bu(t) \\ y(t) &= CQ(t) \end{aligned} \tag{17}$$

where:

$$\begin{aligned} &AQ = a_w\left(\frac{\partial^\beta Q(x,y)}{\partial x^\beta} + \frac{\partial^\beta Q(x,y)}{\partial y^\beta}\right) - R_a Q(x,y), \\ &D(A) = \{Q \in H^2(0,1): Q'(0) = 0, Q'(X) = 0, Q'(Y) = 0\}, \\ &a_w, R_a > 0, \\ &CQ(t) = \langle c, Q(t) \rangle, Bu(t) = bu(t). \end{aligned} \tag{18}$$

The state vector $Q(t)$ is defined as beneath:

$$Q(t) = [q_{0,0}, q_{0,1}, q_{0,2}, \dots, q_{1,0}, q_{1,1}, q_{1,2}, \dots]^T \tag{19}$$

The main difference to the model presented in the paper (24) is that the non-integer order β must be taken into account in the state, control and observation operators. This is presented below.

The state operator A takes the following form:

$$A = diag\{\lambda_{0,0}, \lambda_{0,1}, \lambda_{0,2}, \dots, \lambda_{1,0}, \lambda_{1,1}, \dots, \lambda_{2,2}, \dots, , \lambda_{m,n}, \dots\}. \tag{20}$$

The control operator takes the following form:

$$B = [b_{0,0}, b_{0,1}, b_{0,2}, \dots, b_{1,0}, b_{1,1}, \dots]^T. \tag{21}$$

$$b_{m,n} = \langle H, w_{m,n} \rangle = \int_0^X \int_0^Y b(x,y)w_{m,n}(x,y)dxdy. \tag{22}$$

Taking into account (15) we obtain:

$$b_{m,n} = \begin{cases} S_H, m,n = 0, \\ \frac{2Y^2}{h_{yn}^2} d_{xh} a_{nhy}, m = 0, n = 1,2, \dots \\ \frac{2X^2}{h_{xm}^2} d_{yh} a_{mhx}, n = 0, m = 1,2, \dots \\ \frac{k_{mn}}{h_{xm}h_{yn}} a_{mhx} a_{nhy}, m,n = 1,2, \dots \end{cases} \tag{23}$$

where S_H, d_{xh} and d_{yh} are described by (8), (9) and:

$$\begin{aligned} h_{xm} &= \frac{m\pi}{X}, \\ h_{yn} &= \frac{n\pi}{Y}. \end{aligned} \tag{24}$$

$$k_{mn} = \frac{2}{\pi}\frac{1}{\sqrt[\beta]{\frac{m^\beta}{X^\beta}+\frac{n^\beta}{Y^\beta}}}. \tag{25}$$

$$\begin{aligned} a_{mhx} &= \sin(\square_{xm} x_{\square 2}) - \sin(\square_{xm} x_{\square 1}), \\ a_{nhy} &= \sin(\square_{yn} y_{\square 2}) - \sin(\square_{yn} y_{\square 1}). \end{aligned} \tag{26}$$

The output operator is as beneath:

$$C = \left[c_{0,0}, c_{0,1}, c_{0,2}, \ldots, c_{1,0}, c_{1,1}, \ldots\right]. \tag{27}$$

where:

$$c_{m,n} = \langle S, w_{m,n} \rangle = \int_0^X \int_0^Y c(x,y) w_{m,n}(x,y) dxdy. \tag{28}$$

In (28) each element $c_{m,n}$ is expressed analogically, as (23):

$$c_{m,n} = \begin{cases} S_s, m,n = 0, \\ \frac{2Y^2}{h_{yn}^2} d_{xs} a_{nsy}, m = 0, n = 1,2, \ldots \\ \frac{2X^2}{h_{xm}^2} d_{ys} a_{msx}, n = 0, m = 1,2, \ldots \\ \frac{k_{mn}}{h_{xm} h_{yn}} a_{msx} a_{nsy}, m,n = 1,2, \ldots \end{cases} \tag{29}$$

In (29) $h_{xm,yn}$ and k_{mn} are expressed by (24), (25), S_H, d_{xh} and d_{yh} are described by (10), (11) and:

$$\begin{aligned} a_{msx} &= \sin(\square_{xm} x_{s2}) - \sin(\square_{xm} x_{s1}), \\ a_{nsy} &= \sin\left(\square_{yn} y_{s2}\right) - \sin\left(\square_{yn} y_{s1}\right). \end{aligned} \tag{30}$$

3.1. The decomposition of the model

The decomposition of the model is done by the similar way as for the system considered in the paper (24). Due to the form of the state operator A (20) the system (17)-(29) can be splitted to infinite number of independent scalar subsystems associated to particular eigenvalues:

$$\begin{aligned} & D^\alpha Q(t) = AQ(t) + Bu(t), \\ & Aw_{m,n} = \lambda_{m,n} w_{m,n}. \\ & D^\alpha Q(t) = a_w \left(\frac{\partial^\beta Q}{\partial x^\beta} + \frac{\partial^\beta Q}{\partial y^\beta} \right) - R_a Q + Bu, \\ & D^\alpha \textstyle\sum_{m=0}^\infty \sum_{n=0}^\infty q_{m,n} = a_w \left(\sum_{m=0}^\infty \frac{\partial^\beta q_{m,n}}{\partial x^\beta} + \sum_{n=0}^\infty \frac{\partial^\beta q_{m,n}}{\partial y^\beta} \right) - \\ & -R_a \textstyle\sum_{m=0}^\infty \sum_{n=0}^\infty q_{m,n} + Bu, \\ & D^\alpha q_{m,n} = a_w \left(\frac{\partial^\beta q_{m,n}}{\partial x^\beta} + \frac{\partial^\beta q_{m,n}}{\partial y^\beta} \right) - R_a q_{m,n} + b_{m,n} u. \end{aligned} \tag{31}$$

The form of equation (31) implies the decomposition of the system (7) into systems related to single eigenvalues $\lambda_{m,n}, m, n = 0,1,2, \ldots$. This decomposition allows to easily compute the step and impulse responses of the system as a sum of responses of particular modes. This is presented below.

Assume the homogenous initial condition in the equation (12):$Q(x; y_0) = 0$. Then the step response is as follows:

$$y_\infty(t) = \textstyle\sum_{m=1}^\infty \sum_{n=1}^\infty y_{m,n}(t). \tag{32}$$

where m, n-th mode of response is as follows:

$$y_{m,n}(t) = \frac{E_\alpha(\lambda_{m,n} t^\alpha) - 1(t)}{\lambda_{m,n}} b_{m,n} c_{m,n}. \tag{33}$$

In (33) $E_\alpha(\ldots)$ is the one parameter Mittag-Leffler function (4), $\lambda_{m,n}$, $b_{m,n}$ and $c_{m,n}$ are expressed by (16), (22) and (28) respectively.

Analogically the impulse response takes the following form:

$$g_\infty(t) = \textstyle\sum_{m=1}^\infty \sum_{n=1}^\infty g_{m,n}(t). \tag{34}$$

where m, n-th mode of response is as follows:

$$g_{m,n}(t) = E_{\alpha,\alpha}\left(\lambda_{m,n} t^\alpha\right) b_{m,n} c_{m,n}. \tag{35}$$

In (35) $E_{\alpha,\alpha}(\ldots)$ is the two parameters Mittag-Leffler function (3).

During simulations it is possible to use of the finite - dimensional sums only. Consequently (32) and (34) take the following form:

$$y_{M,N}(t) = \textstyle\sum_{m=0}^M \sum_{n=0}^N y_{m,n}(t). \tag{36}$$

$$g_{M,N}(t) = \textstyle\sum_{m=0}^M \sum_{n=0}^N g_{m,n}(t). \tag{37}$$

The values of M and N assuring the assumed accuracy and convergence of the model can be estimated numerically or analytically.

3.2. The convergence

The convergence of the proposed model can be estimated using Rate of Convergence (ROC) defined as the steady-state value of the strongest damped mode of the step response. The estimation of it in the one dimensional case was relatively simple due to particular modes of a step response are in descending order.

In the two dimensional case particular modes of a step response are not ordered. In addition, eigenvalues (16) can be multiple. This makes the analysis of the convergence of the proposed model a little bit more complicated. The ROC is defined as follows:

$$ROC_{M,N} = \min_{M,N} \left| \frac{b_{M,N} c_{M,N}}{\lambda_{M,N}} \right|. \tag{38}$$

The relation (38) can be employed to numerical analysis of convergence in different places of measurement. This will be presented in the next section.

4. EXPERIMENTAL VALIDATION OF RESULTS

4.1. The experiments

Experiments were done with the use of the heat system shown in the Figure 2. The dimensions of the PCB plate in pixels are: X = 380, Y = 290. The PCB is heated by the heater 170x20 pixels with maximum power 10W located in points: x_{h1} = 100, y_{h1} = 40 The temperature field is read using thermal camera OPTRIS PI 450, connected to computer via USB and installed dedicated software OPTRIS PI CONNECT. The measured temperature covers range 0 – 250 °C, the sampling frequency is 80 Hz. The signal powering the heater is given from computer using NI LabView, NI myRIO and amplifier. The maximum current from amplifier equal 400mA at a voltage of 12V gives the maximum power 4:8W. The tested PCB plate is not isolated from the environment. This implies that measurements strongly depend on ambient temperature. An another cause of noise during The considered experiment was done in hot summer. During experiments the step response of the system was investigated. The "zero" level denotes the heater switched off, the "one" level is the full power of the heater.

DOI 10.2478/ama-2023-0020

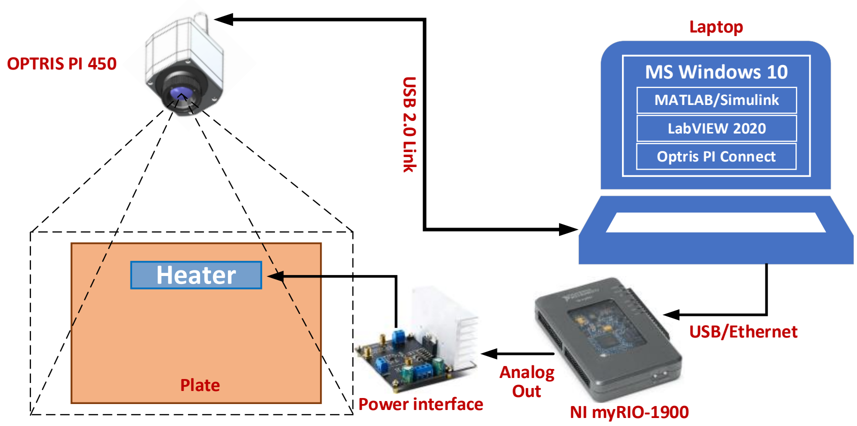


Fig. 2. The experimental system

The temperature fields for both states are shown in the Figure 3. This figure shows also the points of measurement of the step response, marked as "Area 1–4". Areas 1–3 are located in different points of plate, area 4 covers the heater and it describes its mean temperature. Coordinates of all measuring areas are described by the Table 1. The step responses in the selected areas 1, 2, 3 and 4 are shown in the Figure 4. During calculations these coordinates *x*.. and *y*.. were given relative to *X* and *Y*. For example, xs1 = 75 during calculating elements of C matrix with respect to (29) was equal: x_{s1} = 75/380 = 0.1974.

Tab. 1. Coordinates of measuring areas (in pixels)

Area	x_{s1}	y_{s1}	x_{s2}	y_{s2}
1	50	75	52	77
2	200	100	202	102
3	300	200	302	202
4	120	40	250	60

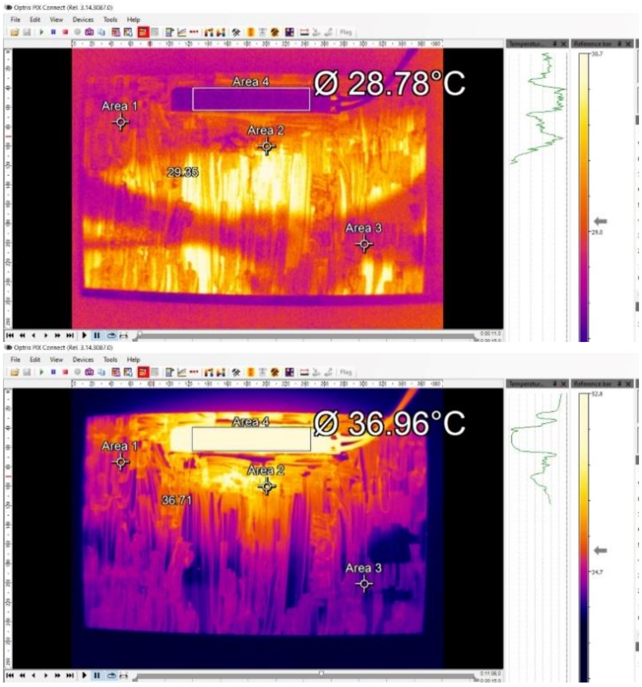


Fig. 3. The steady-state temperature fields for non-heated (top) and heated plate (bottom). The temperature strongly depends on ambient temperature. The colour scale in each case is different

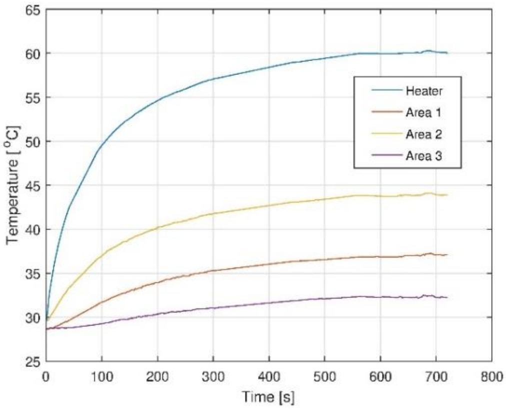


Fig. 4. The step responses of temperature in all tested fields

4.2. The identification of the model

The identification of parameters of the proposed model has been done via minimization of the Mean Square Error (MSE) cost function. This function describes the mean difference between the step responses of the plant and the model at the same time-spatial mesh:

$$MSE = \frac{1}{K}\sum_{k=1}^{K}[y(k) - y_e(k)]^2 \qquad (39)$$

In (39) *K* is the number of all collected samples, *y*(*k*) is the step response of the model, computed using (36), $y_e(k)$ is the experimental response measured in the same place and at the same time instants *k* with the use of thermal camera. The sample time during a step response measurement was equal 1[s]. In each case the mean temperature of the whole area is measured.

The cost function (39) is a function of parameters *α*, *β*, a_w and R_a. Its optimization was done with the use of the MATLAB function *fminsearch*, the step response was calculated using finite-dimensional formula (36). Calculations were done using the following values of both orders: *M* = *N* = 3 and *M* = *N* = 5.

Results are given in the tables 2. The comparison of step responses model vs experiment for M = N = 5 is presented in the Figure 5.

The results in tables 2 and 3 show that the use of 3'rd order model assures practically the same accuracy in the sense of cost function (39) as the use of 5'th order model. This result is confirmed by the convergence analysis presented in the next subsection.

Next, if we recall the results from paper (24), the accuracy of the proposed model is practically the same as the accuracy of 8'th order model with integer order $\beta = 2$.

Tables 4 and 5 compare the values of cost function (39) for all considered models: integer order, fractional order along the time and fractional order along both coordinates.

Tab. 2. Identified parameters of the model for $M = N = 3$

Area	α	β	a_w	R_a	*MSE (39)*
1	1.0794	1.8138	0.0033	0.0032	0.0110
2	0.9356	1.6167	0.0538	0.0089	0.0217
3	1.4878	1.8712	0.0208	0.0003	0.0059
4	0.8156	2.0028	0.0100	0.0235	0.0627

Tab. 3. Identified parameters of the model for $M = N = 5$

Area	α	β	a_w	R_a	*MSE (39)*
1	1.0794	1.8641	0.0032	0.0032	0.0110
2	0.9590	0.3959	0.0357	0.0057	0.0207
3	1.4877	1.8712	0.0208	0.0003	0.0059
4	0.8156	1.2400	0.0098	0.0234	0.0627

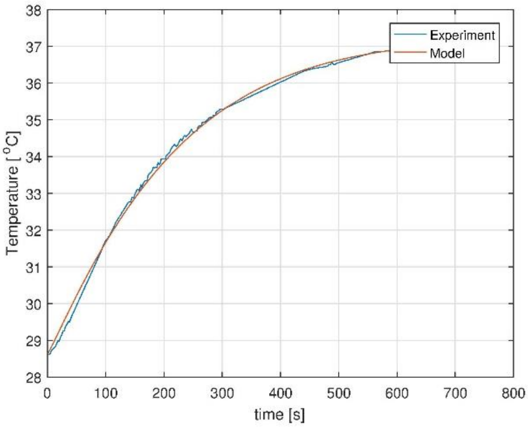
Tab. 4. The cost function MSE for all models and $M = N = 3$

Area	$\alpha = 2, \beta = 2$	$\alpha \in \mathbb{R}, \beta = 2$	$\alpha \in \mathbb{R}, \beta \in \mathbb{R}$
1	0.0233	0.0170	0.0110
2	0.0183	0.0205	0.0217
3	0.0644	0.1429	0.0059
4	1.1448	0.1145	0.0627

Tab. 5. The cost function MSE for all models and $M = N = 5$

Area	$\alpha = 2, \beta = 2$	$\alpha \in \mathbb{R}, \beta = 2$	$\alpha \in \mathbb{R}, \beta \in \mathbb{R}$
1	0.0233	0.0170	0.0110
2	0.0497	0.0205	0.0207
3	0.0665	0.1162	0.0059
4	0.0920	0.1145	0.0627

Quite surprising is the large dispersion of values in the model parameters for points 1, 2 and 3, which are the same in terms of the material. This is probably caused by measurement disturbances related to light reflections and different emissivity of the surface.



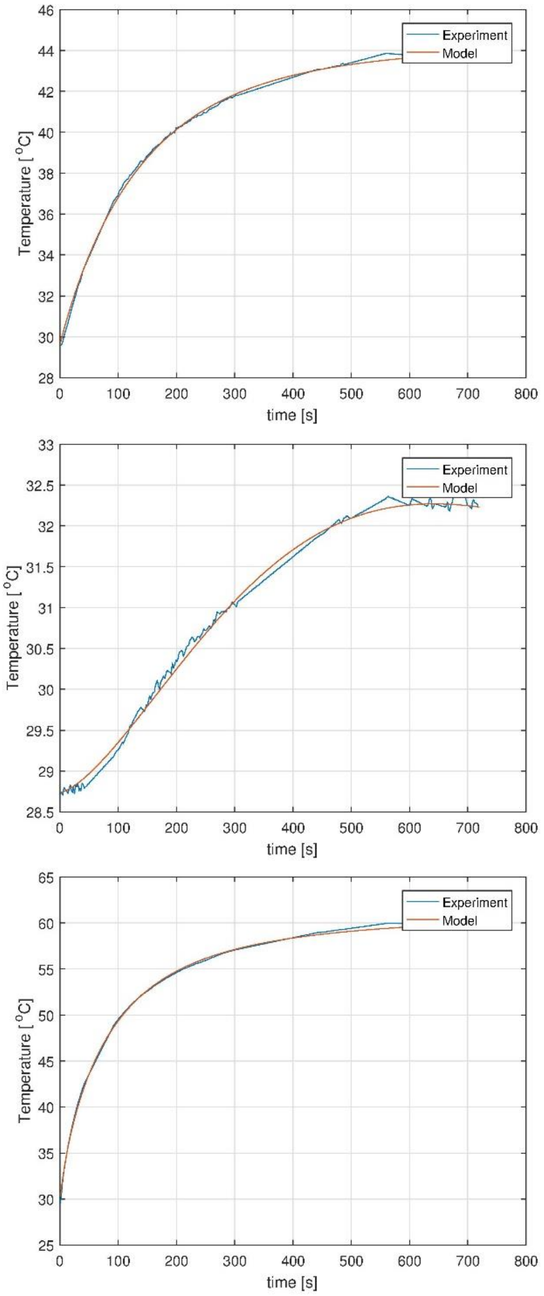


Fig. 5. The step responses of the FO model ($M = N = 5$) vs real plant for area 1(top) to 4 (bottom)

4.3. The numerical analysis of the convergence

The convergence of the proposed model can be estimated using the ROC coefficient expressed by (38). It is a function of place of measurement and a function of both orders of the model M and N. However to simplify the analysis assume that both orders are equal: $M = N$. The value of ROC as a function of order N for all tested places and for parameters given in the table 3 is presented in the Figure 38. From diagrams presented in Figure 38 it can be concluded that the maximum orders $M = N = 4$ assure the good accuracy of the model. Further increasing of orders does not improve the accuracy and increases computational complexity. This is compliant to results of identification, given in the tables 2-5. Next, the ROC depends not only on order N, but also on the location and size of place of measurement.

DOI 10.2478/ama-2023-0020

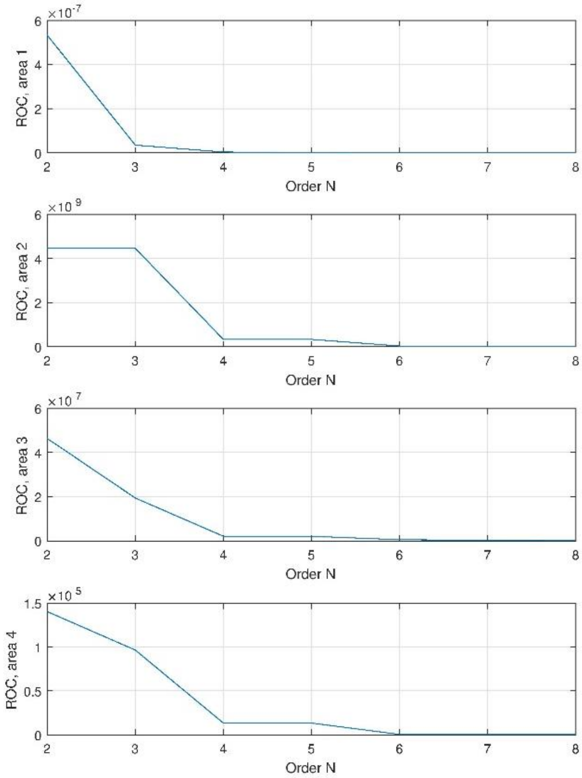


Fig. 6. The Rate of Convergence (38) for all areas and maximum orders *M* = *N* = 8

5. CONCLUSIONS

The main final conclusion from this paper is that the proposed, fully FO model of the distributed parameter system assures better accuracy in the sense of the MSE cost function than the model using fractional derivative only along the time. In addition, the good accuracy can be achieved for relatively low order of model.

The future work will cover deeper analysis of the convergence of the proposed model. Numerical results show, that the convergence depends not only on orders M and N but also on the location and size of area of measurement.

Another important issues are e.g. the positivity analysis as well as the use of a new fractional operators with nonsingular kernel: Atangana-Baleanu and Caputo-Fabrizio to modelling of the presented system.

Next, the use of thermal camera allows to collect many interesting data from different thermal processes, e.g. it allows to investigate thermal processes going in microcontroller system during its work. Such a process can be also described using FO approach and this is planned to describe using the proposed approach.

REFERENCES

1. Podlubny I. Fractional Differential Equations San Diego: Academic Press; 1999.
2. Dzieliński A, Sierociuk D, Sarwas G. Some applications of fractional order calculus. Bulletin of the Polish Academy of Sciences, Technical Sciences. 2010.
3. Caponetto R, Dongola G, Fortuna L, Petra I. Fractional order systems: Modelling and Control Applications. University of California ed. Chua LO, editor. Berkeley: World Scientific Series on Nonlinear Science; 2010.
4. Das S. Functional Fractional Calculus for System Identification and Controls Berlin: Springer; 2010.
5. Gal CG, Warma M. Elliptic and parabolic equations with fractional diffusion and dynamic boundary conditions. Evolution Equations and Control Theory. 2016.
6. Popescu E. On the fractional Cauchy problem associated with a feller semigroup. Mathematical Reports. ; 2010.
7. Sierociuk D, Skovranek T, Macias M, Podlubny I. Diffusion process modelling by using fractional-order models. Applied Mathematics and Computation. 2015.
8. Gómez JF, Torres L, Escobar RF. Fractional derivatives with Mittag-Leffler kernel. Trends and applications in science and engineering Kacprzyk J, editor. Switzerland: Springer; 2019.
9. Boudaoui A, El hadj Moussa Y, Hammouch , Ullah S. A fractional-order model describing the dynamics of the novel coronavirus (covid-19) with nonsingular kernel. Chaos, Solitons and Fractals. 2021; 146(110859):111.
10. Muhammad Farman M, Akgül A, Askar S, Botmart T. Modelling and analysis of fractional order zika model. AIMS Mathematics. 2022.
11. Oprzędkiewicz K, Gawin E, Mitkowski W. Modeling heat distribution with the use of a non-integer order, state space model. International Journal of Applied Mathematics and Com-puter Science. 2016.
12. Oprzędkiewicz K, Gawin E, Mitkowski W. Parameter identification for non-integer order, state space models of heat plant. In MMAR 2016 : 21th international conference on Methods and Models in Automation and Robotics; 2016; Międzyzdroje, Poland. p. 184-188.
13. Oprzędkiewicz K, Stanisławski R, Gawin E, Mitkowski W. A new algorithm for a cfe approximated solution of a discrete-time non integer-order state equation. Bulletin of the Polish Academy of Sciences. Technical Sciences. 2017; 65(4):429-437.
14. Oprzędkiewicz K, Mitkowski W, Gawin E. An accuracy estimation for a non-integer order, discrete, state space model of heat transfer process. In Automation 2017 : innovations in automation, robotics and measurement techniques; 2017; Warsaw, Poland. p. 86-98.
15. Oprzędkiewicz K, Mitkowski W, Gawin E, Dziedzic K. The Caputo vs. Caputo-Fabrizio operators in modelling of heat transfer process. Bulletin of the Polish Academy of Sciences. Technical Sciences. 2018; 66(4):501-507.
16. Oprzędkiewicz K, Gawin E. The practical stability of the discrete, fractional order, state space model of the heat transfer process. Archives of Control Sciences. 2018.
17. Oprzędkiewicz K, Mitkowski W. A memory efficient non in-teger order discrete time state space model of a heat transfer process. International Journal of Applied Mathematics and Computer Science. 2018.
18. Oprzędkiewicz K. Non integer order, state space model of heat transfer process using Atangana-Baleanu operator. Bulletin of the Polish Academy of Sciences. Technical Sciences. 2020; 68(1):43-50.
19. Długosz M, Skruch P. The application of fractional-order models for thermal process modelling inside buildings. Journal of Building Physics. 2015; 1(1):1-13.
20. Ryms M, Tesch K, Lewandowski W. The use of thermal imaging camera to estimate velocity profiles based on tem-perature distribution in a free convection boundary layer. International Journal of Heat and Mass Transfer. 2021.
21. Khan H, Shah R, Kumam P, Arif M. Analytical solutions of fractional order heat and wave equations by the natural transform decomposition method. Entropy. 2019.
22. Olsen-Kettle L. Numerical solution of partial differential equa-tions Brisbane: The University of Queensland; 2011.
23. Al-Omari SK. A fractional Fourier integral operator and its extension to classes of function spaces. Advances in Difference Equations. 2018; 1(195):19.
24. Oprzędkiewicz K, Mitkowski W, Rosół M. Fractional order model of the two dimensional heat transfer process. Energies. 2021.
25. Kaczorek T. Singular fractional linear systems and electrical circuits. International Journal of Applied Mathematics and Computer Science. 2011.

26. Kaczorek T, Rogowski K. Fractional Linear Systems and Electrical Circuits Białystok: Publishing House of the Bialystok University of Technology; 2014.
27. Bandyopadhyay B, Kamal S. Solution, stability and realization of fractional order differential equation. In A Sliding Mode Approach, Lecture Notes in Electrical Engi-neering 317. Switzerland: Springer; 2015. p. 5590.
28. Wyrwas M, Mozyrska D, Girejko E. Comparison of h-difference fractional operators. In Mitkowski W, editor. Advances in the Theory and Applications of Non-integer Order Systems. Switzerland: Springer; 2013. p. 1-178.
29. Berger J, Gasparin S, Mazuroski W, Mende N. An effi-cient two-dimensional heat transfer model for building enve-lopes. An International Journal of Computation and Methodology, Numerical Heat Transfer, Part A: Applications. 2021; 79(3):163194.
30. Moitsheki RJ, Rowjee A. Steady heat transfer through a two-dimensional rectangular straight fin. Mathematical Problems in Engineering. 2011.
31. Yang L, Sun B, Sun X. Inversion of thermal conductivity in two-dimensional unsteady-state heat transfer system based on finite difference method and artificial bee colony. Applied Sciences. 2019.
32. Mitkowski W. Outline of Control Theory Kraków: Publishing House AGH; 2019.
33. Brzek M. Detection and localisation structural damage in selected geometric domains using spectral theory (in Polish). PhD thesis ed. Mitkowski W, editor. Kraków: AGH University of Science and Technology; 2019.
34. Michlin SG, Smolicki CL. Approximate methods for solving differential and integral equations (in Polish) Warszawa: PWN; 1970.

This paper was sponsored by AGH UST project no 16.16.120.773.

Krzysztof Oprzędkiewicz: https://orcid.org/0000-0002-8162-0011

Wojciech Mitkowski: https://orcid.org/0000-0001-5704-8329

Maciej Rosół: https://orcid.org/0000-0003-1176-7904

DOI 10.2478/ama-2023-0021

STANDARD AND FRACTIONAL DISCRETE-TIME LINEAR SYSTEMS WITH ZERO TRANSFER MATRICES

Tadeusz KACZOREK, Andrzej RUSZEWSKI

*Faculty of Electrical Engineering, Bialystok University of Technology, ul. Wiejska 45D, 15-351 Białystok, Poland

t.kaczorek@pb.edu.pl, a.ruszewski@pb.edu.pl

received 8 November 2022, revised 27 December 2022, accepted 28 December 2022

Abstract: The transfer matrix of the standard and fractional linear discrete-time linear systems is investigated. Necessary and sufficient conditions for zeroing of the transfer matrix of the linear discrete-time systems are established. The considerations are illustrated by examples of the standard and fractional linear discrete-time systems.

Key words: fractional, discrete-time, linear system, observability, reachability, zero transfer matrix

1. INTRODUCTION

The notion of controllability and observability and the decomposition of linear systems have been introduced by Kalman [13, 14]. This theory was developed in the following years (e.g. Kailath [12], Klamka [15], Rosenbrock [22]), and became the basic concepts of the modern control theory (e.g. Antsaklis [2], Farina and Rinaldi [5], Poldermann [21]) and modern data-driven system theory and control (see e.g. Dörfler et al. [4], Markovsky and Dörfler [16] and other works cited in this paper). The notion of controllability and observability have been also extended to positive linear systems [5, 9] and electrical circuits [7, 11]. A dynamical system is called positive if its trajectory starting from any non-negative initial state remains forever in the positive orthant for all non-negative inputs. Variety of positive models can be found in electrical engineering, economics, social sciences, biology and medicine, etc.

Fractional calculus is the branch of mathematics that studies integrals and derivatives of non-integer order. Mathematical fundamentals of the fractional calculus are given in various monographs (e.g. Oldham and Spanier [18], Ostalczyk [19], Podlubny [20]). The fractional calculus and its application in many fields of science and engineering have been recently investigated. Numerous applications have been found in mechanics, electricity, chemistry, signal processing, etc. [19, 24, 27]. Fractional-order models of real world phenomena have become more accurate than classical integer order ones. Theory of fractional systems is a rapidly growing field and it concerns properties of processes and control systems, including stability, controllability, observability, realisability, etc. [1, 3, 6, 7, 17, 23, 26, 28]. The standard and positive fractional linear systems have been investigated in monographs by Kaczorek [9] and Kaczorek and Rogowski [11] and the positive linear systems with different fractional orders have been analysed by Kaczorek [8] and Sajewski [25].

Transfer functions (matrices) are very popular in modelling physical phenomena and represent the relation between input and output signals. They are commonly used in the analysis of dynamical systems. In this paper the standard and fractional linear discrete-time systems with zero transfer matrices will be investigated. To the authors' knowledge, this problem for the fractional discrete-time linear systems has not been considered yet. This paper extends the theory of fractional-order systems on this topic.

The remainder of this paper is organised as follows. In Section 2 the basic definitions and theorems concerning the linear discrete-time are given and a class of standard linear discrete-time with zero transfer matrices is investigated. The basic definitions and theorems concerning the fractional linear discrete-time systems and an extension of the results of Section 2 are presented in Section 3. The considerations have been illustrated by linear discrete-time systems. Concluding remarks are given in Section 4.

The following notation will be used: $\Re$ is the set of real numbers; $\Re^{n\times m}$ represents the set of $n\times m$ real matrices; $\Re_+^{n\times m}$ denotes the set of $n\times m$ matrices with non-negative and $\Re_+^n=\Re_+^{n\times 1}$; and I_n is the $n\times n$ identity matrix.

2. STANDARD LINEAR DISCRETE-TIME SYSTEMS

Consider a linear discrete-time system described by the following equations:

$$x_{i+1}=Ax_i+Bu_i,\ i\in Z_+=\{0,\ 1,\ \dots\}, \tag{2.1a}$$

$$y_i=Cx_i, \tag{2.1b}$$

with the initial condition x_0, where $x_i\in\Re^n$, $u_i\in\Re^m$ and $y_i\in\Re^p$ are the state, input and output vectors and $A\in\Re^{n\times n}$, $B\in\Re^{n\times m}$ and $C\in\Re^{p\times n}$.

The transfer matrix of the linear system (2.1) is given by the following equations:

$$T(z)=C[I_nz-A]^{-1}B. \tag{2.2}$$

Definition 2.1. [11, 15] The linear system (2.1) is called reachable in $q\le n$ steps if there exists an input $u_i\in\Re^m$ for $i=0,1,\dots,q\le n-1$ that transfers the state of the system from the initial state $x_0\in\Re^n$ to the given final state $x_f=x_q$ in the q steps.

Theorem 2.1. [11, 15] The linear system (2.1) is reachable in q steps if and only if one of the following equivalent conditions is satisfied:

$$rank[B,\ AB,\ \dots,\ A^{n-1}B] = n, \tag{2.3}$$

$$rank[I_n z - A,\ B] = n \text{ for all } z \in \mathrm{C} \text{ (the field of complex numbers).} \tag{2.4}$$

Definition 2.2. [11] The linear system (2.1) is called observable in q steps if knowing the input u_i and the output y_i in the $q \leq n - 1$ steps it is possible to find its unique initial state x_0.

Theorem 2.2. [11] The linear system (2.1) is observable in q steps if and only if one of the following equivalent conditions is satisfied:

$$rank\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n, \tag{2.5}$$

$$rank\begin{bmatrix} I_n z - A \\ C \end{bmatrix} = n \text{ for all } z \in \mathrm{C}. \tag{2.6}$$

It is well-known [11] that if the linear system (2.1) is unreachable or unobservable then some cancelation of common factors in the numerator and denominator of the transfer matrix (2.2) occurs.

Theorem 2.3. Let the transfer matrix (2.2) of the linear system (2.1) be a zero matrix. Then:

1. if the matrix C is non-zero, then the pair $(A,\ B)$ of the linear system (2.1) is unreachable; and
2. if the matrix B is non-zero, then the pair $(A,\ C)$ of the linear system (2.1) is unobservable.

Proof. It is well-known that the impulse response matrix g_i of the linear system satisfies the condition

$$g_i = CA^{i-1}B = 0 \text{ for } i = 1,\ \dots, n. \tag{2.7}$$

if and only if the transfer matrix (2.2) is a zero matrix.

From Eq. (2.7), we have

$$C[B,\ AB,\ \dots,\ A^{n-1}B] = 0. \tag{2.8}$$

Therefore, if $C \neq 0$ then

$$rank[B,\ AB,\ \dots,\ A^{n-1}B] < n \tag{2.9}$$

and the pair $(A,\ B)$ is unreachable.

Similarly, if $B \neq 0$ then

$$rank\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} < n \tag{2.10}$$

and the pair $(A,\ C)$ is unobservable.

The following theorem gives the necessary and sufficient conditions for zeroing of the transfer matrix (2.2) of the linear system (2.1).

Theorem 2.4. The transfer matrix (2.2) of the unreachable and unobservable linear system (2.1) is a zero matrix if and only if $n \geq m + p$ and

$$CB = 0. \tag{2.11}$$

Proof. It is well known [2, 5, 9, 11] that if the pair $(A,\ B)$ is unreachable then there exists a non-singular matrix $P \in \Re^{n\times n}$ such that

$$PAP^{-1} = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}, PB = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, CP^{-1} = [C_1 \quad C_2] \tag{2.12a}$$

where $A_1 \in \Re^{n_1\times n_1}$, $A_2 \in \Re^{n_1\times n_2}$, $A_3 \in \Re^{n_2\times n_2}$, $B_1 \in \Re^{n_1\times m}$, $C_1 \in \Re^{p\times n_1}$, $C_2 \in \Re^{p\times n_2}$, $n_1 + n_2 = n$,

$$rank[B,\ AB,\ \dots,\ A^{n-1}B] = n_1 \tag{2.12b}$$

and the pair $(A_1,\ B_1)$ is reachable, i.e. $rank[B_1,\ A_1B_1,\ \dots,\ {A_1}^{n_1-1}B_1] = n_1$.

Note that

$$CB = CP^{-1}PB = [C_1 \quad C_2]\begin{bmatrix} B_1 \\ 0 \end{bmatrix} = C_1B_1 = 0, \tag{2.13}$$

since the pair $(A,\ C)$ is unobservable and $C_1 = 0$.

Using (2.2) and (2.12a) we obtain

$$\begin{aligned} T(z) &= C[I_n z - A]^{-1}B = CP^{-1}[PP^{-1}z - PAP^{-1}]^{-1}PB \\ &= [C_1 \quad C_2]\begin{bmatrix} [I_{n_1}z - A_1] & -A_2 \\ 0 & [I_{n_2}z - A_3] \end{bmatrix}^{-1}\begin{bmatrix} B_1 \\ 0 \end{bmatrix} \\ &= [C_1 \quad C_2]\begin{bmatrix} [I_{n_1}z - A_1]^{-1} & * \\ 0 & [I_{n_2}z - A_3]^{-1} \end{bmatrix}\begin{bmatrix} B_1 \\ 0 \end{bmatrix} \\ &= C_1[I_{n_1}z - A_1]^{-1}B_1 = 0 \end{aligned} \tag{2.14}$$

if and only if the condition (2.11) is satisfied, where * denotes a matrix unimportant in these considerations. Therefore, the transfer matrix (2.2) of the unreachable and unobservable linear system (2.1) is a zero matrix if and only if $n \geq m + p$ and the condition (2.11) is satisfied.

Example 2.1. Consider the linear system (2.1) with the matrices

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}, C = [-0.5 \quad 1 \quad 0]. \tag{2.15}$$

The pair $(A,\ B)$ with (2.15) is unreachable and the pair $(A,\ C)$ is unobservable, since

$$rank[B,\ AB,\ A^2B] = rank\begin{bmatrix} 4 & 8 & 16 \\ 2 & 4 & 8 \\ 1 & 3 & 5 \end{bmatrix} = 2 < n = 3 \tag{2.16a}$$

and

$$rank\begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = rank\begin{bmatrix} -0.5 & 1 & 0 \\ 0.5 & -1 & 0 \\ -0.5 & 1 & 0 \end{bmatrix} = 1 < n = 3. \tag{2.16b}$$

The condition (2.11) is also satisfied, since

$$CB = [-0.5 \quad 1 \quad 0]\begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} = 0. \tag{2.17}$$

In this case the matrix P has the form

$$P = \begin{bmatrix} 1 & -2 & 0 \\ 0.5 & 0 & 0 \\ -0.5 & 1 & 0 \end{bmatrix} \tag{2.18}$$

and

$$\bar{A} = PAP^{-1} = \begin{bmatrix} -1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix}, \bar{B} = PB = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix},$$

$$\bar{C} = [0 \quad 0 \quad 1]. \tag{2.19}$$

The transfer function of the linear system with (2.15) has the form

$$T(z) = C[I_3 z - A]^{-1}B =$$

$$[-0.5 \quad 1 \quad 0]\begin{bmatrix} z-1 & -2 & 0 \\ -1 & z & 0 \\ -1 & 0 & z+1 \end{bmatrix}^{-1}\begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} = 0 \qquad (2.20)$$

and

$$\bar{T}(z) = \bar{C}[I_3 z - \bar{A}]^{-1}\bar{B} =$$

$$[0 \quad 0 \quad 1]\begin{bmatrix} z+1 & -2 & 0 \\ 0 & z-2 & -1 \\ 0 & 0 & z+1 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = 0. \qquad (2.21)$$

This confirms Theorem 2.4.

3. FRACTIONAL LINEAR DISCRETE-TIME SYSTEMS

Consider the fractional discrete-time linear system described by the equations

$$\Delta^{\alpha} x_{i+1} = Ax_i + Bu_i, \; 0 < \alpha < 1, \; i \in Z_+, \qquad (3.1a)$$

$$y_i = Cx_i, \qquad (3.1b)$$

where

$$\Delta^{\alpha} x_i = \sum_{j=1}^{i} c_j x_{i-j} \qquad (3.2a)$$

$$c_j = (-1)^j \binom{\alpha}{j}, \quad \binom{\alpha}{j} = \begin{cases} 1 & \text{for} \quad j = 0 \\ \frac{\alpha(\alpha-1)\ldots(\alpha-j+1)}{j!} & \text{for} \quad j = 1,2,\ldots \end{cases} \qquad (3.2b)$$

is the fractional α-order difference of x_i and $x_i \in \Re^n$, $u_i \in \Re^m$ and $y_i \in \Re^p$ are the state, input and output vectors, $x_0$is the initial condition and $A \in \Re^{n\times n}$, $B \in \Re^{n\times m}$ and $C \in \Re^{p\times n}$. Eq. (3.2a) is the definition of the Grünwald–Letnikov fractional derivatives.

Substitution of Eq. (3.2) into Eq. (3.1a) yields

$$x_{i+1} = A_\alpha x_i - \sum_{j=2}^{i+1} c_j x_{i-j+1} + Bu_i, i \in Z_+, \qquad (3.3a)$$

where

$$A_\alpha = A + I_n \alpha. \qquad (3.3b)$$

The solution of Eq. (3.1a) is given by

$$x_i = \Phi_i x_0 + \sum_{j=0}^{i-1} \Phi_{i-j-1} Bu_j, \qquad (3.4a)$$

where the matrices Φ_i are defined by

$$\Phi_{i+1} = \Phi_i A_\alpha + \sum_{j=2}^{i+1} (-1)^{j+1} \binom{\alpha}{j} \Phi_{i-j+1}, \qquad \Phi_0 = I_n,$$
$$i = 0,1,\ldots \qquad (3.4b)$$

It is well-known [9] that if $0 < \alpha < 1$ then

1) $-c_j > 0$ for $j = 1,2,\ldots$ (3.5a)

2) $\sum_{j=1}^{n} c_j = -1$ (3.5b)

The transfer matrix of the fractional linear discrete-time system is given by

$$T(\bar{z}) = C[I_n \bar{z} - A_\alpha]^{-1}B, \qquad (3.6a)$$

where

$$\bar{z} = z - c_\alpha, \; c_\alpha = \sum_{j=2}^{i+1} (-1)^{j-1} \binom{\alpha}{j} z^{1-j}. \qquad (3.6b)$$

Definition 3.1. [9] The fractional linear discrete-time system (3.1) is called reachable in q steps if for any given final state $x_f \in \Re^n$ there exists an input sequence u_i for $i \in [0, \; q]$ that steers the state of the system from $x_0 = 0$ to the given final state $x_q = x_f$.

Theorem 3.1. [9] The fractional linear discrete-time system (3.1) is reachable if and only if one of the equivalent conditions is satisfied:

1) $rank[B, \; A_\alpha B, \; \ldots, \; {A_\alpha}^{q-1}B] = n,$ (3.7a)

2) $rank[I_n z - A_\alpha, \; B] = n$ *for all* $z \in$ C
(the field of complex numbers). (3.7b)

Definition 3.2. [9] The fractional linear discrete-time system (3.1) is called observable in q steps if knowing the input $u_i \in \Re^m$ and the output $y_i \in \Re^p$ in the q steps it is possible to find its unique initial state $x_0 \in \Re^n$.

Theorem 3.2. [9] The fractional linear discrete-time system (3.1) is observable in q steps if and only if one of the following equivalent conditions is satisfied:

$$\textit{1) } rank \begin{bmatrix} C \\ CA_\alpha \\ \vdots \\ CA_\alpha^{q-1} \end{bmatrix} = n, \qquad (3.8a)$$

$$\textit{2) } \; rank \begin{bmatrix} I_n z - A_\alpha \\ C \end{bmatrix} = n \textit{ for all } z \in \text{C}. \qquad (3.8b)$$

Theorem 3.3. The transfer matrix (3.6) of the unreachable and unobservable linear system (3.1) is a zero matrix if and only if $n \geq m + p$ and

$$CB = 0. \qquad (3.9)$$

Proof. It is well-known that the transfer matrix $T(z) = 0$ if and only if the corresponding matrix of impulse responses $g_i = 0$. If the pair $(A_\alpha, \; B)$ is unreachable then the pair by similarity transformation can reduced to the form

$$A_\alpha = \begin{bmatrix} A_{1\alpha} & A_{2\alpha} \\ 0 & A_{3\alpha} \end{bmatrix}, \; A_{1\alpha} \in \Re^{n_1\times n_1}, A_{3\alpha} \in \Re^{n_2\times n_2}, \; B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix},$$
$$B_1 \in \Re^{n_1\times m}, n_1 + n_2 = n \qquad (3.10a)$$

and

$$C = [C_1 \quad C_2], C_1 \in \Re^{p\times n_1}, C_2 \in \Re^{p\times n_2}. \qquad (3.10b)$$

In this case, from Eq. (3.4b), it follows that

$$\Phi_q = \begin{bmatrix} \Phi_{1\alpha} & \Phi_{2\alpha} \\ 0 & \Phi_{3\alpha} \end{bmatrix}, \Phi_{1\alpha} \in \Re^{n_1\times n_1}, \Phi_{3\alpha} \in \Re^{n_2\times n_2}, \qquad (3.11)$$

$$g(q) = C\Phi_q B = [C_1 \quad C_2]\begin{bmatrix} \Phi_{1\alpha} & \Phi_{2\alpha} \\ 0 & \Phi_{3\alpha} \end{bmatrix}\begin{bmatrix} B_1 \\ 0 \end{bmatrix} = C_1 \Phi_{1q} B_1 =$$
$$0, \qquad (3.12)$$

since the system is unobservable and $C_1 = 0$.

Therefore, the transfer matrix is a zero matrix if the system is an unreachable and unobservable system and the condition (3.9) is satisfied.

Note that the Theorem 3.3 can be also proved in a manner similar to Theorem 2.4.

Example 3.1. Consider the fractional discrete-time linear system (3.1) for $\alpha = 0.4$ with the matrices

$$A = \begin{bmatrix} 0.4 & 0.1 & 0.2 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.4 \end{bmatrix}, \quad B = \begin{bmatrix} 0.4 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad C = [0 \quad 0 \quad 1]. \qquad (3.13)$$

In this case we have

$$A_\alpha = A + I_3\alpha = \begin{bmatrix} 0.6 & 0.1 & 0.2 \\ 0 & 0.7 & 0 \\ 0 & 0 & 0.8 \end{bmatrix} = \begin{bmatrix} A_{1\alpha} & A_{2\alpha} \\ 0 & A_{3\alpha} \end{bmatrix},$$

$$A_{1\alpha} = \begin{bmatrix} 0.6 & 0.1 \\ 0 & 0.7 \end{bmatrix}, A_{3\alpha} = [0.8]. \tag{3.14}$$

Note that the pair $(A_\alpha,\ B)$

$$rank[B, \quad A_\alpha B]$$
$$= rank \begin{bmatrix} 0.4 & 0 & 0.24 & 0.1 \\ 0 & 1 & 0 & 0.7 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 2 < n = 3 \tag{3.15}$$

is unreachable and the pair $(A_\alpha,\ C)$

$$rank \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = rank \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0.4 \\ 0 & 0 & 0.32 \end{bmatrix} = 1 < n = 3 \tag{3.16}$$

is unobservable.

The transfer matrix of the system with Eq. (3.13) can be given as

$$T(\bar{z}) = C[I_3\bar{z} - A_\alpha]^{-1}B$$
$$= [0 \quad 0 \quad 1] \begin{bmatrix} \bar{z} - 0.6 & -0.1 & -0.2 \\ 0 & \bar{z} = 0.7 & 0 \\ 0 & 0 & \bar{z} - 0.8 \end{bmatrix}^{-1} \begin{bmatrix} 0.4 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$
$$= [0 \quad 0]. \tag{3.17}$$

This confirms Theorem 3.3.

From a comparison of the considerations presented in Sections 2 and 3, we have the following important conclusion.

Conclusion 3.1. The zeroing of the transfer matrix of the linear systems is invariant under the order of the differential equations describing the linear discrete-time systems.

The above considerations can be extended to normal linear discrete-time systems [10].

4. CONCLUDING REMARKS

The problem of zeroing of the transfer matrix of standard and fractional linear discrete-times has been investigated. Necessary and sufficient conditions for the zeroing of the transfer matrices of linear discrete-time systems have been established (Theorem 2.3 and 2.4).These conditions have been extended to fractional linear discrete-time systems (Theorem 3.3). It has been shown that the necessary and sufficient conditions are invariant under the fractional orders of the linear discrete-time systems. The considerations have been illustrated by standard and fractional examples of linear discrete-time systems. The considerations can be extended to linear discrete-time systems characterised by different orders.

REFERENCES

1. Abu-Saris R, Al-Mdallal Q. On the asymptotic stability of linear system of fractional-order difference equations. Fract. Calc. Appl. Anal. 2013; 16: 613-629.
2. Antsaklis E, Michel A. Linear Systems. Birkhauser, Boston, 2006.
3. Cermak J, Gyori I, Nechvatal L. On explicit stability conditions for a linear fractional difference system. Fract. Calc. Appl. Anal. 2015; 18: 651-672.
4. Dörfler F, Coulson J, Markovsky I. Bridging direct & indirect data-driven control formulations via regularizations and relaxations. Trans. Automat. Contr., 2023.
5. Farina L, Rinaldi S. Positive Linear Systems: Theory and Applications. J. Wiley & Sons, New York, 2000.
6. Goodrich C, Peterson A. Discrete Fractional Calculus. Springer, Cham, 2015.
7. Kaczorek T. Positivity and reachability of fractional electrical circuits. Acta Mechanica et Automatica. 2011; 5(2): 42-51.
8. Kaczorek T. Positive linear systems consisting of n subsystems with different fractional orders. IEEE Trans. Circuits and Systems. 2011; 58(6): 1203-1210.
9. Kaczorek T. Selected Problems of Fractional Systems Theory. Berlin, Germany: Springer-Verlag, 2011.
10. Kaczorek T. Normal positive electrical circuits. IET Control Theory Appl. 2015; 9(5): 691–699.
11. Kaczorek T, Rogowski K. Fractional Linear Systems and Electrical Circuits. Studies in Systems, Decision and Control, Vol. 13, Springer, 2015.
12. Kailath T. Linear systems. Prentice Hall, Englewood Cliffs, New York, 1980.
13. Kalman R. Mathematical description of linear systems. SIAM J. Control. 1963; 1(2): 152-192.
14. Kalman R. On the general theory of control systems. Proc. First Intern. Congress on Automatic Control. London, UK: Butterworth, 1960; 481-493.
15. Klamka J. Controllability of Dynamical Systems. Kluwer, Dordrecht, Netherlands, 1981.
16. Markovsky I, Dörfler F. Behavioral systems theory in data-driven analysis, signal processing, and control. Annual Reviews in Control. 2021; 52: 42–64.
17. Mozyrska D, Wyrwas M. The Z-transform method and delta type fractional difference operators. Discrete Dyn. Nat. Soc. 2015; (2-3): 1-12.
18. Oldham K, Spanier J. The fractional calculus: integrations and differentiations of arbitrary order. New York, USA: Academic Press, 1974.
19. Ostalczyk P. Discrete Fractional Calculus: Applications in Control and Image Processing; Series in Computer Vision, World Scientific Publishing, Hackensack, New York, 2016.
20. Podlubny I. Fractional differential equations. San Diego, USA: Academic Press, 1999.
21. Poldermann JW, Willems J.C. Introduction to Mathematical Systems Theory. Texts in Applied Mathematics, vol. 26. Springer, New York, NY, 1998.
22. Rosenbrock H. State-space and multivariable theory. New York, USA: J. Wiley, 1970.
23. Ruszewski A. Stability of discrete-time fractional linear systems with delays, Archives of Control Sciences. 2019; 29(3): 549-567.
24. Sabatier J, Agrawal OP, Machado JAT. Advances in Fractional Calculus, Theoretical Developments and Applications in Physics and Engineering. Springer, London, 2007.
25. Sajewski Ł. Stabilization of positive descriptor fractional discrete-time linear systems with two different fractional orders by decentralized controller. Bull. Pol. Acad. Sci. Techn. 2017; 65(5): 709-714.
26. Song TT, Wu GC, Wei JL. Hadamard fractional calculus on time scales, Fractals. 2022; 30(7), 2250145.
27. Sun HG, Zhang Y, Baleanu D, Chen W, Chen YQ. A new collection of real world applications of fractional calculus in science and engineering. Commun. Nonlinear Sci. Numer. Simul. 2018; 64: 213-231.
28. Wu GC, Abdeljawad T, Liu J, Baleanu D, Wu KT. Mittag-Leffler stability analysis of fractional discrete-time neural networks via fixed point technique. Nonlinear Analysis: Model. Contr. 2019; 24: 919-936.

Tadeusz Kaczorek: https://orcid.org/0000-0002-1270-3948

Andrzej Ruszewski: https://orcid.org/0000-0003-0095-6486

EXPERIMENTAL AND NUMERICAL STUDY OF THE EFFECT OF THE PRESENCE OF A GEOMETRIC DISCONTINUITY OF VARIABLE SHAPE ON THE TENSILE STRENGTH OF AN EPOXY POLYMER

Khalissa SAADA*,, Salah AMROUNE*,**, Moussa ZAOUI*,**
Amin HOUARI***, Kouider MADANI***, Amina HACHAICHI******

*Department of Mechanical Engineering, University of Mohamed Boudiaf-M'Sila, M'sila, Algeria
**Laboratoire de Matériaux et Mécanique des Structures (LMMS), Université de M'sila, M'sila, Algérie
***Laboratoire de Matériaux, et Mécanique des Structures (LMMS), Université SBA. Sidi Bel Abesse, Algérie
**** Faculty of Technology, M'hamed Bougara, University, Boumerdes 35000, Algeria

khalissa.saada@univ-msila.dz, salah.amroune@univ-msila.dz, moussa.zaoui@univ-msila.dz
houari.latif2016@gmail.com, koumad10@yahoo.fr, a.hachaichi@univ-boumerdes.dz

received 25 October 2022, revised 2 January 2023, accepted 2 January 2023

Abstract: The presence of geometric discontinuity in a material reduces considerably its resistance to mechanical stresses, therefore reducing the service life of materials. The analysis of structural behaviour in the presence of geometric discontinuities is important to ensure the proper use, especially if it is regarding a material of weak mechanical properties such as a polymer. The objective of the present work is to analyse the effect of the notch presence of variable geometric shapes on the tensile strength of epoxy-type polymer specimens. A series of tensile tests were carried out on standardised specimens, taking into account the presence or absence of a notch. Each series of tests contains five specimens. Two notch shapes were considered: circular (hole) and elliptical. The experimental results in terms of stress–strain clearly show that the presence of notches reduces considerably the resistance of the material, where the maximum stress for the undamaged specimen was 41.22 MPa and the lowest stress for the elliptical-notched specimen was 11.21 MPa. A numerical analysis by the extended finite element method (XFEM) was undertaken on the same geometric models; in addition, the results in stress–strain form were validated with the experimental results. A remarkable improvement was obtained (generally an error within 0.06%) for strain, maximum stress, Young's modulus and elongation values. An exponential decrease was noted in the stress, strain, and Young's modulus in the presence of a notch in the material.

Key words: Tensile test, Hole-notched, Elliptical-notched, XFEM, Finite element method

1. INTRODUCTION

In the recent years, scientific researchers and industrial experts have focused their efforts on biocomposite materials, due to their properties of being sustainable, renewable, biodegradable and biocompatible [1-4]. Biocomposite materials can be used in different areas of applications, especially in light structures that have a bolted assembly before using. It is necessary to determine their mechanical properties through the different experimental techniques used under different types of loading such as traction [5–7], compression [8, 9], torsion [10–12], fatigue [13–15] and impact [16]. The study of the geometry effect of the tensile specimens on the mechanical properties has gained importance from several authors. Notably, Baykan et al. [17] conducted an experimental study on the effect of hole size and position on the mechanical properties of tensile specimens using peridynamic (PD) theory and compared them numerically using code developed on MATLAB and ANSYS. The authors noticed that PD can accurately capture fracture stress, strain and hole interaction in composite laminates.

In another study presented by Hao et al. [18], plastic specimens reinforced with Kenaf fibres containing holes of different diameters were produced. These specimens were tested in uniaxial tensile, open hole tensile, tension at different strain rates, bending and in-plane shear. The obtained results indicated that those treated at a higher temperature of 230°C, but at a shorter time of 60 s, had the best mechanical performance. Also, the linear elastic finite element (FE) model of KPNCs agreed well with the experimental results in the valid strain range of 0%–0.5% for the uniaxial tensile test and 0%–1% for the bending test. Tensile specimens according to ASTM D1822 were manufactured using a 3D printer by Galeta et al. [19]. In this study, the authors determined the impact of three different lattice-like hollow structures considered as honeycombs, drills and scratches on the tensile strength of 3D printed samples. The test samples were prepared on a 3D printer, with variations of the geometric structure internal. The results of the tensile test revealed that the honeycomb structure and the structured samples exhibited the greatest strength.

Khosravani et al. [20] found that the tensile strength of samples of polylactic acid (PLA) and acrylonitrile butadiene styrene (ABS) indicates that increasing the diameter of the hole leads to a decrease in the strength of the part. The finite element method (FEM) has been generally used for predicting the mechanical behaviour of composite materials [21–23], while Mohammadi et al. [24] established standard open-hole tensile (OHT) laminated composites using the FEM to quantify the damage mechanism. Moreover, Ghezzo–Fabrizia et al. [25] presented a numerical and experimental study conducted on T300/epoxy carbon fibre thin laminates with multiple cutouts subjected to in-plane loads. Liao and Adanur [26] studied a new model, which is based on the geometric modelling of woven and braided fabric structures in three dimensions using a computer-aided geometric design (CAD) technique. Recently, Recement Bogrekci et al. [27] analysed the impact of modified PLA specimen geometries on structural strength using FE analysis. Moreover, several types of adhesives such as epoxy, polyester resin, phenolic and polyurethane have been used for manufacturing of sandwich joints [28, 29].

Most of these studies did not take into account in their analysis the effect of the presence of notches on a material with weak mechanical properties such as polymer. Our work is a part of this context. The aim of this paper is to propose a mixed extended finite element method (XFEM) formulation for the elasto-plastic analysis of stress and strain. The objective is to see the effect of the presence of a geometric discontinuity on the tensile strength of an epoxy type-polymer and to see how its mechanical properties evolve compared with the shape of the notch, and on the other hand, to try to use the XFEM to analyse the variation of the stress applied according to the deformation. Tensile tests were carried out on specimens in the presence of different shapes of geometric discontinuity (hole, elliptical notch). The effect of these geometric shapes (hole and ellipse) on the mechanical properties of the polymer was examined and compared with undamaged specimens. The experimental results were compared with those obtained with the FE numerical analysis using the ABAQUS software using XFEM. The effect of the geometric shape of the notch was evaluated in terms of the maximum stress and maximum deformation. On the other hand, the calculation of the stress concentration factor K_t in the different epoxy specimens has been highlighted.

2. MATERIALS AND METHODS

2.1. Specimen geometry

The epoxy polymer specimens intended for the tensile test are shown in Fig. 1. The dimensions are standardised according to the ASTM D638-14 standard (Fig. 1a). Two notch shapes were considered: a circular shape (Fig. 1b) and an elliptical shape (Fig. 1c). The presence of the notch aims to concentrate the stresses and locate a plastification, which will be a source of initiation of the damage.

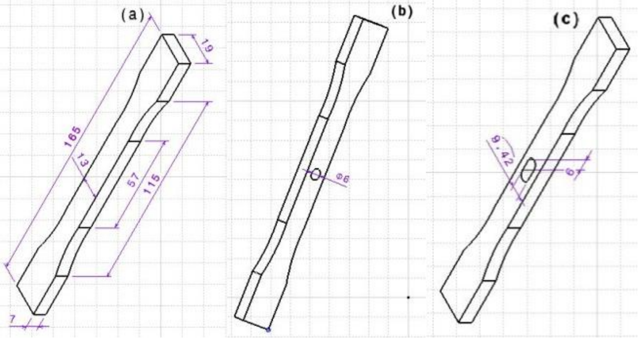


Fig 1. (a) Undamaged specimen, (b) specimen with hole and (c) specimen with elliptical notch

The dimensions of the samples were 165 × 19 × 7 mm3, where the hole diameter is about 6 mm and for the elliptical notch the dimensions are r^1 = 3 mm, r^2 = 6 mm.

2.2. Test procedure

The samples were connected to the handles of the device and then the displacement was measured by the computer program. The stress values relate to the force (F) and section area (S) of samples. Also, the force–displacement curves of the samples are recorded followed by incorporation of the total energy absorption (EA) into the force–displacement curve with the following relation [30]:

$$EA = \int_0^S F dS = F_m S \tag{1}$$

For displacement u and field q to be used for the XFEM of the model [31].

$$U^h = \sum_{i\in\mathcal{M}} N_i u_i^{(\mathcal{M})} + \sum_{i\in\mathcal{H}} \widetilde{N}_i H u_i^{\mathcal{H}} + \sum_{i\in\mathcal{K}} \sum_{j=1}^{4} \widetilde{N}_i F_j^{(u)} u_{ij}^{(\mathcal{K})} = H\hat{u} \tag{2}$$

$$\mathrm{q}^\mathrm{h} = \sum_{\mathrm{i}\in\mathcal{M}} \mathrm{N}_\mathrm{i} \mathrm{q}_\mathrm{i}^{\mathcal{M}} + \sum_{i\in\mathcal{H}} \widetilde{N}_i H q_i^{\mathcal{H}} + \sum_{i\in\mathcal{K}} \sum_{j=1}^{2} \widetilde{N}_i F_j^{(q)} q_{ij}^{(\mathcal{K})} = H_q \hat{q} \tag{3}$$

where: i and j are a numbered node, $\hat{\mathrm{u}}, \hat{\mathrm{q}}$ are the global vectors, and $\mathrm{N_i}$ corresponds to the functions of quadratic shape approximation or associated with the linear continuum, while $\widetilde{\mathrm{N}}_\mathrm{i}$ is the linear shape function of the FEs which construct the partition of unity. $\mathrm{u}_\mathrm{i}^{(\mathcal{M})}, \mathrm{u}_\mathrm{i}^{\mathcal{H}}, \mathrm{q}_\mathrm{i}^{\mathcal{M}}, \mathrm{u}_\mathrm{ij}^{(\mathcal{K})}, \mathrm{q}_\mathrm{i}^{\mathcal{H}}$ and $\mathrm{q}_\mathrm{ij}^{(\mathcal{K})}$ denote the unknowns at i, H and $\mathrm{F}_\mathrm{j}^{(\mathrm{q})}$ are the functions on the sides.

In this work, we used the XFEM technique which allows us to initiate and predict the propagation of the crack in the structure to be damaged. Our structure is fixed in their extremities in order to quickly cause the damage, adding to this the mode of loading that appears as an important effect and quickly causes the damage in our analysis. Not only does it accumulate the load to quickly cause the damage, but it also gives us a broad understanding of the behaviour of our structure for analysis. The XFEM technique is implemented in the standard ABAQUS calculation code. In technique XFEM the damage is presented in the following forms:

$$u(x) = \sum_{i=N} N_i(x) u_i + \sum_{i\in N_d} N_i(x) H(x) a_i + \sum_{i\in N_p} N_i(x) \left(\sum_{j=1}^{4} F_i(x) b_i^j\right) \tag{4}$$

where N is all the nodes of the mesh and u_i is the classical degree of freedom at node I; N_i (x) are the classical FE shape function associated with node I, where a and b are the corresponding degrees of freedom, H(x) is a Heaviside type enrichment function and $\mathrm{F_i(x)}$ Enrichment functions represent the singularity in the vicinity of the crack front as follows:

$$\{F_i(x)\} = \left\{\sqrt{r} sin\frac{\theta}{2}, \sqrt{r} cos\frac{\theta}{2}, \sqrt{r} sin\frac{\theta}{2} sin\frac{\theta}{2}, \sqrt{r} cos\frac{\theta}{2} sin\frac{\theta}{2}\right\} \tag{5}$$

In the deformation zone, there is a point (yield point) which, if exceeded by the stress value, does not return to its original value [32] as follows:

$$\sigma = \sigma_e(1+\varepsilon) \tag{6}$$

$$\int_0^\varepsilon d\varepsilon = \int_{l_0}^{l_i} \frac{dl}{l} \tag{7}$$

$$\varepsilon_{true} = ln\left(\frac{l_i}{l_0}\right) \tag{8}$$

$$\varepsilon_{true} = ln\left(\frac{l_0+\Delta l}{l_0}\right) \tag{9}$$

$$\varepsilon_{true} = ln(1+\varepsilon) \tag{10}$$

Isotropic materials have the same mechanical properties in all directions in all points of the material as follows:

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12} & C_{13} & 0 & 0 & 0\\ & C_{11} & 0 & 0 & 0 & 0\\ & & C_{11} & 0 & 0 & 0\\ & & & \frac{C_{11}-C_{12}}{2} & 0 & 0\\ & & & & \frac{C_{11}-C_{12}}{2} & 0\\ & \text{sym} & & & & \frac{C_{11}-C_{12}}{2}\end{bmatrix}\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\end{bmatrix} \tag{11}$$

For the ellipse shape which is parallel to the axis and the plate α = 0 or α = π/2, which indicated the concentration of stresses around the ellipse and the concentration of stress at the edge of the given hole [33]:

$$\bar{\sigma} = \sigma^{\infty}\frac{1-m^2-2m+2\cos 2\theta}{1-2m\cos 2\theta+m^2} \tag{12}$$

so that $0 \leq m \leq 1$.

The maximum stress equals:

$$\sigma_{max} = k_t\sigma_{nom} \tag{13}$$

The stress concentration factor kt:

$$k_t = 1 + 2\sqrt{\frac{a}{R}} \tag{14}$$

R: Curve radius

2.3. Materials tested

Three sample types of epoxy resin with and without fibres were prepared, and the notches were created in the shape of elliptical holes (Fig. 2). During the manufacturing process, the samples were exposed to a polymerisation temperature of 70° and for 5 h in order to improve their mechanical properties. The testing conditions were as follows: these samples will be subjected to a tensile test under a displacement speed estimated at F = 1 mm/min and the diameter of the hole was droll = 6 mm for hole-notched specimen and the dimensions of the samples were 165 × 19 × 7 mm^3.

2.4. Design of experiments

The purpose of determining the tensile curves for the different specimens was to take into account the location of rupture of the specimens by the presence of notches. The different geometric notch shapes have the purpose of defining the sensitivity of the specimen with respect to the tensile loading. For this, five specimens were considered for each type, as shown in Fig. 2.

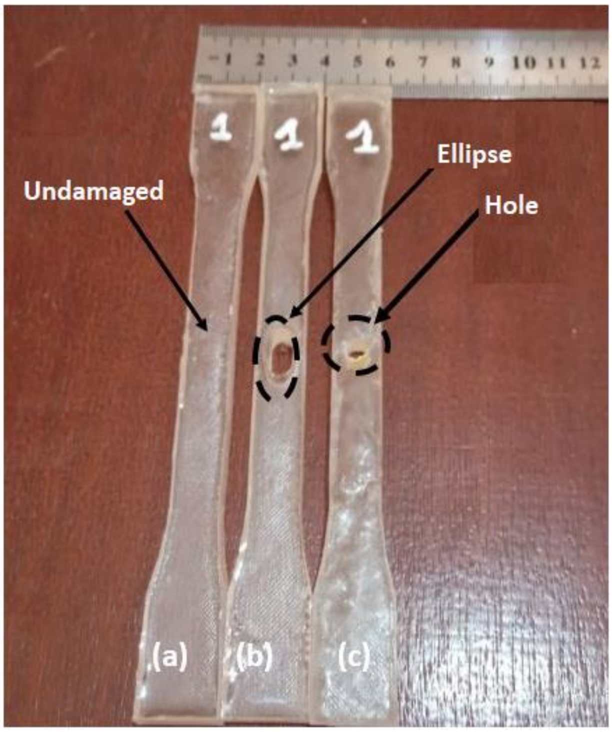


Fig. 2. (a) The undamaged sample of epoxy. (b) Epoxy specimens with ellipse. (c) Epoxy specimens with hole

3. EXPERIMENTAL ANALYSIS

The experimental results obtained refer to five samples for each of the specimens: complete Fig. 3(a), hole-notched specimen Fig. 3(b) and elliptical-notched specimen Fig. 3(c). Obtaining a reproducibility of the results was a little difficult, especially for the samples in the presence of elliptical notch. Machining this shape was a bit difficult. For the other specimens the shape of the curves was the same. We notice a variation in the tensile curves for the three specimens, where the mechanical properties of resistance (stress and Young's modulus) as well as the deformation were influenced by the presence of notches. A considerable drop in these properties is noted compared with the results of the solid samples (Fig. 4). The maximum stress is dropped from 40 MPa to 21.06 MPa (a reduction of almost 50%) for the samples with holes and 11.21 MPa (a reduction of almost 75%) for the samples with elliptical notch.

The experimental results in Fig. 4, including Young's modulus and maximum stress value, indicate that the highest value for them is at the undamaged specimen (1,793.80 MPa and 41.22 MPa, respectively). The value of the hole-notched specimen is lower with a Young's modulus of 1,423.36 MPa and a maximum stress equal to 21.06 MPa; the smallest of them is the elliptical-notched specimen with a maximum stress of 11.21 MPa and a Young's modulus of 547.59 MPa.

The experimental results indicate that the elasticity area of the undamaged specimen is larger than the elasticity area of the hole-notched specimen and the elliptical-notched specimen.

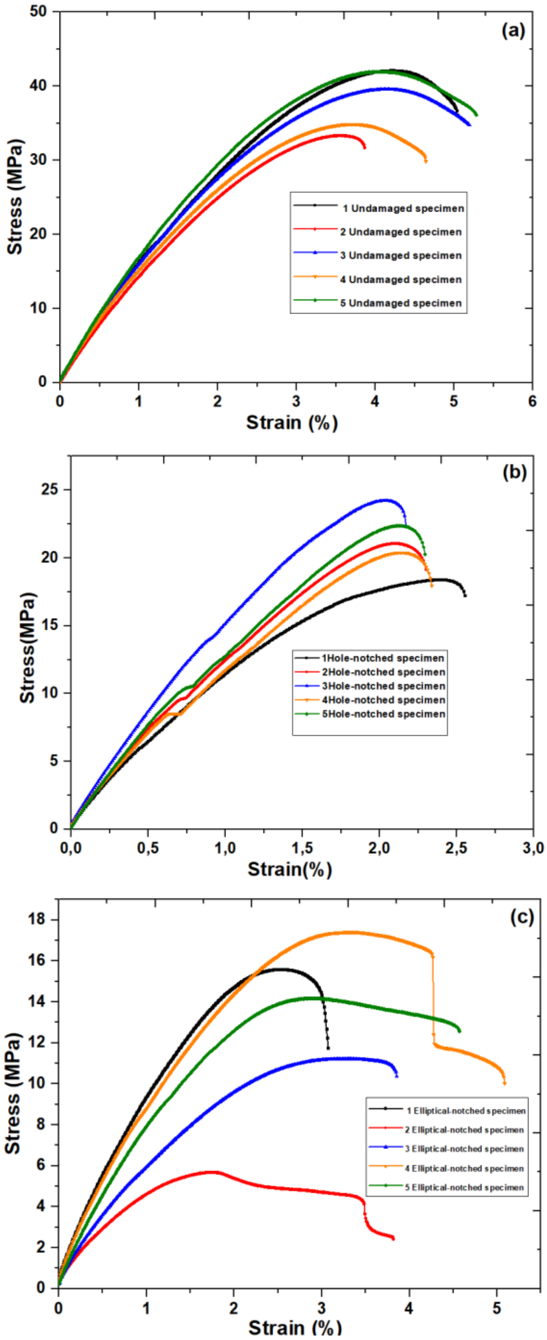
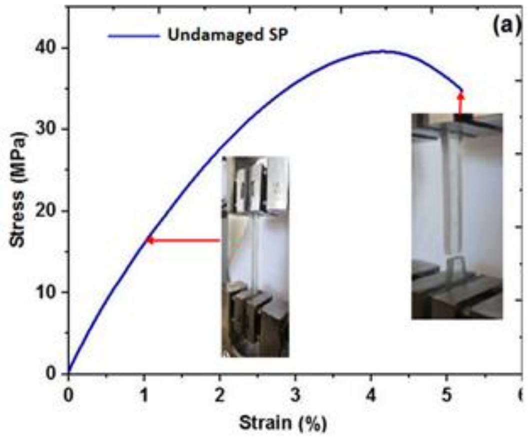


Fig. 3. Different experimental results for the five samples. (a) The five undamaged specimens. (b) The five hole-notched specimen. (c) The five elliptical-notched specimens



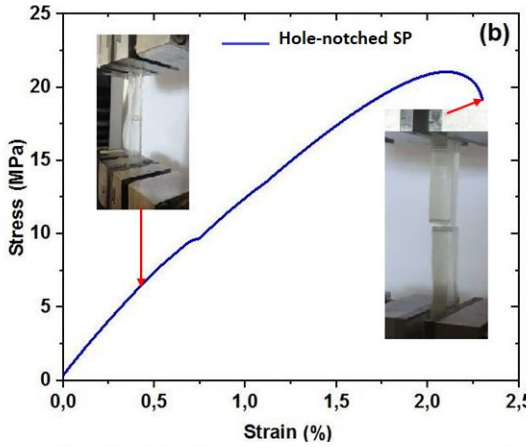


Fig. 4. Comparison of the experimental results of the mean of the three specimens: (a) The undamaged sample. (b) The hole-notched specimen. (c) The elliptical-notched specimen

3.1. Analysis of tensile test specimens

The three-epoxy specimens were analysed by performing a (Fig. 4) tensile test to obtain the mechanical properties and entering the values for numerical analysis using the Abaqus program; the results are shown in Tab. 1.

Tab. 1. Mechanical properties of Epoxy specimen under tensile test

Features	Young's modulus (MPa)	ε-Fmax (%)	εBreak (%)	σBreak (MPa)	Maximal stress (MPa)
Undamaged specimen	1,793.80	4.12	5.19	34.74	39.58
Hole-notched specimen	1,423.36	2.11	2.30	19.15	21.06
Elliptical-notched specimen	547.59	3.30	3.86	10.36	11.21

Fig. 5 summarises the tensile test results for the three epoxy samples, where the results of standard deviation showed the effect of the samples on the mechanical properties. We notice that the larger the hole opening like ellipse, the lower the results of stress and Young's modulus, as well as strain. We also notice a convergence between the strain results for each of the hole-notched specimen and elliptical-notched specimens.

The three test samples for the undamaged specimen and the hole-notched specimen as well as the specimen containing the ellipse shape had the average Young's modulus for the undamaged specimen 1,635.96 ± 341.66 MPa compared with the hole-notched specimen, Young’s modulus 1,450.41 ± 162.51 MPa. While for the elliptical-notched specimen the Young’s modulus 710.06 ± 260.48 MPa as shown in Fig. 5(a), while Fig. 5(b) shows the average stresses for the three samples, which were 35.68 ± 8.60 MPa and for the undamaged specimen, and stress 21.28 ± 2.17 MPa and 12.81 ± 4.58 MPa for each of the hole-notched specimen and elliptical-notched specimens, respectively. Fig. 5(c) presents the deformation for the three samples. It showed less strain for the hole-notched specimen with an average strain of 2.16 ± 0.13%, followed by the elliptical-notched specimen with an average strain of 2.75 ± 0.65% and the largest of these was the undamaged specimen with an average strain of 3.68 ± 0.48% .The standard deviations of the three test samples for the undamaged specimen, the hole-notched specimen and the elliptical-notched specimen were in the range of 6%–37%.

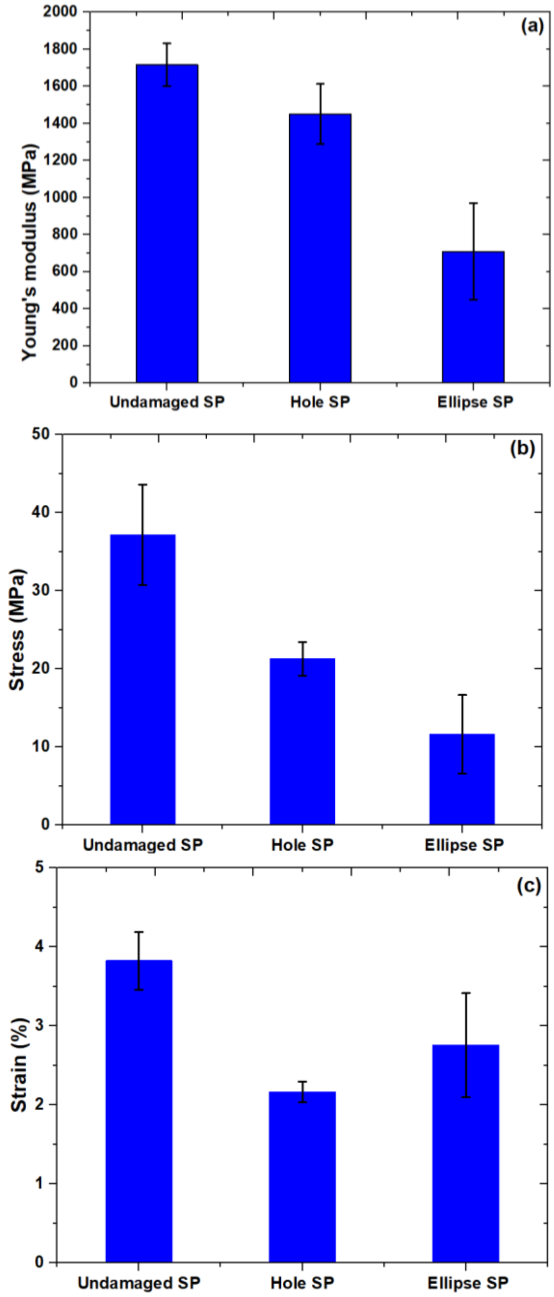


Fig. 5. Young’s modulus and stress and strain for all three samples

4. NUMERICAL ANALYSIS

4.1. Mesh view

The numerical model was realised by FE under the code Abaqus. A fine mesh has been undertaken for the different models.

There are two types of mesh: a mesh around the hole and a uniform 3 × 3 mesh as shown in Fig. 6. The mesh around the circular hole and around the elliptical hole has been refined. The mesh optimisation provides accurate values compared with the normal mesh in the FEM, and thus gives better results to ensure convergence between the experimental and numerical results. They tried to model the three specimens with the same number of mesh elements, as shown in Fig. 6.

In order to be close to the real loading conditions in the traction machine, the following boundary conditions have been assumed. One end of the specimen was embedded and the other subjected to an applied stress. Application of force perpendicular force F = 1 mm/min to the tensile samples was done, as shown in Fig. 7.

Tab. 2. Input parameters of numerical simulation

Specimen	Total number of nodes	Displace ment speed	Mesh type	Young’s modulus (MPa)
Undamaged	1,008	1	Hexagon (C3D8R)	1,793.80
Hole-notched	1,050	1	Hexagon (C3D8R)	1,423.36
Elliptical-notched	1,050	1	Hexagon (C3D8R)	547.59

The tensile test of the specimens was carried out numerically and the behaviour of the specimens was studied according to the presence of the notch. In the numerical analysis in the ABAQUS software, it showed a convergence between the experimental results and the results of the numerical analysis for the three specimens (Figs 8–10). Moreover, the maximum stress value (red colour) was 41.20 MPa and the minimum stress (blue colour) was around 0 MPa for the undamaged specimen, as shown in Fig. 8. The maximum stress value (red colour) was 20.51 MPa and the minimum stress (blue colour) was around 0 MPa for the hole-notched specimen (Fig. 9). The maximum stress value (red colour) was 11.17 MPa and the minimum stress (blue colour) was around 0 MPa for the elliptical-notched specimen (Fig. 10).

A high concentration of stress is noted in the vicinity of the notch which reduces the resistance of the specimen. However, for the specimen without notch, the stresses are distributed uniformly along its useful length. The elliptical notch represents a higher concentration of stress than that in the presence of a circular notch. The stresses in the complete sample are greater than in the other samples. A comparison of the results was made with respect to the Young’s modulus, maximum stress, and deformation between the experimental results and those obtained numerically (see Tab. 5). There is clearly a slight difference between the different properties between the simulation and experimental results.

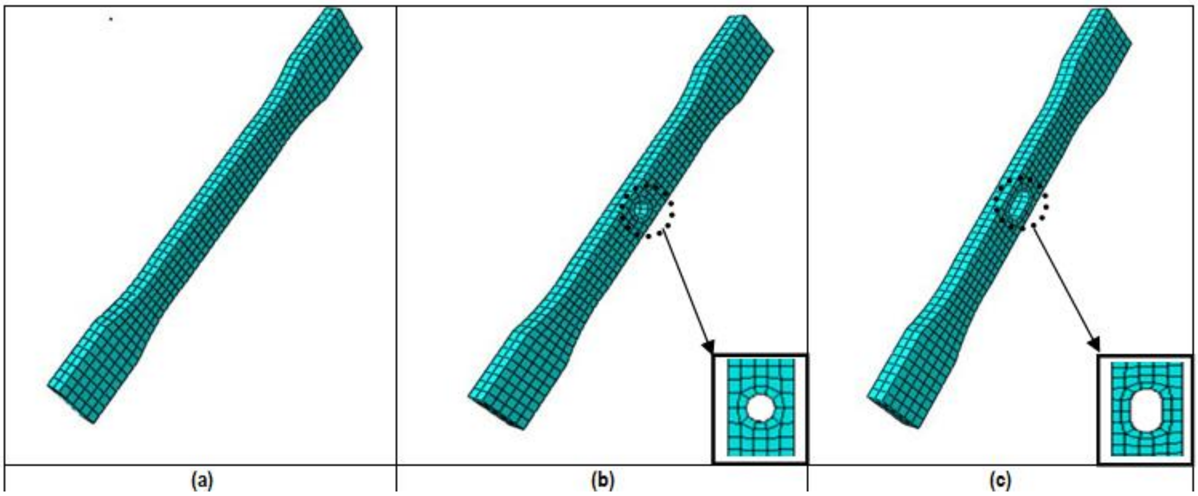


Fig. 6. Mesh view. (a) The undamaged specimen. (b) The hole-notched specimen. (c) The elliptical-notched specimen

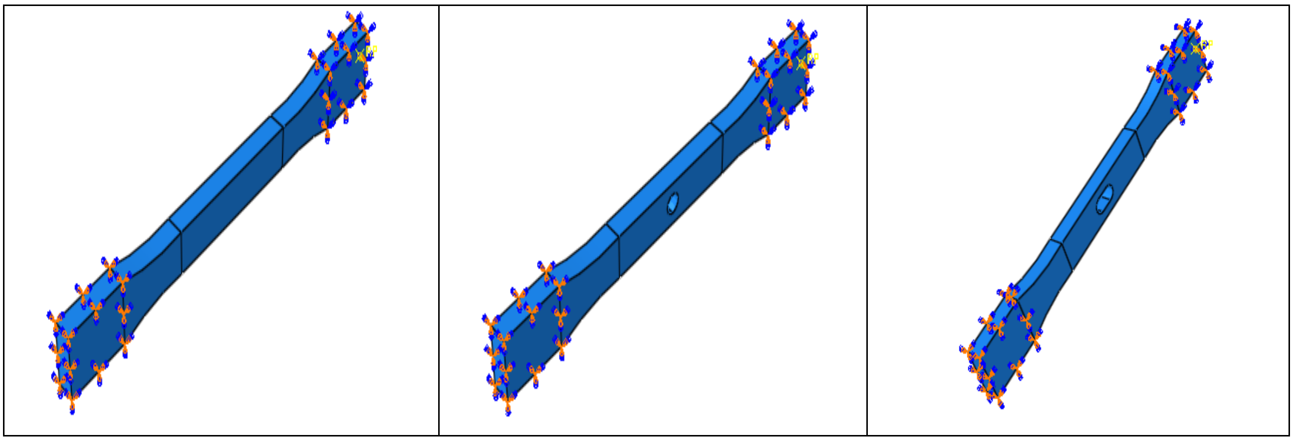


Fig. 7. Application of force to specimens in ABAQUS program

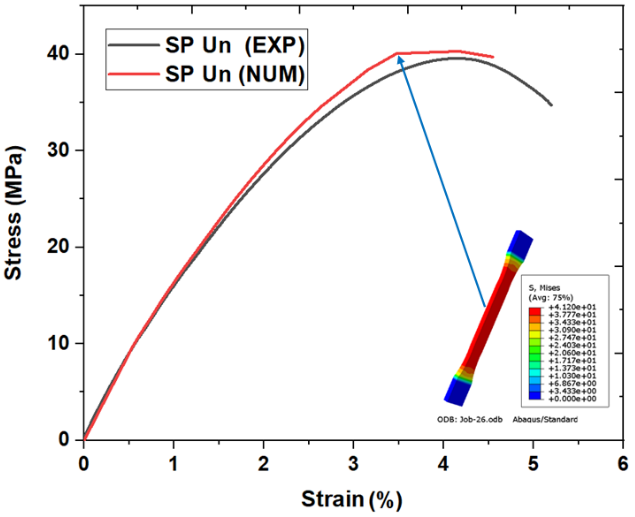


Fig. 8. Comparison between the experimental and numerical results for the undamaged specimen

It can be seen in Tab. 3 that the effect of stresses is higher in the undamaged specimen and lower in the elliptical-notched specimen. Stresses in the undamaged specimen were 41.20 MPa and stresses in the elliptical-notched specimen were 11.17 MPa. As for the results of PEMAG (the magnitude of equivalent plastic strains) and PEEQ (equivalent plastic strain) in the Abaqus software, the results were close and similar in the perforated and elliptical-notched specimen and slightly higher in the undamaged specimen.

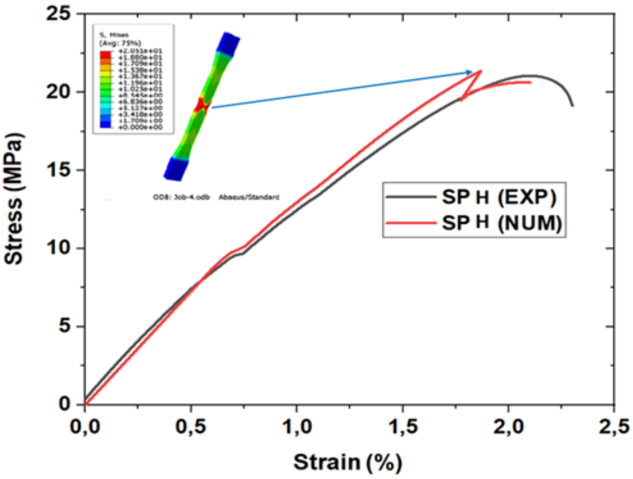


Fig. 9. Comparison between the experimental and numerical results for hole-notched specimen

Tab. 5 shows, comparing the results of the reactions generated in the three specimens along the steepest region, that the highest value of RF magnitude is 310.7 N in the undamaged specimen, 83.49 N in the hole-notched specimen and 55.26 N in the elliptical-notched specimen. The displacement is close for the hole-notched specimen and the undamaged specimen of 5.993 mm and 4.580 mm, respectively, and it is lower for the elliptical-notched specimen of 1.421 mm.

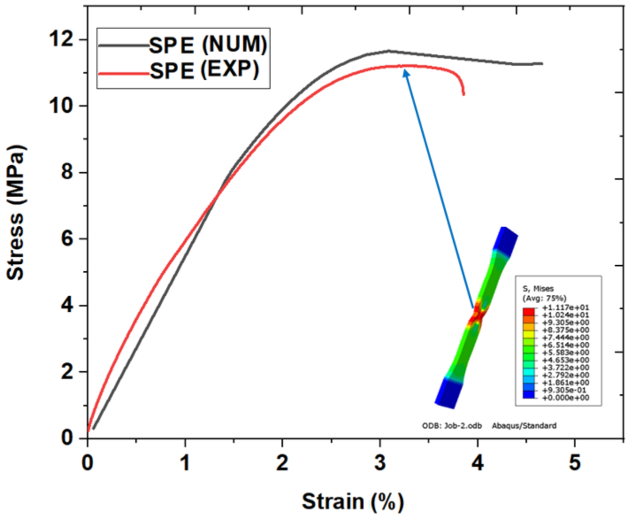


Fig. 10. Comparison between the experimental and numerical results for elliptical-notched specimen

Tab. 3. Experimental and simulation result with error for the stress

Specimen	Experimental	ABAQUS	Error (%)
Undamaged	41.20	41.22	0.04
Hole-notched	20.51	21.06	2 .61
Elliptical-notched	11.17	11.21	0.35

Tab. 4. Experimental and simulation result with error for the Young's modulus

Specimen	Experimental	ABAQUS	Error (%)
Undamaged	1,793.80	1,650.01	8.08
Hole-notched	1,423.36	1,185.92	16.89
Elliptical-notched	547.59	474.47	13.35

Fig. 11 shows that the cracks propagated directly from the edge of the hole to the nearest edge of the samples, due to the quasi-isotropic characteristic of the samples of the epoxy composites. This can be clearly seen in the optical images included in Fig. 11, the tensile specimens after fractures for all three types (undamaged, circular and elliptical). The damage occurred across the width of the specimen on either side of the hole at the notched specimen and the ellipse notched specimen, while different damage could be observed at various locations on solid specimens, the fracture of the destroyed specimen occurred along the plane perpendicular to the direction of maximum tensile stress.

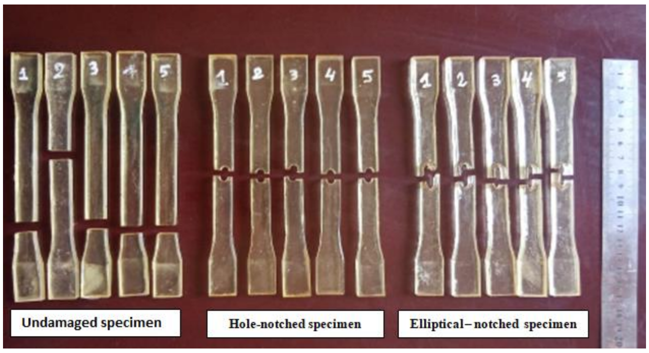


Fig. 11. Tensile specimens after fractures

Tab. 5. Comparison between numerical results for the three specimens on ABAQUS

Specimen	STRESS MAX (MPa)	E MAX(%) principal	PEMAG MAX (%)	PEEQ MAX	RF Magnitude MAX(N)	U Magnitude MAX(mm)	STATUSXFEM MAX
Undamaged	41.20	0.05828	0.03739	0.03740	310.7	4.580	1.000
Hole-notched	20.51	0.04016	0.02635	0.02640	83.49	5.993	1.000
Elliptical-notched	11.17	0.04185	0.02660	0.02660	55.26	1.421	0.4001
	STRESS min (MPa)	E MIN (%) principal	PEMAG min (%)	PEEQ min	RF Magnitude min(N)	U Magnitude min(mm)	STATUSXFEM min
Undamaged	3.333	4.86	0.003116	0.003116	25.89	0.3816	0.0800
Hole-notched	1.709	0.003347	0.002.196	0.002200	6.958	0.499	0.0833
Elliptical-notched	0.9305	0.003488	0.002261	0.0002217	4.602	0.1184	0.03334

5. CONCLUSION

The study undertaken in this work aims to use the XFEM technique for modelling the behaviour of an epoxy-type polymer to see the influence of the presence of notch on the tensile response. The obtained experimental results indicate that the undamaged specimen has ultimate tensile strength of 41.22 MPa and Young's modulus of 1,793.80 MPa, which are the strongest parts among the studied specimens. Comparison of the ultimate tensile strength indicates that an increase in hole diameter leads to a decrease in the strength of the part. The ultimate maximum stress is a decrease of 11.21 MPa and a Young's modulus of 547.59 MPa for the elliptical-notched specimen. The weakest examined is the elliptical-notched specimen. The standard deviations of the three test samples for the undamaged specimen, the hole-notched specimen and the elliptical-notched specimen are in the range 6%–37%. The comparison of the numerical results with the experimental values revealed good agreement and that the error ratios are less than 3% for the maximum stress, while the error rate is less than 17% for the Young's modulus.

FE analysis using the XFEM technique is efficient and presents good results if the mesh of the specimen is well optimised.

A high stress concentration is noted around the elliptical and circular notches. These geometric discontinuities reduce the width of the plate and provide a considerable drop in the value of the maximum tensile stress.

REFERENCES

1. Hermansson F, Janssen M, Gellerstedt F. Environmental evaluation of Durapulp Bio-composite using LCA: comparison of two applications. J For. 2016; 5: 68-76.
2. Kahl S, Peng RL, Calmunger M, Olsson B, Johansson S. In situ EBSD during tensile test of aluminum AA3003 sheet. Micron. 2014. 58: 15-24.
3. Mohanty AK, Misra M, Drzal LT. Sustainable Bio-Composites from Renewable Resources: Opportunities and Challenges in the Green Materials World. J Polym Environ. 2002; 10(1): 19-26.
4. Calì M , Pascoletti G, Gaeta M , Milazzo G, Ambu R.A New Generation of Bio-Composite Thermoplastic Filaments for a More Sustainable Design of Parts Manufactured by FDM. Appl Sci. 2020; 10(17): 5852.
5. Paiva JMd, Mayer S, Rezende MC. Comparison of tensile strength of different carbon fabric reinforced epoxy composites. Mater Res. 2006; 9(1): 83-90.
6. Goutham ERS, Vamshi Y, Namratha M, Gupta KB, Chandrasekar M, Naveen J. Influence of glass fibre hybridization on the open hole tensile properties of pineapple leaf fiber/epoxy composites. AIP Conf Proc;2022.
7. Larbi Chaht F, Mokhtari M, Benzaama H. Using a Hashin Criteria to predict the Damage of composite notched plate under traction and torsion behavior. Frat.Integrità.Strut. 2019; 13(50): 331-341.
8. SaadallahY Modeling of mechanical behavior of cork in compression. Frat.Integrità.Strut.2020; 14(53): 417-425.
9. Huang Y, Frings P, Hennes E. Mechanical properties of Zylon/epoxy composite. Composites, Part B.2002; 33(2): 109-115.
10. Duc F, Bourban PE, Månson JAE The role of twist and crimp on the vibration behaviour of flax fibre composites. Compos Sci Technol 2014; 102: 94-99.
11. uillén-Rujano R, Avilés F, Vidal-Lesso A, Hernández-Pérez A.Closed-form solution and analysis of the plate twist test in sandwich and laminated composites. Mech Mater. 2021; 155: 103753.
12. Tretyakova TV, Wildemann VE, Strungar EM. Deformation and failure of carbon fiber composite specimens with embedded defects during tension-torsion test. Frat.Integrità.Strut .2018; 12(46):295-305.
13. Liang S, Gning PB, Guillaumat L.A comparative study of fatigue behaviour of flax/epoxy and glass/epoxy composites. Compos Sci Technol. 2012; 72(5): 535-543.
14. Lu Z , Feng B, Loh C.Fatigue behaviour and mean stress effect of thermoplastic polymers and composites. Frat Integrità Strut 2018;12(46): 150-157.
15. Banaszkiewicz M, Dudda W. Applicability of notch stress-strain correction methods to low-cycle fatigue life prediction of turbine rotors subjected to thermomechanical loads. acta mech autom. 2018;12(3).
16. Panettieri E, Fanteria D , Montemurro M . Low-velocity impact tests on carbon/epoxy composite laminates: A benchmark study. Compos B Eng. 2016;107: 9-21.
17. Baykan BM, Yolum U , Özaslan E , Güler MA, Yıldırım B. Failure Prediction of Composite Open Hole Tensile Test Specimens Using Bond Based Peridynamic Theory. Procedia Struct Integ. 2020; 28: 2055-2064.
18. Hao A, Zhao H, Chen JY. Kenaf/polypropylene nonwoven composites: The influence of manufacturing conditions on mechanical, thermal, and acoustical performance. Compos B Eng. 2013; 54: 44-51.
19. Galeta T, Raos P , Stojšić J , Pakši I. Influence of Structure on Mechanical Properties of 3D Printed Objects. Procedia Eng. 2016; 149: 100-104.
20. hosravani MR, Rezaei S , Faroughi S , Reinicke T. Experimental and numerical investigations of the fracture in 3D-printed open-hole plates. Theor Appl Fract Mech.2022; 121:103543.
21. Zako M, Uetsuji Y, Kurashiki T. Finite element analysis of damaged woven fabric composite materials. Compos Sci Technol. 2003; 63(3):507-516.
22. Dixit A, Mali HS. Modeling techniques for predicting the mechanical properties of woven-fabric textile composites: a review. Mech compos Mater. 2013; 49(1): 1-20.
23. Eshraghi S, Das S. Micromechanical finite-element modeling and experimental characterization of the compressive mechanical properties of polycaprolactone–hydroxyapatite composite scaffolds prepared by selective laser sintering for bone tissue engineering. Acta biomater. 2012; 8(8): 3138-3143.
24. Mohammadi R , Najafabadi MA, Saeedifar M . Correlation of acoustic emission with finite element predicted damages in open-hole tensile laminated composites. Compos B Eng. 2017; 108: 427-435.
25. Ghezzo, F, Giannini G, Cesari F, Caligiana G, Numerical and experimental analysis of the interaction between two notches in carbon fibre laminates. Compos Sci Technol. 2008; 68(3): 1057-1072.
26. Liao T, Adanur S. A Novel Approach to Three-Dimensional Modeling of Interlaced Fabric Structures. Text Res J. 1998; 68(11): 841-847.
27. Bogrekci I, Demircioglu P, Sucuoglu HS,Altun E, Sakar B, Durakbasa MN, Topology Optimization of a Tensiletest Specimen. Int Sci Bk; 2020.
28. Khosravani MR. Influences of defects on the performance of adhesively bonded sandwich joints. Key eng mater; 2018.
29. Kojnoková T, Nový F, Markovičová L. Evaluation of tensile properties of carbon fiber reinforced polymers produced from commercial prepregs. Mater Today Proc. 2022; 62: 2663-2668.
30. Xu P, Yang C, Peng Y , Yao S, Zhang D, Li B .Crash performance and multi-objective optimization of a gradual energy-absorbing structure for subway vehicles. int j mech sci. 2016; 107:1-12.
31. Feulvarch E, Lacroix R, Deschanels H, A 3D locking-free XFEM formulation for the von Mises elasto-plastic analysis of cracks. Comput Methods Appl Mech Eng. 2020; 361: 112805.
32. Frolov AS, Fedotov IV, Gurovich BA. Evaluation of the true-strength characteristics for isotropic materials using ring tensile test. nucl eng technol. 2021; 53(7): 2323-2333.
33. Pilkey WD, Pilkey DF, Bi Z. Peterson's stress concentration factors;2020.

Khalissa Saada: https://orcid.org/0000-0002-3025-1287

Salah Amroune: https://orcid.org/0000-0002-9565-1935

Moussa Zaoui: https://orcid.org/0009-0005-7178-2542

Amin Houari: https://orcid.org/0009-0004-2617-2182

Kouider Madani: https://orcid.org/0000-0003-3277-1187

Amina Hachaichi: https://orcid.org/0000-0002-4905-7599

DOI 10.2478/ama-2023-0023

EFFECTIVE SHAPING OF A STEPPED SANDWICH BEAM WITH CLAMPED ENDS

Krzysztof MAGNUCKI*, Joanna KUSTOSZ*, Damian GOLIWĄS*

*Łukasiewicz Research Network – Poznan Institute of Technology, Rail Vehicles Center,
ul. Warszawska 181, 61-055 Poznań, Poland

krzysztof.magnucki@pit.lukasiewicz.gov.pl, joanna.kustosz@pit.lukasiewicz.gov.pl, damian.goliwas@pit.lukasiewicz.gov.pl

received 25 November 2022, revised 10 January 2023, accepted 10 January 2023

Abstract: The aim of this work is to propose a sandwich beam with stepped layer thickness in three parts along its length. The total depth, width of the cross-section and its mass are constant. The beam is under a uniformly distributed load. The system of two equilibrium equations was formulated for each part based on the literature. This system was analytically solved for the successive parts of the beam and the functions of the shear effect and deflection were determined in them. The effective stepped layer thicknesses was determined on the basis of the adopted criterion for minimizing the maximum deflection of the beam. The example calculations were made for two elected beams. The effective shapes of these beams are shown in the figures. Moreover, FEM numerical calculations of the deflections of these beams are performed.

Key words: analytical modelling, bending, shaping criterion, FEM calculation

1. INTRODUCTION

Sandwich constructions initiated in the 20th century are intensively developed in the 21st century. Vinson [1] presented a general introduction to the mechanics of sandwich structures with reference to the 174 articles. Icardi [2] developed a sublaminate model taking into account the zig-zag theory for the analysis of laminated and sandwich beams. The aim is to show the advantages of using higher-order approximations of displacements in sublaminates. Yang and Qiao [3] developed an analytical high-order impact model of a soft-core sandwich beam to analyse its response to foreign body impact. The results of analytical tests were compared with the results of FEM (Finite Element Method) numerical tests. Magnucka-Blandzi and Magnucki [4] presented the problem of effective shaping of a sandwich beam with a metal foam core with variable properties along its thickness. The optimal dimensionless parameters of the beam were determined on the basis of the adopted criterion. Kreja [5] described, based on a review of 246 articles, the state of the art in the field of analytical and numerical FEM methods used in the calculations of laminated composite and sandwich panels. Wang and Li [6] presented a theoretical analysis of bending of two types of sandwich beams with aluminium or steel facings and cores made of shape memory polymers. Nguyen et al. [7] studied sandwich panels with stepped facings and honeycomb cores. They have demonstrated in numerous examples that stepped linings can increase the strength and rigidity of sandwich structures. Phan et al. [8], taking into account the high-order theory of sandwich panels (HSAPT), developed a one-dimensional high-order theory for elastic orthotropic sandwich beams, taking into account the transverse compressibility of the core. Magnucki et al. [9] developed analytical and numerical FEM models of a five-layer sandwich beam and analysed its strength and stability. Sayyad and Ghugal [10] made a critical review of analytical and numerical studies described in selected 515 papers on bending, buckling and free vibration of homogeneous, laminated composite and sandwich beams, taking into account the applied theories. Birman and Kardomateas [11] presented, based on a review of 363 articles, contemporary trends in the development of sandwich structures in terms of theory and their practical application, with an emphasis on aviation, civil and marine engineering, electronics and biomedicine. In addition, sandwich structures are used in ships, which was described in detail by Kozak in 2018 [12]. Magnucki [13] presented analytical studies of bending of sandwich and I-beams with a symmetrical structure using two deformation models of flat cross-sections. Magnucki et al. [14] studied analytical and numerical FEM bending, buckling and free vibration of a sandwich beam with an asymmetric structure. The analytical model was developed taking into account the classic broken-line theory. Sayyad and Ghugal [15] presented a review of research, including 250 articles, on the modelling and analysis of functionally stepped sandwich beams and indicated directions for further research. Chinh et al. [16] analysed the bending of sandwich beams with a symmetrical structure with functionally stepped facings and a porous core subjected to a uniformly distributed load. Four types of supports for these beams, simply supported, clamped-clamped, clamped-hinged and clamped-free, were included. Magnucki et al. [17] developed three models of a simply supported sandwich beam and studied analytical and numerical FEM bending, buckling and free vibration. Kustosz et al. [18] analysed analytical and numerical FEM bending of a stepped sandwich beam with fixed ends under the action of a uniformly distributed load along its length. This work is a theoretical study presenting a generalized model of a sandwich beam, thanks to which it is possible to test the bending strength of beams with stepped structures.

The aim of this work is to propose the effective shaping of a symmetrically stepped sandwich beam along its length. This work is a continuation of the studies presented in the paper [18].

2. ANALYTICAL STUDIES

2.1. Analytical model of the stepped sandwich beam

The subject of the studies is the sandwich beam stepped in three parts arranged symmetrically along its length. This beam with clamped ends of the length *L*, the total depth *h* and width *b* is subjected to the uniform load of intensity *q* (Fig. 1).

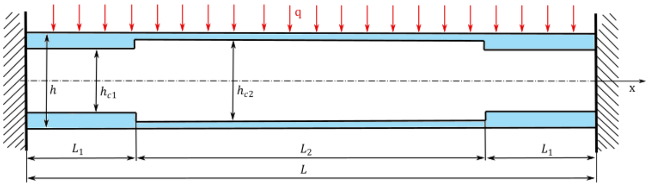


Fig. 1. Scheme of the stepped sandwich beam and the load

The volume of the core of the classical sandwich beam with constant layer thicknesses is given as follows

$$V_c^{(cl)} = bh_cL, \tag{1}$$

where h_c is the thickness of the core.

However, the volume of the core of the stepped sandwich beam (Fig. 1) is given in the following form:

$$V_c^{(st)} = b(2h_{c1}L_1 + h_{c2}L_2), \tag{2}$$

where h_{c1} is the thickness, L_1 is the length of the first part of the beam, h_{c2} is the thickness and L_2 is the length of the second-middle part of the beam.

Equating the volumes of the core of the classical beam (1) and the stepped beam (2), $V_c^{(cl)} = V_c^{(st)}$, after simple transformation, the following is obtained:

$$2\chi_{c1}\lambda_1 + \chi_{c2}\lambda_2 = \chi_c\lambda = \text{const}, \tag{3}$$

from which

$$\chi_{c2} = \frac{\chi_c\lambda - 2\chi_{c1}\lambda_1}{\lambda - 2\lambda_1}, \tag{4}$$

where $\chi_c = h_c/h$, $\chi_{c1} = h_{c1}/h$, $\chi_{c2} = h_{c2}/h$ is the relative thicknesses of the cores and $\lambda = L/h$, $\lambda_1 = L_1/h$, $\lambda_2 = L_2/h$ is the relative lengths of the beam and its parts, also $2\lambda_1 + \lambda_2 = \lambda$.

It is noted that the expression (3) also provides a constant volume of the faces, so it is a condition for constant mass of the stepped sandwich beam.

The system of two differential equations of equilibrium for individual parts of the stepped sandwich beam, based on the paper [18], is formulated in the following form:

$$C_{vvi}\frac{d^2\bar{v}^{(i)}}{d\xi^2} - C_{v\psi i}\frac{d\psi_f^{(i)}}{d\xi} = -6[(\xi - \xi^2)qL^2 - 2M_0]\frac{\lambda}{E_fbh^2}, \tag{5}$$

$$C_{v\psi i}\frac{d^3\bar{v}^{(i)}}{d\xi^3} - C_{\psi\psi i}\frac{d^2\psi_f^{(i)}}{d\xi^2} + C_{\psi i}\lambda^2\psi_f^{(i)}(\xi) = 0, \tag{6}$$

where $\xi = x/L$ is the dimensionless coordinate, $i = 1, 2$ is the number of the beam part, $\bar{v}^{(i)}(\xi) = v^{(i)}(\xi)/L$ is the relative deflection, $\psi_f^{(i)}(\xi)$ is the dimensionless longitudinal displacements in faces, $C_{vvi} = 1 - (1 - e_c)\chi_{ci}^3$, $C_{v\psi i} = 3 - (3 - 2e_c)\chi_{ci}^2$, $C_{\psi\psi i} = 4[3 - (3 - e_c)\chi_{ci}]$, $C_{\psi i} = \frac{24}{1+\nu_c}\frac{e_c}{\chi_{ci}}$, $e_c = E_c/E_f$ is the dimensionless coefficient, E_f, E_c is the Young modulus of faces and the core, ν_c is the Poisson ratio of the core and M_0 is the clamped-ends moment.

This system of two differential equations, Eqs (5) and (6), after simple transformation, is reduced to one differential equation in the following form:

$$\frac{d^2\psi_f^{(i)}}{d\xi^2} - (\alpha_i\lambda)^2\psi_f^{(i)}(\xi) = -6(1 - 2\xi)\frac{C_{v\psi i}}{C_{vvi}C_{\psi\psi i} - C_{v\psi i}^2}\lambda^3\frac{q}{E_fb}, \tag{7}$$

where $\alpha_i = \sqrt{\frac{C_{v\psi i}C_{\psi i}}{C_{vvi}C_{\psi\psi i} - C_{v\psi i}^2}}$ is the dimensionless coefficient.

The solution of this equation is given as the following function:

$$\psi_f^{(i)}(\xi) = k_{ci}[C_{1i}\sinh(\alpha_i\lambda\xi) + C_{2i}\cosh(\alpha_i\lambda\xi) + \\ +6(1 - 2\xi)]\lambda\frac{q}{E_fb}, \tag{8}$$

where $k_{ci} = \frac{C_{v\psi i}}{C_{vvi}C_{\psi i}}$ is the coefficient and C_{1i}, C_{2i} are integration constants.

Eq. (5), after the first integration, is given as follows:

$$C_{vvi}\frac{d\bar{v}^{(i)}}{d\xi} = C_{3i} + C_{v\psi i}\psi_f^{(i)}(\xi) - 6\left(\frac{1}{2}\xi^2 - \frac{1}{3}\xi^3 - 2\xi\bar{M}_0\right)\frac{q\lambda^3}{E_fb}, \tag{9}$$

where C_{3i} is the integration constant and $\bar{M}_0 = M_0/qL^2$ is the dimensionless clamped-ends moment.

2.2. Analytical solution

The analytical solution is realized in the individual parts of the stepped beam.

The dimensionless longitudinal displacements in faces:

– the first part $(i = 1)$, $0 \le \xi \le \lambda_1/\lambda$

The function (8), with consideration of the boundary condition $\psi_f^{(1)}(0) = 0$, from which $C_{21} = -6$, is provided in the following form:

$$\psi_f^{(1)}(\xi) = k_{c1}[C_{11}\sinh(\alpha_1\lambda\xi) - 6\cosh(\alpha_1\lambda\xi) + \\ +6(1 - 2\xi)]\frac{q\lambda}{E_fb}. \tag{10}$$

This function for $\xi = \lambda_1/\lambda$ is as follows

$$\psi_f^{(1)}\left(\frac{\lambda_1}{\lambda}\right) = k_{c1}\left[C_{11}\sinh(\alpha_1\lambda_1) - 6\cosh(\alpha_1\lambda_1) + 6\frac{\lambda_2}{\lambda}\right]\frac{q\lambda}{E_fb}, \tag{11}$$

– the second-middle part $(i = 2)$, $\lambda_1/\lambda \le \xi \le 1/2$

The function (8), with consideration of the condition $\psi_f^{(2)}(1/2) = 0$ and simplification, is provided in the following form:

$$\psi_f^{(2)}(\xi) = 6k_{c2}(1 - 2\xi)\frac{q\lambda}{E_fb}. \tag{12}$$

Taking into account the continuity condition for the longitudinal displacements in faces $\psi_f^{(1)}(\lambda_1/\lambda) = \psi_f^{(2)}(\lambda_1/\lambda)$, the integration constant is obtained as follows:

$$C_{11} = \frac{6}{\sinh(\alpha_1\lambda_1)}\left[\cosh(\alpha_1\lambda_1) + \frac{k_{c2}-k_{c1}}{k_{c1}}\frac{\lambda_2}{\lambda}\right]. \tag{13}$$

Consequently, the function (10) is provided in the following form:

$$\psi_f^{(1)}(\xi) = 6k_{c1}\left\{-\frac{\sinh[\alpha_1(\lambda_1-\lambda\xi)]}{\sinh(\alpha_1\lambda_1)} + \frac{\sinh(\alpha_1\lambda\xi)}{\sinh(\alpha_1\lambda_1)}\frac{k_{c2}-k_{c1}}{k_{c1}}\frac{\lambda_2}{\lambda} + \right.$$
$$\left. +1 - 2\xi\right\}\frac{q\lambda}{E_f b}. \tag{14}$$

Therefore, the function for $\xi = \lambda_1/\lambda$ is given as follows:

$$\psi_f^{(1)}\left(\frac{\lambda_1}{\lambda}\right) = 6k_{c2}\lambda_2\frac{q}{E_f b} \tag{15}$$

The relative deflection:

- the first part $(i = 1)$, $0 \le \xi \le \lambda_1/\lambda$

Eq. (9), with consideration of the boundary condition $d\bar{v}^{(1)}/d\xi\,|_0 = 0$, from which $C_{31} = 0$, is provided in the following form:

$$C_{vv1}\frac{d\bar{v}^{(1)}}{d\xi} = C_{v\psi1}\psi_f^{(1)}(\xi) - 6\left(\frac{1}{2}\xi^2 - \frac{1}{3}\xi^3 - 2\xi\bar{M}_0\right)\frac{q\lambda^3}{E_f b}. \tag{16}$$

Therefore, the derivative of the relative deflection curve for $\xi = \lambda_1/\lambda$ is given as follows

$$\frac{d\bar{v}^{(1)}}{d\xi}\Big|_{\frac{\lambda_1}{\lambda}} =$$
$$\left\{6C_{v\psi1}k_{c2}\frac{\lambda_2}{\lambda} - \left[\left(2 + \frac{\lambda_2}{\lambda}\right)\frac{\lambda_1}{\lambda} - 12\bar{M}_0\right]\lambda_1\lambda\right\}\frac{q\lambda}{C_{vv1}E_f b}. \tag{17}$$

Eq. (16) after integration is given as follows:

$$C_{vv1}\bar{v}^{(1)}(\xi) = C_{41} + 6k_{c1}C_{v\psi1}\Phi_\psi^{(1)}(\xi)\frac{q\lambda}{E_f b} -$$
$$-6\left(\frac{\xi^3}{6} - \frac{\xi^4}{12} - \xi^2\bar{M}_0\right)\frac{q\lambda^3}{E_f b}, \tag{18}$$

where

$$\Phi_\psi^{(1)}(\xi) = \frac{\cosh[\alpha_1(\lambda_1-\lambda\xi)]}{\alpha_1\lambda_1\sinh(\alpha_1\lambda_1)} + \frac{\cosh(\alpha_1\lambda\xi)}{\alpha_1\lambda_1\sinh(\alpha_1\lambda_1)}\frac{k_{c2}-k_{c1}}{k_{c1}}\frac{\lambda_2}{\lambda} + \xi - \xi^2. \tag{19}$$

Based on the boundary condition $\bar{v}^{(1)}(0) = 0$, the integration constant is given as follows:

$$C_{41} = -\frac{6k_{c1}}{\sinh(\alpha_1\lambda_1)}\frac{C_{v\psi1}}{\alpha_1\lambda}\left[\cosh(\alpha_1\lambda_1) + \frac{k_{c2}-k_{c1}}{k_{c1}}\frac{\lambda_2}{\lambda}\right]\frac{q\lambda}{E_f b}. \tag{20}$$

Therefore, the relative deflection of this part for $\xi = \lambda_1/\lambda$ is given as follows:

$$\bar{v}^{(1)}\left(\frac{\lambda_1}{\lambda}\right) = \left\{6k_{c1}C_{v\psi1}\left[\frac{k_0}{\alpha_1\lambda}\left(\frac{k_{c2}-k_{c1}}{k_{c1}}\frac{\lambda_2}{\lambda} - 1\right) + \left(1 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda}\right] - \right.$$
$$\left. - \left[\frac{1}{2}\left(2 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda} - 6\bar{M}_0\right]\lambda_1^2\right\}\frac{q\lambda}{C_{vv1}E_f b}, \tag{21}$$

where $k_0 = \frac{\cosh(\alpha_1\lambda_1)-1}{\sinh(\alpha_1\lambda_1)}$ is the dimensionless coefficient.

- the second-middle part $(i = 2)$, $\lambda_1/\lambda \le \xi \le 1/2$

Taking into account Eq. (9) with consideration of the function (12), based on the condition $d\bar{v}^{(2)}/d\xi\,|_{1/2} = 0$, the integration constant is obtained as follows:

$$C_{32} = \frac{1}{2}(1 - 12\bar{M}_0)\frac{q\lambda^3}{E_f b}. \tag{22}$$

Therefore, Eq. (9) is provided in the following form:

$$C_{vv2}\frac{d\bar{v}^{(2)}}{d\xi} = \left\{6C_{v\psi2}k_{c2}(1 - 2\xi) + \left[\frac{1}{2} - 3\xi^2 + 2\xi^3 - \right.\right.$$
$$\left.\left. -6(1 - 2\xi)\bar{M}_0\right]\lambda^2\right\}\frac{q\lambda}{E_f b}. \tag{23}$$

Thus, the derivative of the relative deflection curve for $\xi = \lambda_1/\lambda$ is given as follows:

$$\frac{d\bar{v}^{(2)}}{d\xi}\Big|_{\frac{\lambda_1}{\lambda}} = \left\{6C_{v\psi2}k_{c2}\frac{\lambda_2}{\lambda} + \left[\frac{1}{2} - \left(3 - 2\frac{\lambda_1}{\lambda}\right)\left(\frac{\lambda_1}{\lambda}\right)^2 - \right.\right.$$
$$\left.\left. -6\frac{\lambda_2}{\lambda}\bar{M}_0\right]\lambda^2\right\}\frac{q\lambda}{C_{vv2}E_f b}. \tag{24}$$

Based on the continuity condition for the derivative of the relative deflection curve $d\bar{v}^{(1)}/d\xi\,|_{\lambda_1/\lambda} = d\bar{v}^{(2)}/d\xi\,|_{\lambda_1/\lambda}$, the dimensionless clamped-ends moment is obtained as follows:

$$\bar{M}_0 = \frac{\mathrm{N}_{M0}}{12[2(C_{vv1}-C_{vv2})\lambda_1 - C_{vv1}\lambda]\lambda}, \tag{25}$$

where the numerator of this expression is given as follows:

$$\mathrm{N}_{M0} = 12\left(C_{vv2}C_{v\psi1} - C_{vv1}C_{v\psi2}\right)\frac{\lambda_2}{\lambda}k_{c2} - C_{vv1}\lambda^2 +$$
$$+2(C_{vv1} - C_{vv2})\left(2 + \frac{\lambda_2}{\lambda}\right)\lambda_1^2. \tag{26}$$

Eq. (23) after integration is given as follows:

$$C_{vv2}\bar{v}^{(2)}(\xi) = C_{42} + \left\{6C_{v\psi2}k_{c2}(\xi - \xi^2) + \left[\frac{1}{2}\xi - \xi^3 + \right.\right.$$
$$\left.\left. +\frac{1}{2}\xi^4 - 6(\xi - \xi^2)\bar{M}_0\right]\lambda^2\right\}\frac{q\lambda}{E_f b}. \tag{27}$$

Based on the continuity condition for the relative deflection curve $\bar{v}^{(1)}(\lambda_1/\lambda) = \bar{v}^{(2)}(\lambda_1/\lambda)$, the integration constant is obtained as follows:

$$C_{42} = [6(\bar{C}_{421} - \bar{C}_{422}) - (\bar{C}_{423} + \bar{C}_{424})\lambda^2]\frac{q\lambda}{E_f b}, \tag{28}$$

where

$$\bar{C}_{421} = C_{v\psi1}\frac{C_{vv2}}{C_{vv1}}k_{c1}\left\{\frac{k_0}{\alpha_1\lambda}\left(\frac{k_{c2}-k_{c1}}{k_{c1}}\frac{\lambda_2}{\lambda} - 1\right) + \left(1 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda}\right\},$$

$$\bar{C}_{422} = C_{v\psi2}k_{c2}\left(1 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda},$$

$$\bar{C}_{423} = \frac{C_{vv2}}{C_{vv1}}\left[\frac{1}{2}\left(2 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda} - 6\bar{M}_0\right]\left(\frac{\lambda_1}{\lambda}\right)^2,$$

$$\bar{C}_{424} = \left\{\frac{1}{2}\left[1 + \left(1 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda}\right] - 6\bar{M}_0\right\}\left(1 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda}.$$

Taking into account Eq. (27), the maximum relative deflection of the stepped sandwich beam is given as follows:

$$\bar{v}_{\max} = \bar{v}^{(2)}\left(\frac{1}{2}\right) = \tilde{v}_{\max}\frac{q}{E_f b}, \tag{29}$$

where

$$\tilde{v}_{\max} = \left(\tilde{v}_\psi + \tilde{v}_v\lambda^2\right)\lambda, \tag{30}$$

and

$$\tilde{v}_\psi = 6\left\{\frac{C_{v\psi1}}{C_{vv1}}\frac{k_0}{\alpha_1\lambda}\left[(k_{c2} - k_{c1})\frac{\lambda_2}{\lambda} - k_{c1}\right] + \frac{1}{4}\frac{C_{v\psi2}}{C_{vv2}}k_{c2} + \right.$$
$$\left. + \left(\frac{C_{v\psi1}}{C_{vv1}}k_{c1} - \frac{C_{v\psi2}}{C_{vv2}}k_{c2}\right)\left(1 - \frac{\lambda_1}{\lambda}\right)\frac{\lambda_1}{\lambda}\right\}, \tag{31}$$

$$\tilde{v}_v = \left(\frac{5}{32} - \frac{3}{2}\bar{M}_0 - \bar{C}_{423} - \bar{C}_{424}\right)\frac{1}{C_{vv2}}. \qquad (32)$$

Thus, the criterion of effective shaping of the stepped sandwich beam was assumed as the minimization of the maximum dimensionless deflection of this beam, with considering two expressions (4) and $\lambda_2 = \lambda - 2\lambda_1$, and is obtained in the following form:

$$\min_{\lambda_1, \chi_{c1}}[\tilde{v}_{\max}(\lambda_1, \chi_{c1})]. \qquad (33)$$

The detailed calculations are carried out for the exemplary stepped sandwich beams.

3. DETAILED CALCULATIONS

3.1. Beam B-1

The data of the classical sandwich beam B-1 are specified in Tab. 1. However, the results of the calculations, the effective dimensionless sizes and maximal deflection, are specified in Tab. 2.

Tab. 1. The classical sandwich beam – B-1

λ	e_c	ν_c	χ_c	$\tilde{v}_{\max}$
20	1/20	0.3	17/20	728.66

Tab. 2. The effective dimensionless sizes and maximal deflection – B-1

$\lambda_{1,ef}$	$\lambda_{2,ef}$	$\chi_{c1,ef}$	$\chi_{c2,ef}$	$\tilde{v}_{\max}$
2.4	15.2	14.66/20	17.74/20	676.02

Moreover, the graph of the dimensionless longitudinal displacements – shear effect functions (12) and (14) – is shown in Fig. 2, and the scheme of the effective shape of the stepped sandwich beam is shown in Fig. 3.

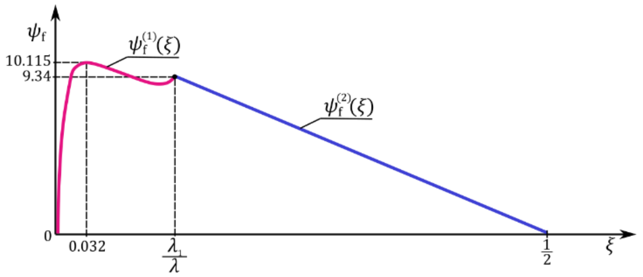


Fig. 2. The graph of the dimensionless longitudinal displacements – shear effect functions

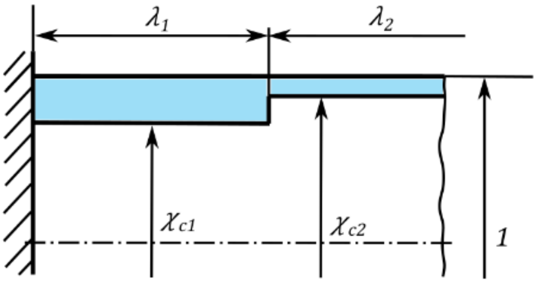


Fig. 3. The scheme of the effective shape of the stepped sandwich beam

3.2. Beam B-2

The data of the classical sandwich beam B-2 are specified in Tab. 3. Moreover, the results of the calculations, the effective dimensionless sizes and maximal deflection, are specified in Tab. 4, and the graph of the dimensionless longitudinal displacements – shear effect functions (12) and (14) – is shown in Fig. 4. The scheme of the effective shape of this stepped sandwich beam is similar to Fig. 3.

Tab. 3. The classical sandwich beam – B-2

λ	e_c	ν_c	χ_c	$\tilde{v}_{\max}$
20	1/20	0.3	16/20	613.14

Tab. 4. The effective dimensionless sizes and maximal deflection – B-2

$\lambda_{1,ef}$	$\lambda_{2,ef}$	$\chi_{c1,ef}$	$\chi_{c2,ef}$	$\tilde{v}_{\max}$
2.5	15.0	13.22/20	16.93/20	576.16

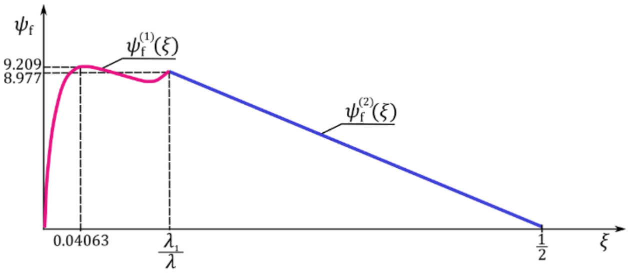


Fig. 4. The graph of the dimensionless longitudinal displacements – shear effect functions

4. NUMERICAL FEM STUDIES

4.1. Beam B-1

The numerical model of the example effective stepped sandwich beam B-1 was developed in the ABAQUS 6.12 system using 84,000 hexahedral linear finite elements (C3D8R type). The model of the beam is solid and represents only half of the beam due to the symmetry. The longitudinal x-axis is collinear with the beam neutral axis, the y-axis is directed and the z-axis is parallel to the cross-section of the beam. The beam is under a continuous load and its ends are clamped (Fig. 5).

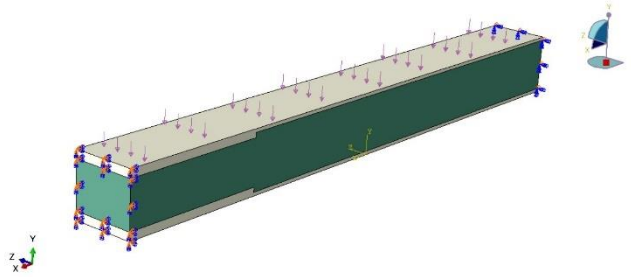


Fig. 5. The scheme of the numerical FEM model of the effective stepped sandwich beam B-1

The value of the maximum deflection determined numerically is as follows: $\tilde{v}_{\max} = 678.96$. The difference between analytical (An) and numerical (FEM) results is 0.43% in these exemplary stepped sandwich beam.

4.2. Beam B-2

The numerical model of the B-2 beam is analogous to the model of the B-1 beam. The value of the maximum deflection determined numerically is given as follows: $\tilde{v}_{\max} = 581.58$. The difference between analytical (An) and numerical (FEM) results is 0.94% in these exemplary stepped sandwich beams.

5. CONCLUSIONS

The detailed calculations for the exemplary stepped sandwich beams provide the following conclusions:

- The stiffness of the sandwich structures can be increased by introducing stepped facings, which is expressed in smaller maximum deflections compared to the maximum deflection of the classical sandwich beam, and so for the beam B-1 by 7.2% and for the beam B-2 by 6.0%.
- The developed analytical and numerical FEM models of this beam are equivalent, the differences between the values of the maximum deflections of the exemplary beams calculated based on these two models are less than 1%, and so for the beam B-1 is 0.43% and for the beam B-2 is 0.94%.
- The facings thicknesses of the effective sandwich beams in the first part, at clamped beam ends, are greater than the facings thicknesses in the middle part of these beams (Figs. 3 and 5).
- In future works related to this paper, the problem of effective shaping of sandwich beams with a stepped structure, taking into account the local buckling (face wrinkling), could be considered.

REFERENCES

1. Vinson JR. Sandwich structures. Applied Mechanics Reviews. 2001;54(3):201–214.
2. Icardi U. Applications of Zig-Zag theories to sandwich beams. Mechanics of Advanced Materials and Structures. 2003;10(1):77–97.
3. Yang M, Qiao P. Higher-order impact modeling of sandwich structures with flexible core. International Journal of Solids and Structures. 2005;42(20):5460–90.
4. Magnucka-Blandzi E, Magnucki K. Effective design of a sandwich beam with a metal foam core. Thin-Walled Structures. 2007;45(4):432-8.
5. Kreja I. A literature review on computational models for laminated composite and sandwich panels. Central European Journal of Engineering. 2011;1(1):59-80.
6. Wang ZD, Li ZF. Theoretical analysis of the deformation of SMP sandwich beam in flexure. Archive of Applied Mechanics. 2011;81(11):1667–78.
7. Nguyen CH, Chandrashekhara K, Birman V. Enhanced static response of sandwich panel with honeycomb cores through the use of stepped facings. Journal of Sandwich Structures and Materials. 2011;13(2):237-60.
8. Phan CN, Frostig Y, Kardomateas GA. Analysis of sandwich beams with a compliant core and with in-plane rigidity–extended high-order sandwich panel theory versus elasticity. ASME: Journal of Applied Mechanics. 2012;79:041001–1-11.
9. Magnucki K, Jasion P, Szyc W, Smyczynski M. Strength and buckling of a sandwich beam with thin binding layers between faces and a metal foam core. Steel and Composite Structures. 2014;16(3):325-37.
10. Sayyad AS, Ghugal YM. Bending, buckling and free vibration of laminated composite and sandwich beams: a critical review of literature. Composite Structures. 2017;171:486–504.
11. Birman V, Kardomateas GA. Review of current trends in research and applications of sandwich structures. Composites Part B. 2018;142:221-40.
12. Kozak J. Steel sandwich panels in ship structures. Gdańsk Tech Publishing House, 2018, Gdańsk. ISBN 978-83-7348-742-0 (in polish).
13. Magnucki K. Bending of symmetrically sandwich beams and I-beams – Analytical study. International Journal of Mechanical Sciences. 2019;150:411-9.
14. Magnucki K, Magnucka-Blandzi E, Lewiński J, Milecki S. Analytical and numerical studies of an unsymmetrical sandwich beam - bending, buckling and free vibration. Engineering Transactions. 2019;67(4):491-512.
15. Sayyad AS, Ghugal YM. Modeling and analysis of functionally graded sandwich beams: A review. Mechanics of Advanced Materials and Structures. 2019;26(21):1776-95.
16. Chinh TH, Tu TM, Duc DM, Hung TQ. Static flexural analysis of sandwich beam with functionally graded face sheets and porous core via point interpolation meshfree method based on polynomial basic function. Archive of Applied Mechanics. 2021;91(3):933–47.
17. Magnucki K, Magnucka-Blandzi E, Wittenbeck L. Three models of a sandwich beam: Bending, buckling and free vibration. Engineering Transactions. 2022;70(2):97-122.
18. Kustosz J, Magnucki K, Goliwąs D. Bending of a stepped sandwich beam: The shear effect. Engineering Transactions. 2022;70(4):373-390.

Krzysztof Magnucki: https://orcid.org/0000-0003-2251-4697

Joanna Kustosz: https://orcid.org/0000-0002-9408-2099

Damian Goliwąs: https://orcid.org/0000-0003-3280-0400

sciendo

DOI 10.2478/ama-2023-0024

acta mechanica et automatica, vol.17 no.2 (2023)
Special Issue "Machine Modeling and Simulations 2022"

THE INFLUENCE OF FRICTION FORCE AND HYSTERESIS ON THE DYNAMIC RESPONSES OF PASSIVE QUARTER-CAR SUSPENSION WITH LINEAR AND NON-LINEAR DAMPER STATIC CHARACTERISTICS

Zbyszko KLOCKIEWICZ*, Grzegorz ŚLASKI*

*Faculty of Mechanical Engineering, Institute of Machine Design, Poznan University of Technology,
ul. Piotrowo 3, 61-138, Poznan, Poland

zbyszko.klockiewicz@put.poznan.pl, grzegorz.slaski@put.poznan.pl

received 8 August 2022, revised 11 January 2023, accepted 22 January 2023

Abstract: Vehicle passive suspensions consist of two major elements generating force – spring and passive damper. Both possess non-linear characteristics, which are quite often taken into account in simulations; however, the friction forces inside the hydraulic damper and the damping force's hysteresis are usually left out. The researchers in this paper present the results of examination of the influence of using complex damper models – with friction and hysteresis; and with linear and non-linear static characteristics – on the chosen dynamic responses of a suspension system for excitations in the typical exploitation frequency range. The results from the simulation tests of the simplified and advanced versions of the damper model – different transfer functions and their relation to the reference model's transfer functions – are compared. The main conclusion is that friction and hysteresis add extra force to the already existing damping force, acting similar to damping increase for the base static characteristics. But this increase is not linear – it is bigger for smaller frequencies than for higher frequencies. The research shows the importance of including non-linear characteristics and proposed modules in modelling passive dampers.

Key words: vehicle vertical dynamics, damper model, friction, hysteresis, transfer function

1. INTRODUCTION

The car suspension – in a sense of mechanical components – is a group of elements such as springs, shock absorbers and linkages that connect the vehicle body or frame to its wheels and transmit longitudinal, lateral and vertical forces between them while simultaneously allowing relative vertical motion between the wheel and the body or frame of the vehicle.

In the sense of functionality in an area of vehicle vertical dynamics, the suspension must support both road holding/handling and ride quality, which are at odds with each other. The structure of suspension, wheels and the body forms an oscillatory system, affected by kinematic excitation z_r caused by the road surface profile (Fig. 1) and also by the force excitations caused by inertial forces acting on vehicle body during manoeuvres in the longitudinal or lateral directions (1).

In detail, the functionality of the suspension in the area of vertical dynamics involves keeping the car wheel in contact with the road surface as much as possible, because all the road or ground forces acting on the vehicle do so through the contact patches of the tyres. The suspension also protects the vehicle itself and any cargo or luggage from damage and wear, working as a vibro-isolation system that minimises the effects of road kinematic excitations.

The kinematic excitations processed by the suspension's dynamic structure produce responses such as bounce displacement of the sprung mass z_M and the unsprung mass z_m; the relative displacement of both masses – the suspension deflection (rattle space) $z_M - z_m$; the acceleration of the sprung and unsprung masses ($\ddot{z}_M$,$\ddot{z}_m$); as well as the forces F_t of tyre–road contact. The relations between road excitation and listed suspension responses, in a function of excitation frequency, usually are called suspension transmissibility functions (2), the dynamic characteristics of suspension, the frequency response function or the magnitude–frequency characteristics (3,4).

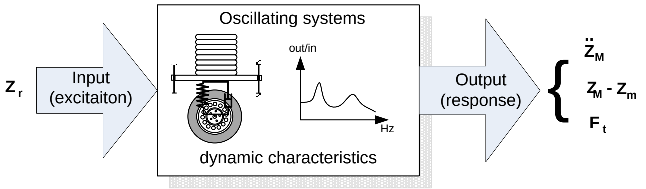


Fig. 1. A car suspension as an oscillatory system with its dynamic characteristics

The shape and amplitude values of the frequency response functions (transfer functions) allow the easy assessment of suspension performance in terms of the criteria of comfort, safety and suspension rattle space, which is a technical limitation of suspensions. Ride comfort can be assessed using the sprung mass acceleration transfer function $\ddot{z}_M$(ω), the safety potential using the dynamic wheel load transfer function $F_t(\omega)$ and technical limitations with the wheel rattle space amplification function. This function is also important from the point of view of the kinematic performance of the suspension – possible changes of wheel camber and steer angles – and the influence of vertical dynamics on lateral dynamics.

These functions can be shaped by tuning the parameters of the suspension components – sprung and unsprung masses, spring characteristics and shock absorber (damper) characteristics and the tyre stiffness and damping coefficients or characteristics. Thus, it can be said that synthesis of suspension performance involves finding the right compromise between fulfilling all the tasks in terms of vertical dynamics by setting adequate suspension stiffness and damping parameters or characteristics to fulfil all the tasks in an optimised way.

A lot of research concerning improving the ride comfort and handling performance has been done purely on theoretical, simplified models of vehicle vertical dynamics, usually linear, where the stiffness and damping characteristics are described by only linear parameters. This method of modelling suspensions is justified when using shock absorbers and springs with linear characteristics in earlier periods of motor car history and also while narrowing the range of analysis only to the use of the linear part of the spring or characteristics of the shock absorber. Finally, the use of linear parameters could be justified by the use of substitute linear parameters in place of real non-linear characteristics.

The experimental tests of shock absorber characteristics show that not only are non-linear static characteristics often not implemented in simple mathematical models but also additional phenomena, such as dry friction or hysteresis in a suspension, are not considered.

Thus, more advanced models include non-linear characteristics and/or asymmetrical characteristics. Sometimes, dry friction is added (5–7), while taking into account damping force hysteresis is rare. The authors therefore decided to research the influence of inclusion of friction and hysteresis modules into the shock absorber model on changes in the estimated frequency responses of suspensions with linear and non-linear static shock absorber characteristics.

2. RESEARCH METHOD

2.1. Research matrix and model parameters

In the research, computer simulation of a vertical dynamics model of a quarter car was implemented in MatLab/Simulink software. Specifically, modified quarter-car models with different damper models were used for the different cases, as presented in Tab. 1.

Tab. 1. Shared features of a quarter-car model used in the research

Case no.	Linear characteristic	Non-linear characteristic	Friction	Hysteresis
1	✔	–	–	–
2	✔	–	✔	–
3	✔	–	–	✔
4	✔	–	✔	✔
5	–	✔	–	–
6	–	✔	✔	–
7	–	✔	–	✔
8	–	✔	✔	✔

The investigation began by analysing Case 1 – the simplest possible damping characteristic – passive linear one with minimum (900 Ns/m) and maximum (3,100 Ns/m) damping coefficient. The case involving just the non-linear static characteristic without friction and hysteresis, namely Case 5, was chosen as the easiest way to model real damper characteristics in both maximum and minimum damping modes (Fig. 4). Cases 2–4 and 6–8 were also tested as passive suspension, with constant shock absorber static damping characteristics but including additionally either friction or hysteresis, or both of them.

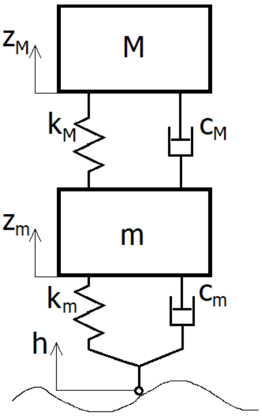


Fig. 2. Quarter-car model used in the simulations

The notations in the figure mean: M – the sprung mass, m – the unsprung mass, k_M – the suspension stiffness coefficient, k_m – the tyre stiffness coefficient, c_M – the suspension damping coefficient, c_m – the tyre damping coefficient, z_M – the sprung mass displacement, k_m – the unsprung mass displacement, h - the kinematic excitation.

All other suspension parameters were shared between all versions of the model: the linear stiffness characteristics of a tyre and suspension were used (Tab. 2), while the quarter-car model used is shown in Fig. 2.

Tested models were subjected to the excitation that enables calculation of the suspension dynamic responses in the form of transfer functions (frequency response function) between excitation and responses important for evaluation of suspension dynamic performance:

- suspension deflection for evaluation of necessary rattle space;
- sprung mass acceleration and sprung mass displacement for evaluation of ride comfort;
- cumulative tyre force for evaluation of safety potential.

Tab. 2. Shared features of a quarter-car model used in the research

Parameter	Unsprung mass	Sprung mass	Tyre stiffness	Tyre damping	Suspension stiffness
Unit	(kg)	(kg)	(kN/m)	(Ns/m)	(kN/m)
Value	50	400	200	350	30

The road excitation used in the research was a vertical sinusoidal displacement with a constant amplitude of 3 mm and a variable frequency starting from 0.0001 Hz up to 40 Hz (Fig. 3). It is similar to that used in European Shock Absorber Manufactures Association (EUSAMA) testers during periodical technical inspections of car suspensions.

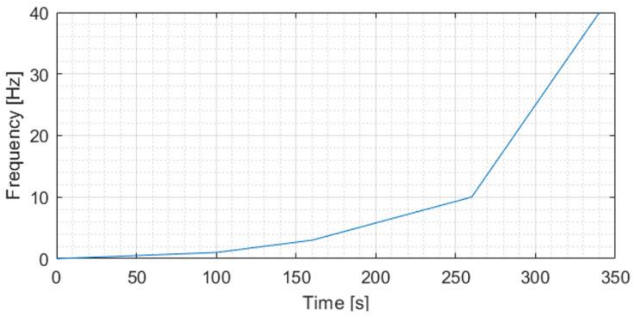


Fig. 3. Changes in the frequency of the input signal over simulation time

Important frequencies when analysing suspension dynamics cover a range from 0.5 Hz to 25 Hz. The frequency values changed in a non-linear fashion in order to allow more cycles to occur in a lower range, which gives better results when calculating transfer functions (8). Additionally, frequencies <0.5 and >25 Hz were added to the simulation in order to further stabilise the results of the tfestimate MatLab function, which was used to estimate the transfer function of the suspension.

2.2. Damper models including friction and hysteresis

The damper model for all cases, besides Case no. 1 (labelled as “linear”), had non-linear, asymmetric characteristics identified after empirical testing of damping forces of a real damper and averaging these forces to obtain the static characteristics (9) (Fig. 4).

All the damper model versions, besides Case no. 1, included a model of an adjustable damper with hysteresis, friction and also actuation delay modelled but not used in this research (Fig. 5).

Four main modules were applied to model the total damper force:

1. the static damping force,
2. hysteresis force,
3. friction force and
4. dynamic response.

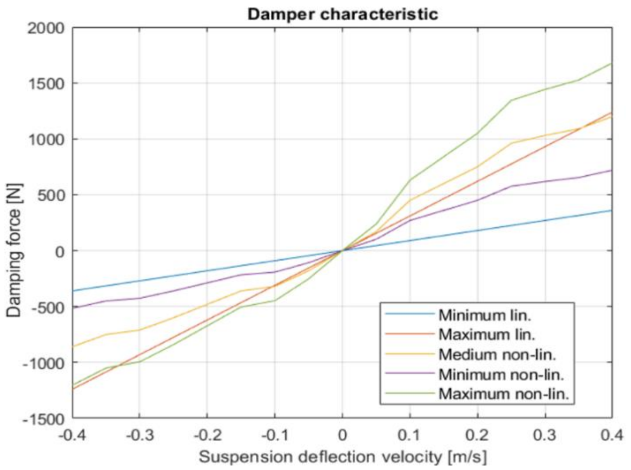


Fig. 4. Damper model static characteristics; lin. – linear, non-lin. – non-linear

The static damping characteristics module models the damping force as a function of deflection speed, differing for the compression and rebound and also on the control current if it models the electrically adjusted damper. Its implementation in Simulink is shown in Fig. 6.

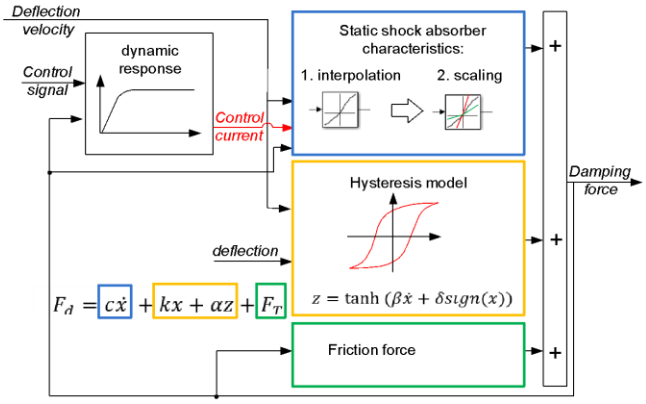


Fig. 5. Damper model diagram

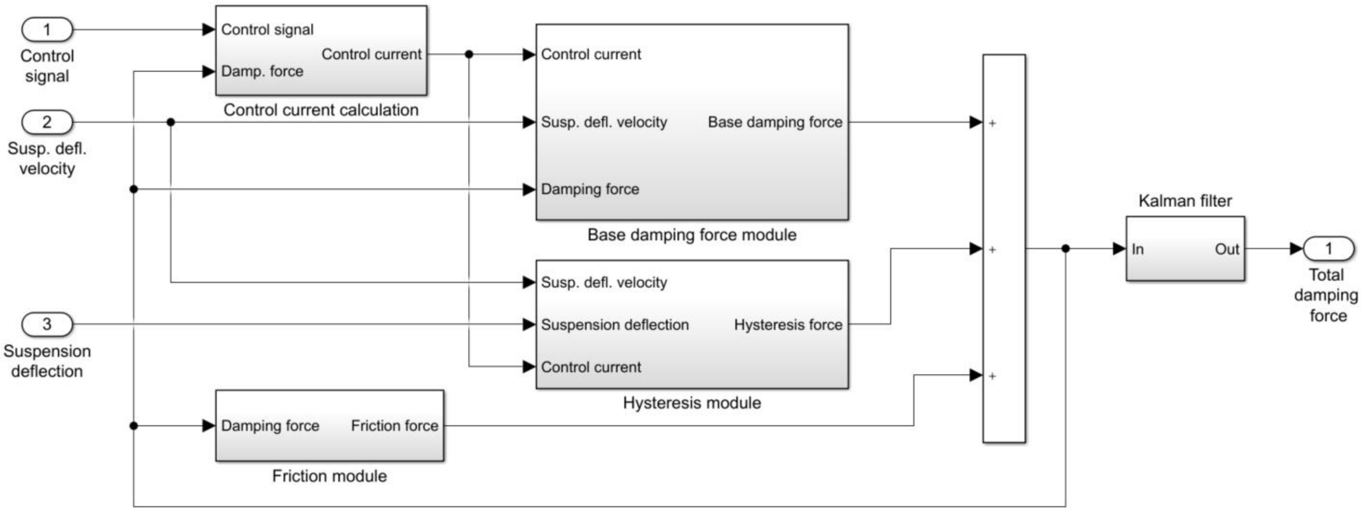


Fig. 6. Advanced damper simulation model
Damp. – Damping; Susp. defl. – Suspension deflection

For linear and symmetrical characteristics, the damping force can be modelled using simple equations (it was done for linear damper for Case no. 1):

$$F_d = c\dot{x}, \quad (1)$$

where c is the damping coefficient and $\dot{x}$ is the damper compression/extension velocity.

In cases of non-linear and asymmetrical characteristics (cases no. 2–4 and 6–8), interpolation of experimental characteristics was used by using the "Look-up table" block during MatLab/Simulink software implementation.

For the adjustable damper, the interpolation was also necessary for the value of damping force in relation to the control current. This also can be accomplished with two-dimensional "Look up table" block for the three-dimensional shock absorber characteristics. Assuming a linear relation between the control current and the damping forces, the medium damping F_{dS_m} static characteristic was used along with the coefficient K_I to increase or decrease damping force according to the value of the valve coil current and the state of the damper work – compression or rebound:

$$F_{dS\_I} = F_{dS\_m} \cdot K_I, \quad (2)$$

where: $F_{d\_S\_I}$ – interpolated value of damping force from the static characteristic for a given current, F_{dS_m} – the middle static characteristics damping force (for middle value of valve coil current), K_I – the coefficient to increase or decrease the damping force according to the value of valve coil current and state of damper work – compression or rebound. K_I values for the modelled shock absorber formulas, according to the value of the current I_c (0.6 ≤ I_c ≤1.6 A), were determined respectively for compression and rebound as follows:

$$\mathrm{K_{IC}} = -0.55\mathrm{I_c} + 1.59 \quad (3)$$

and

$$\mathrm{K_{IR}} = -0.71\mathrm{I_c} + 1.74 \quad (4)$$

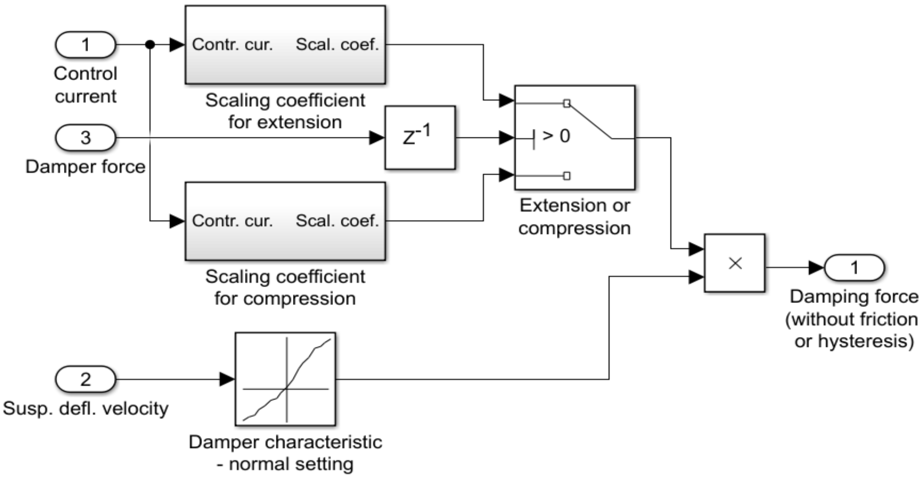

1
Control current
3
Damper force
Contr. cur. Scal. coef.
Scaling coefficient for extension
z^{-1}
> 0
Extension or compression
Contr. cur. Scal. coef.
Scaling coefficient for compression
×
1
Damping force (without friction or hysteresis)
2
Susp. defl. velocity
Damper characteristic - normal setting

Fig. 7. Base damping force calculation subsystem
Contr. cur., control current; Scal. coeff., scaling coefficient; Susp. defl., suspension deflection

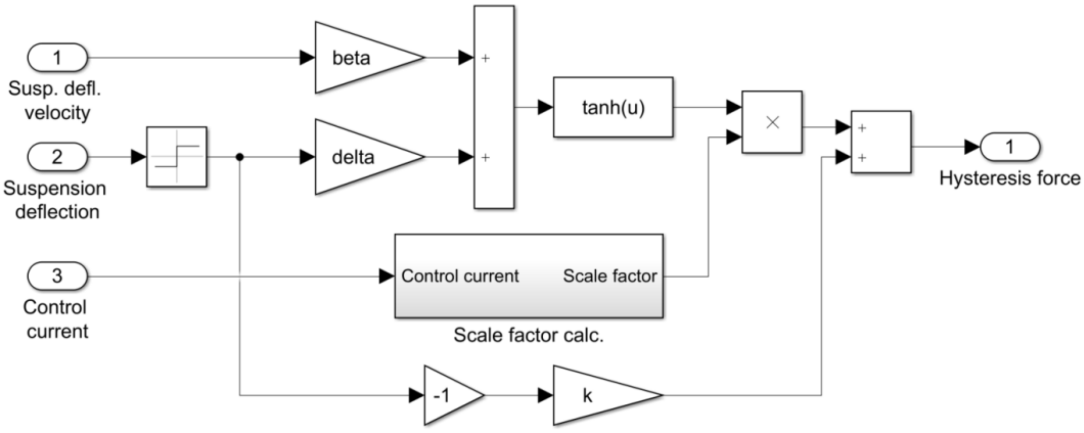


Fig. 8. Hysteresis module of the Simulink damper model
calc., calculation; Susp. defl., suspension deflection

The damper hysteresis module (Fig. 8) is important for high damping forces and high velocities. A simple model based on a previous work (10) was proposed to simulate the hysteretic force–velocity characteristic of the damper. This model is given by the following formulas:

$$F_h = kx + \alpha z, \quad (5)$$

$$z = F_0 \cdot tanh\big(\beta\dot{x} + \delta sign(x)\big), \quad (6)$$

where: k – the stiffness coefficient, which is responsible for the hysteresis opening found from the vicinity of zero velocity; a large value of k corresponds to the hysteresis opening of the ends; z – the hysteretic variable given by the hyperbolic tangent function; β – the scale factor of the damper velocity defining the hysteretic slope; a large value of β gives a steep hysteretic slope. δ – factor determining the width of the hysteresis through the term δsign(x); a wide hysteresis results from a large value of δ, α – scale factor of the hysteresis that determines the height of the hysteresis; its value depends on the control current.

Based on the dynamic characteristic analysis of the tested shock absorber, a formula for the relation between scale factor α and valve coil current I_C was developed as follows:

$$\alpha = \alpha_0 \cdot (-2.15 I_C + 4.45) \quad (7)$$

where: α_0 – scale factor α of the hysteresis for middle static characteristics damping force.

The hysteresis force's value was dependent on the suspension deflection and its velocity, as well as on the control current's value and a number of empirically obtained parameters [9].

The internal friction module (Fig. 9) models the force F_T and consists of two elements – the value of the kinetic friction force and a signum function due to the model friction force with opposite sign to the damping force. The friction force calculation depends on the suspension deflection velocity: if it was greater than a given threshold, then the friction force had a value equal to the defined kinematic friction (35 N); if it was smaller, then the kinematic friction value was multiplied by the ratio of current suspension deflection velocity to the threshold value.

$$F_f = \begin{cases} 35 \; if \; v_{defl} > 0.1 \; m/s \\ 35 \cdot \frac{v_{defl}}{0.1} \; if \; v_{defl} < 0.1 \; m/s \end{cases} \; [N] \quad (8)$$

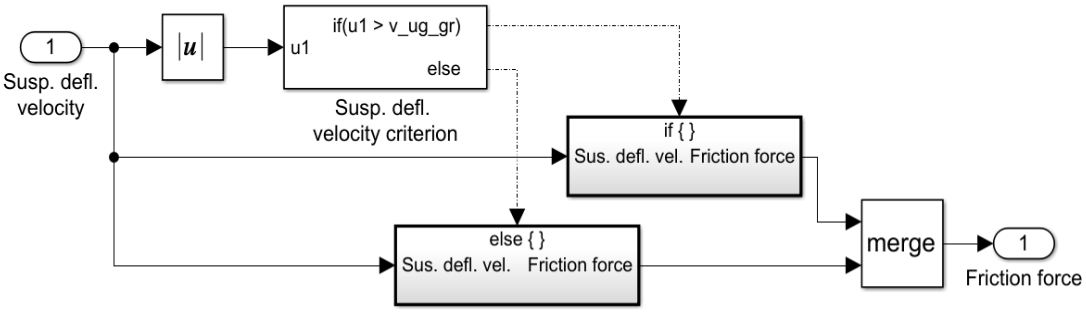

1
Susp. defl. velocity
|u|
if(u1 > v_ug_gr)
u1
else
Susp. defl. velocity criterion
if { }
Sus. defl. vel. Friction force
else { }
Sus. defl. vel. Friction force
merge
1
Friction force

Fig. 9. Friction force calculation subsystem
Susp. defl., suspension deflection; vel., velocity

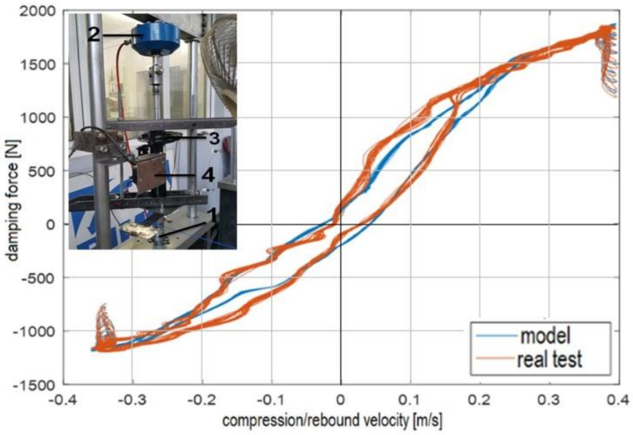


Fig. 10. Damper forces from the experimental test performed with the presented stand (1 – material test system [MTS] electrohydraulic excitator, 2 – force transducer, 3 –tested damper, 4 – travel sensor) and modelled

Forces modelled by the static non-linear model with added hysteresis and friction modules are presented in Fig. 10 and compared with the results from the experiment, as presented in the photograph in Fig. 10. Experimental tests of the characteristics of the three various types of shock absorbers, using a material test system (MTS) electrohydraulic actuator, performed by one of the authors, are described in a previous publication (11).

The comparison of damping forces calculated from static characteristics with those calculated from friction and hysteresis modules shows that the influence of including friction and hysteresis should be much bigger for lower frequencies due to their much larger share in the total damping force, as seen in Figs. 11 and 12.

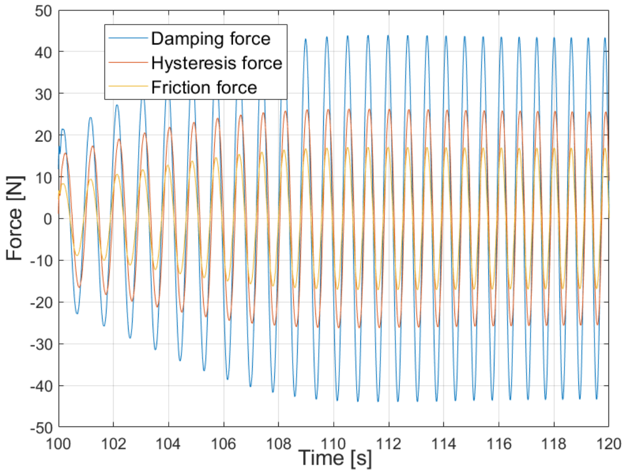


Fig. 11. Comparison of damping forces calculated from static characteristics with those calculated from friction and hysteresis modules for the first resonant frequency, which is approximately equal to 1 Hz

As previously stated, there were a few versions of a quarter-car model used in the research: one only static and linear version, which was set to maximum damping force; one with minimum force (reference model); and six versions of advanced models – linear and non-linear, which had switched-on modules of friction, hysteresis, or friction and hysteresis.

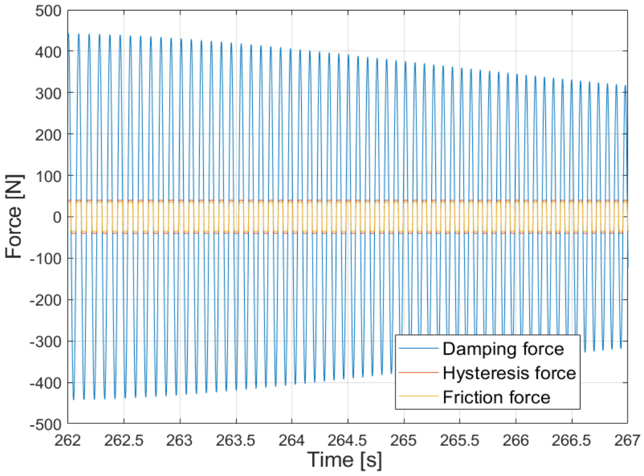


Fig. 12. Comparison of damping forces calculated from static characteristics with those calculated from friction and hysteresis modules for the second resonance: approximately 11 Hz

3. METHODOLOGY FOR ANALYSIS OF THE TESTING AND SIMULATION RESULTS

In the simulation tests, a quarter-car model that includes a non-linear damper model with controllable friction and hysteresis modules, which could be turned on or off, was used. Besides the damper module, the remainder of the model was linear, with the parameters of the model presented in Tab. 2. For each variant, the same excitation was applied – a changing-frequency sine wave of amplitude 3 mm, with the course of frequency's variability shown in Fig. 3.

The influence of friction and hysteresis implemented separately or simultaneously was analysed for three indicators: 1) suspension deflections; 2) cumulative force between the tyre and the road surface; and 3) sprung mass accelerations for linear passive dampers and non-linear passive dampers. The analysed indicators allow for the evaluation of the suspension performance in terms of ride comfort, ride safety and rattle space of suspension required for its work. The tools chosen for the analysis of suspension performance were the transfer functions (frequency responses) between the given indicator and the kinematic excitation. They were chosen because these indicators are not defined by a single value but are expressed as a function of the excitation frequency.

These functions were calculated using the response signals (deflection, cumulative tyre force and sprung mass acceleration) obtained during simulations, as would be done for an experimental testing of suspension frequency responses. These functions between particular responses and kinematic excitation were calculated using the MatLab tfestimate function. The reason why these functions were not calculated analytically using suspension element parameters was the non-linear character of the damper model. The resulting frequency responses were then plotted as graphs, which show their magnitude as a function of frequency, with frequencies ranging from 0.5 Hz to 25 Hz being investigated. The results for the relative values between a given case and the reference model also were used.

4. RESULTS

4.1. Influence of friction and hysteresis on a linear damper with minimum damping coefficient

The investigation began by analysing the influence of inclusion of friction force, hysteresis and their combined effects on a suspension with the simplest possible damping characteristic – passive linear one with minimum (900 Ns/m) and maximum (3,100 Ns/m) damping coefficient. This was meant both to test the method of creating the transfer function graphs and to observe the effects of hysteresis and friction on the frequency response of a model. This resulted in the overall shape and behaviour of the linear damper matching the expectations and theory that was previously presented in literature (3,12) (blue characteristics in Figs. 13–18) – the reference transfer function for the model considering only the damping force.

First, the influence of friction force, hysteresis and their combined effects on the behaviour of the linear damper of minimum assumed damping, with damping coefficient of 900 Ns/m (dimensionless damping coefficient of 0.13) was analysed.

The resulting transfer functions for the suspension deflection are shown in Fig. 13.

For other suspension performance assessment criteria, analogous information was extracted from the graphs for the most crucial frequencies:

- sprung mass resonant frequency – approximately 1 Hz;
- unsprung mass resonant frequency – approximately 10 Hz;
- frequency between resonances for sprung and unsprung masses – approximately 3 Hz; and
- maximum tested frequency – 25 Hz.

The relative values were calculated as the ratio between a given dynamic response's value for a damper model with friction, hysteresis or both of them over the corresponding value for a simplified model (only static damper characteristics).

The suspension deflection transfer function analysis showed that adding hysteresis and friction to the damping force yields similar results as the increasing of the damping coefficient – for minimal damping, lowering the magnitude from 3.2 m/m to around 1.9 m/m for the first resonant frequency and from 3 m/m to 2.5 m/m for the second frequency (Fig. 13a).

Comparing the suspension deflection transfer function values for the four chosen frequency ranges, it is visible that the biggest influence of friction and hysteresis can be seen for unsprung and (especially) sprung mass resonant frequencies, with hysteresis affecting the results more than does friction (Fig. 13b – in relation to the reference model).

The next analysis was done for the transfer function of cumulative tyre force, which is synonymous with the contact force between the road and the tyre, which is used as an indicator for ride safety criteria. Adding friction and hysteresis decreases the values of the transfer function in the first and second resonance ranges: in the first, even to almost 40%; in the second, from about 6% to 12% (Fig. 14). The biggest influence is for the frequencies between the first and second resonances – from 40% to 90% increase – but the absolute values are still much smaller (>10 times) than for the second resonance.

The analysis of transfer function of sprung mass acceleration (Fig. 15), which is used as an indicator for the ride comfort criterion, shows that the influence of friction and hysteresis is visible for three of the four analysed frequency ranges. For two of them, the sprung mass acceleration is greater than for the reference model. For the sprung mass resonant frequency range, the trend is reversed – the accelerations are smaller in value, which translates to a better ride comfort for low frequencies, while being worse for higher ones. Once again, hysteresis plays a bigger role than friction for all frequencies.

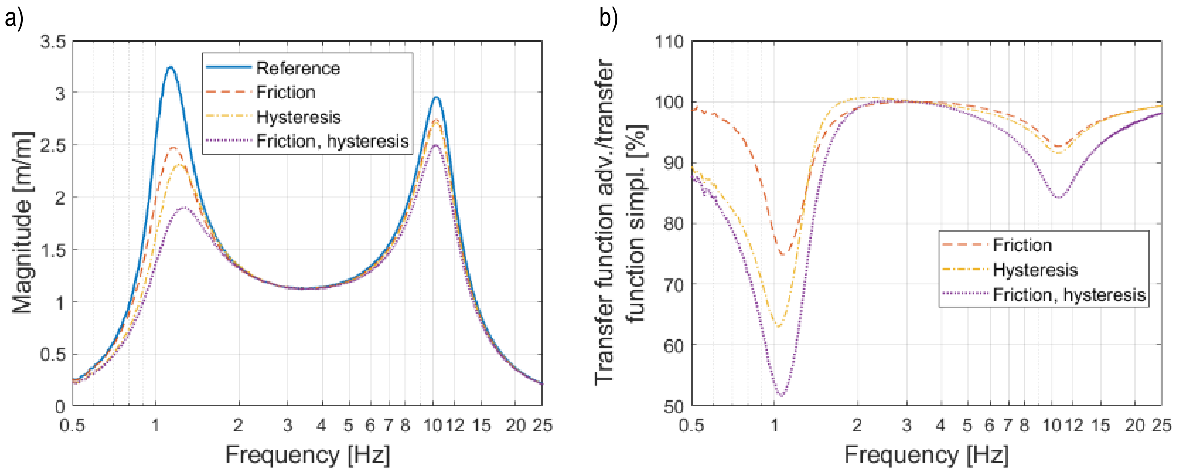


Fig. 13. Transfer function (a) between road excitation and suspension deflection and ratio (b) between suspension deflection transfer functions of reference (simpl.) and tested (adv.) models for minimal linear damping - 900 Ns/m (adv., advanced; simpl., simplified)

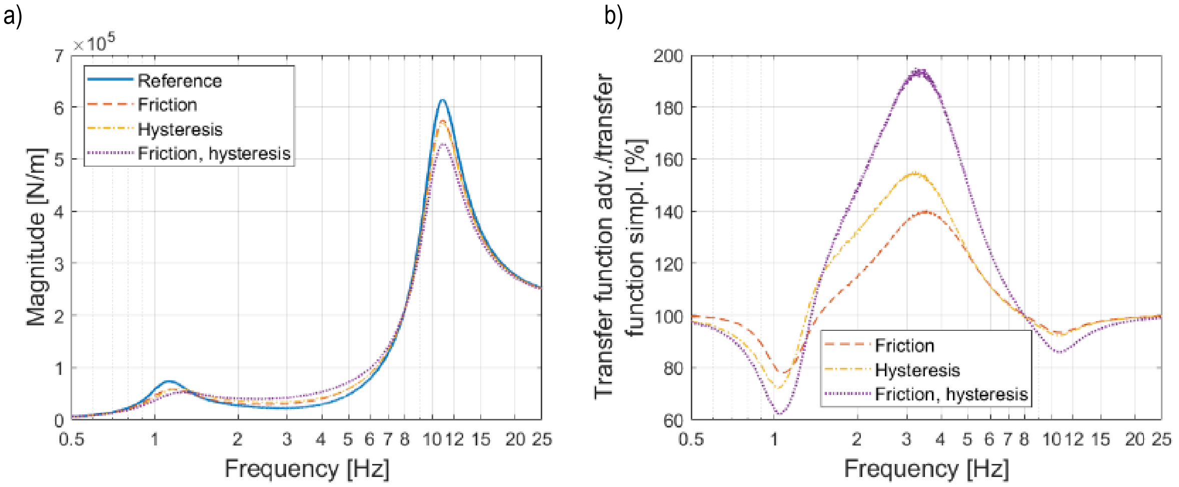


Fig. 14. Transfer function (a) between road excitation and cumulative tyre force and ratio (b) between cumulative tyre force transfer functions of reference (simpl.) and tested (adv.) models for minimal linear damping - 900 Ns/m (adv., advanced; simpl., simplified)

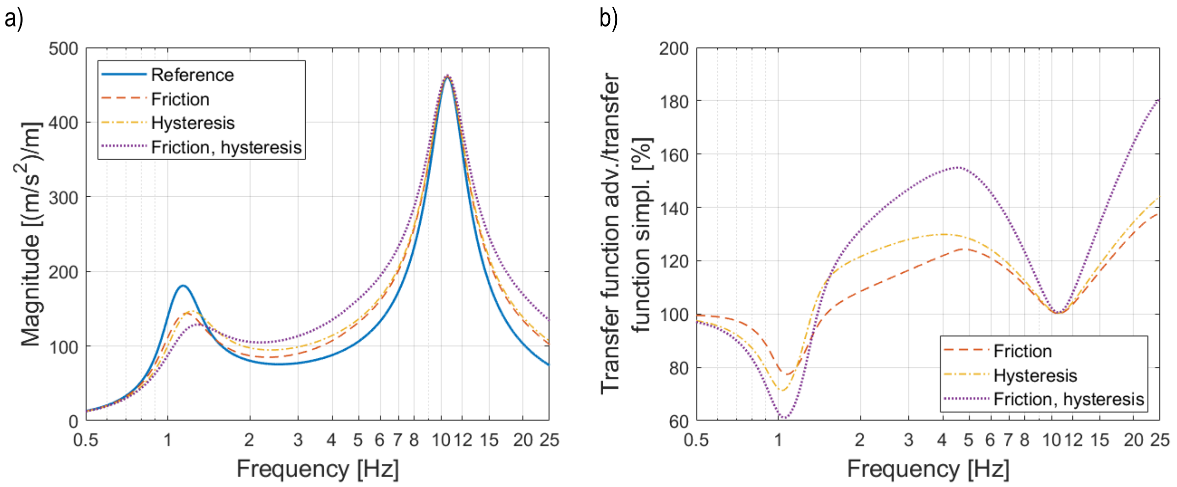


Fig. 15. Transfer function (a) between road excitation and sprung mass acceleration and ratio (b) between sprung mass acceleration transfer functions of reference (simpl.) and tested (adv.) models for minimal linear damping - 900 Ns/m (adv., advanced; simpl., simplified)

4.2. Influence of friction and hysteresis on a linear damper with maximum damping coefficient

Similarly to the linear damper of the minimum damping coefficient, the influence of the components of an advanced damper model on the transfer functions for different dynamic responses was analysed for a suspension with linear damping of maximum damping coefficient of 3,100 Ns/m (dimensionless damping coefficient of 0.45).

In the case of maximal damping and suspension deflection transfer function analysis (Fig. 16), it is visible that adding hysteresis (which also had a bigger effect on the minimal damping case than friction) causes the gain to drop to <1.1 m/m for all frequencies and removes two separate resonances, leaving the one around 3 Hz.

The suspension deflection transfer function values for the four analysed frequency ranges and the comparison for those frequencies with the reference model (Fig. 17) show that the range around 3 Hz is not sensitive to the inclusion of friction and hysteresis. For ranges of lower and higher frequencies, the influence of hysteresis is much bigger than the influence of friction. Hysteresis decreases the amplitudes of the first resonance about 40% and of the second about 25%, while reducing friction by about 10% and 7%, respectively (Fig. 17).

The next transfer function analysed was the cumulative tyre force for maximum linear damping. Compared to the minimum damping coefficient, a big difference is visible for low frequencies (Fig. 14 vs Fig. 17). In the first resonant frequency, the transfer function values are actually greater for the more advanced model by around 25% (Fig. 17), while for minimal damping, they are lower by around the same amount for friction and hysteresis, and even almost 40% lower for friction and hysteresis (Fig. 14). Friction's influence is also less pronounced for all frequencies, which is linked with its maximum value (35 N) being smaller, in comparison with the damping forces for higher damping coefficients.

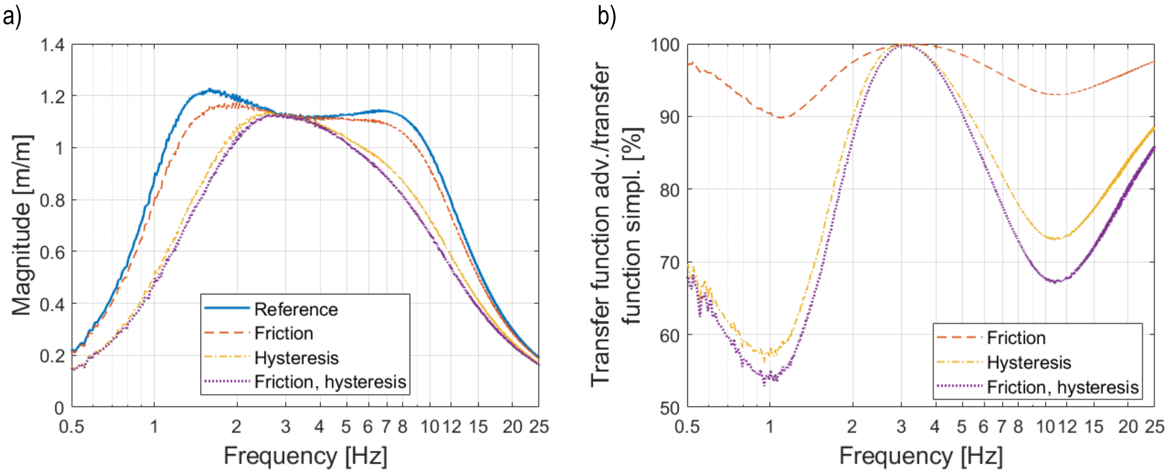


Fig. 16. Transfer function (a) between road excitation and suspension deflection and ratio (b) between suspension deflection transfer functions of reference (simpl.) and tested (adv.) models for maximal linear damping – 3,100 Ns/m (adv., advanced; simpl., simplified)

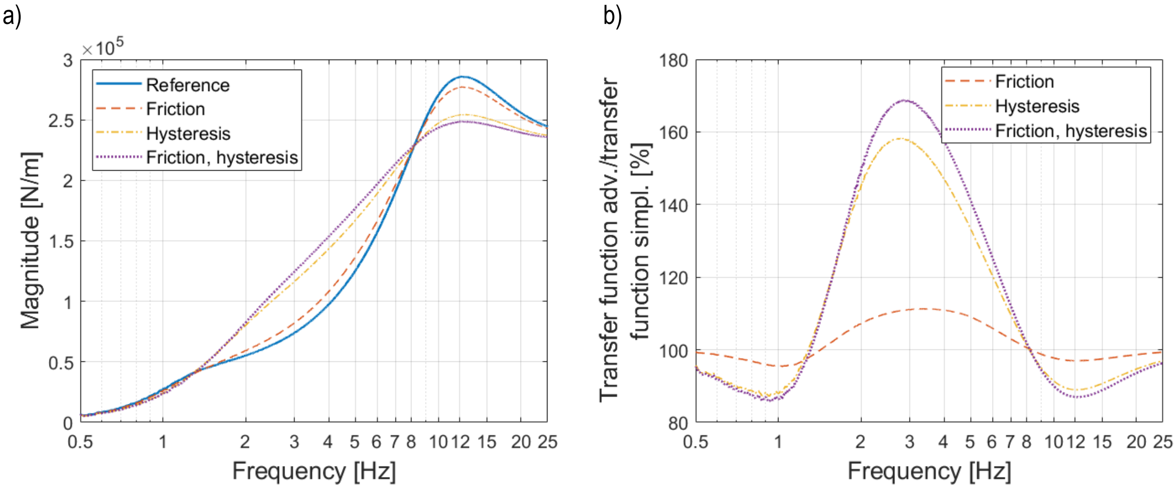


Fig. 17. Transfer function (a) between road excitation and cumulative tyre force and ratio (b) between cumulative tyre force transfer functions (adv., advanced; simpl., simplified)

The results for sprung mass accelerations are similar to those for the cumulative tyre force for 1 Hz, while the higher frequencies show higher influence of friction and hysteresis for almost all cases compared to the results for minimum linear damping (Fig. 18). The effects are best visible for 3 Hz, while also being very noticeable for 25 Hz. For all frequencies, the transfer function values are higher than in a reference model, indicating higher sprung mass accelerations, which translates to lower ride comfort.

Seeing these effects, a question arose regarding whether it is possible to find an equivalent damping coefficient, which could be used instead of more complicated models with friction and hysteresis modules. Simulation was conducted for a number of values (Fig.19), and the results suggest that finding just one value that could simulate the behaviour of a model with friction and damping is impossible – big differences will arise either in the first or second resonant frequency. The potential solution to this problem would be the use of bilinear models, which have been previously extensively studied (8).

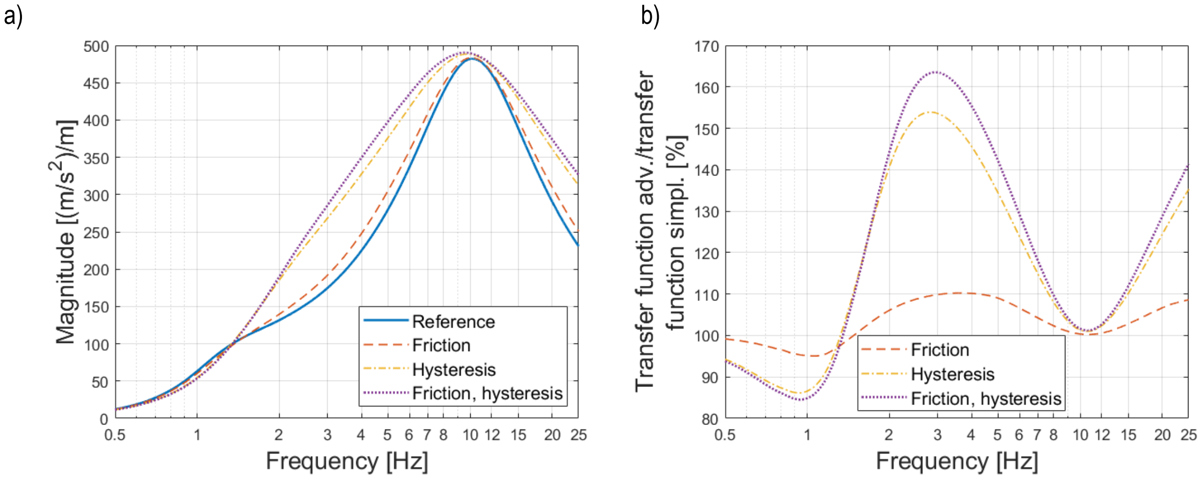


Fig. 18. Transfer function (a) between road excitation and sprung mass acceleration and ratio (b) between sprung mass acceleration transfer functions of reference (simpl.) and tested (adv.) models for maximal linear damping – 3,100 Ns/m (adv., advanced; simpl., simplified)

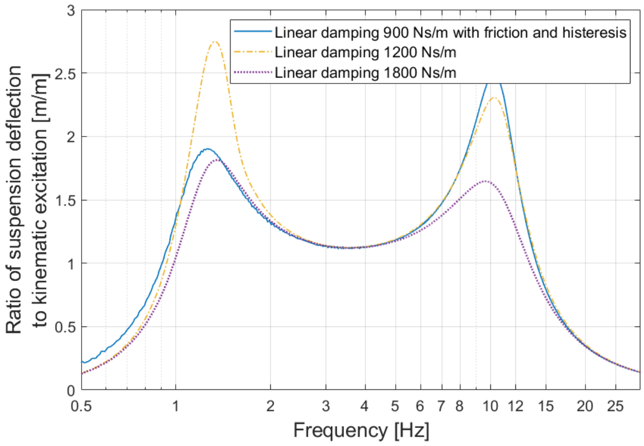


Fig. 19. Search for equivalent linear damping to friction and hysteresis - damping coefficient 1,800 Ns/m and 1,200 Ns/m

4.3. Influence of friction and hysteresis on a non-linear damper with the lowest damping mode

The next step was to look at the effects of friction and hysteresis on passive models that had non-linear characteristics implemented – for both the highest and the lowest possible damping modes. The results proved to be similar to those achieved in a fully linear model – friction and hysteresis acted as if the damping force was increased. It can be also observed that the total effect of both those forces is almost equal to a sum of both of them for all analysed suspension responses. Both friction and hysteresis had a slightly bigger impact near the first resonant frequency for minimal damping mode, decreasing the values of all analysed responses – tyre force, suspension deflection and both sprung mass acceleration and displacement.

For the minimum non-linear damper model, the influence of friction and hysteresis on suspension deflections is less pronounced than for a linear model for 1 Hz, while being almost identical in other frequencies (Fig. 20). In general, the influence of a more advanced model does not exceed 20% for the minimum damping non-linear model.

Hysteresis and friction do not show a substantial influence on the cumulative tyre forces for the minimum damping non-linear model (especially compared to the linear version), the biggest effects being seen for the 3-Hz range, where the transfer function is higher by 35% for both forces implemented into the model.

For the remaining frequencies, the resulting cumulative tyre force transfer functions are slightly lower than in the reference model (Fig. 21).

Transfer functions for sprung mass accelerations show slightly bigger influence of friction and hysteresis (Fig. 22). The difference is that models with friction and hysteresis have transfer functions bigger than the reference model – especially for 3 Hz and 25 Hz.

Summarising the analysis for all responses, we can state that as the frequency rises, all the relations between the transfer functions shifted to values >1 (meaning that the dynamic responses for model with friction and/or hysteresis were greater than those for the reference ones), with the exception of suspension deflection, which remained <1 for the entire range. Maximum gain was achieved for all transfer functions for 3 Hz. After the first resonant frequency, the response for the maximum damping model was slightly greater compared to that for minimal damping – being smaller by 5–10 percentage points in the maximum for the model with both friction and hysteresis.

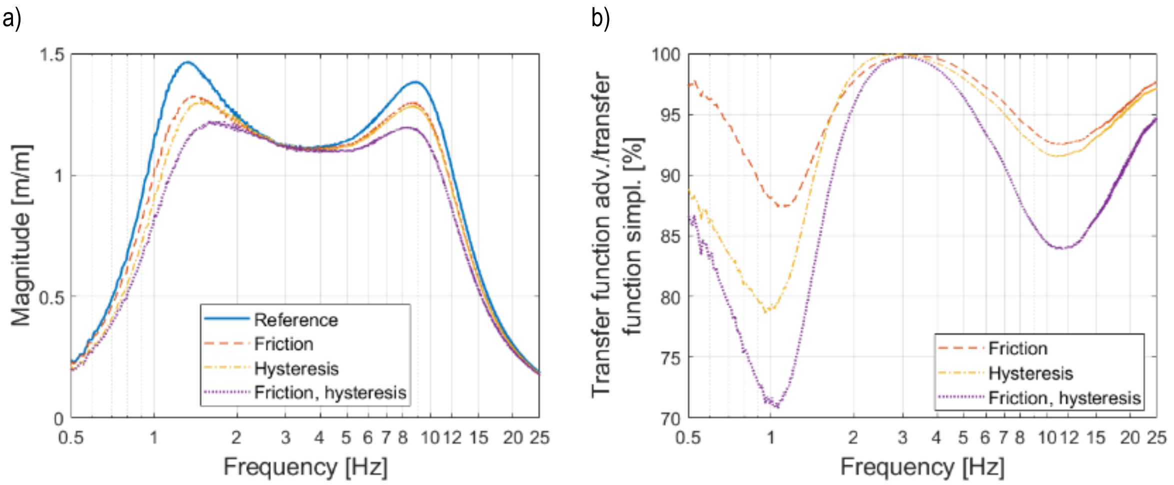


Fig. 20. Transfer function (a) between road excitation and suspension deflection and ratio (b) between suspension deflection transfer functions of reference (simpl.) and tested (adv.) models for minimal non-linear damping (adv., advanced; simpl., simplified)

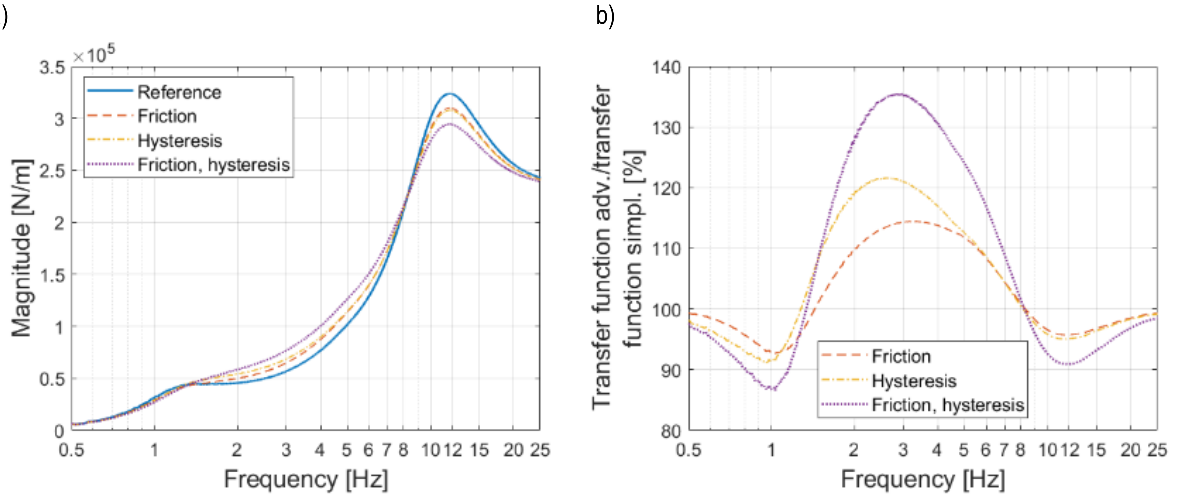


Fig. 21. Transfer function (a) between road excitation and cumulative tyre force and ratio (b) between cumulative tyre force transfer functions of reference (simpl.) and tested (adv.) models for minimal non-linear damping (adv., advanced; simpl., simplified)

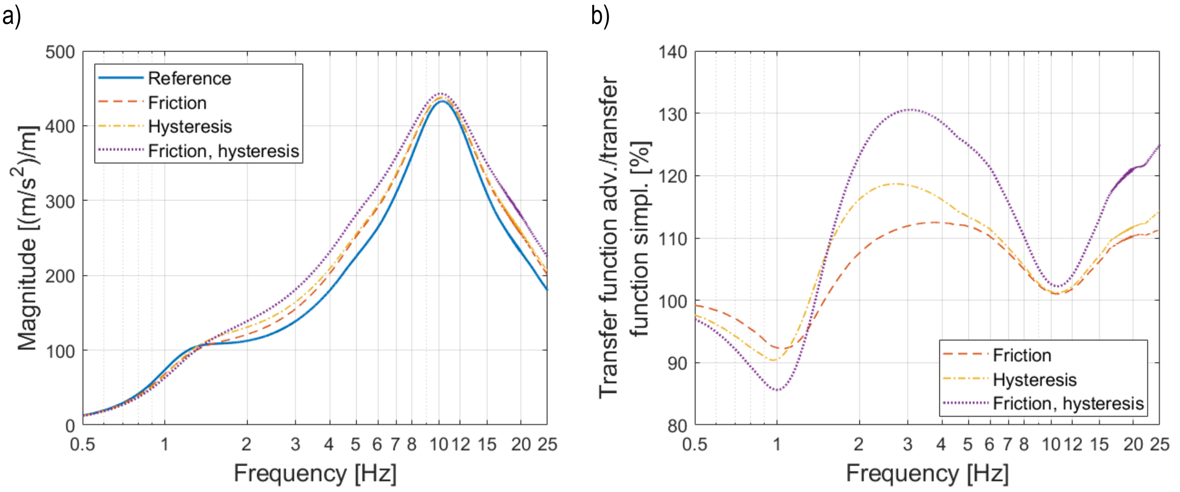


Fig. 22. Transfer function (a) between road excitation and sprung mass acceleration and ratio (b) between sprung mass acceleration transfer functions of reference (simpl.) and tested (adv.) models for minimal non-linear damping (adv., advanced; simpl., simplified)

4.4. Influence of friction and hysteresis on a non-linear damper with highest damping mode

For the maximum non-linear damping model, the effects of hysteresis are much more visible for the entire range of frequencies, with the exception of frequencies around 3 Hz (Fig. 23). This influence ranges from 32% for the first resonance to around 22% for the second resonance, while the influence of friction does not exceed 6%. Additional elements in the damper model for suspension deflection cause the transfer function values to drop below reference values.

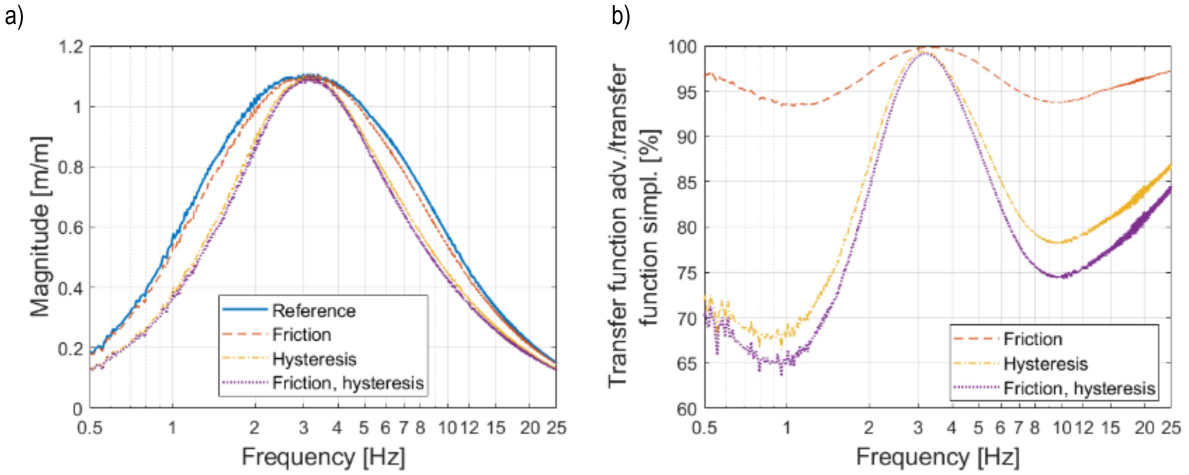


Fig. 23. Transfer function (a) between road excitation and suspension deflection and ratio (b) between suspension deflection transfer functions of reference (simpl.) and tested (adv.) models for maximal non-linear damping (adv., advanced; simpl., simplified)

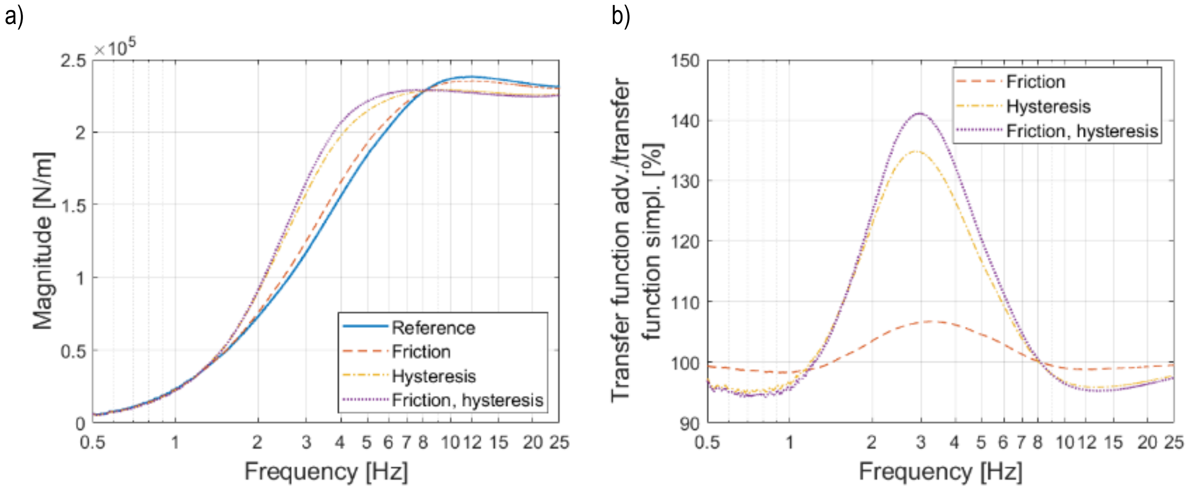


Fig. 24. Transfer function (a) between road excitation and cumulative tyre force and ratio (b) between cumulative tyre force transfer functions of reference (simpl.) and tested (adv.) models for maximal non-linear damping (adv., advanced; simpl., simplified)

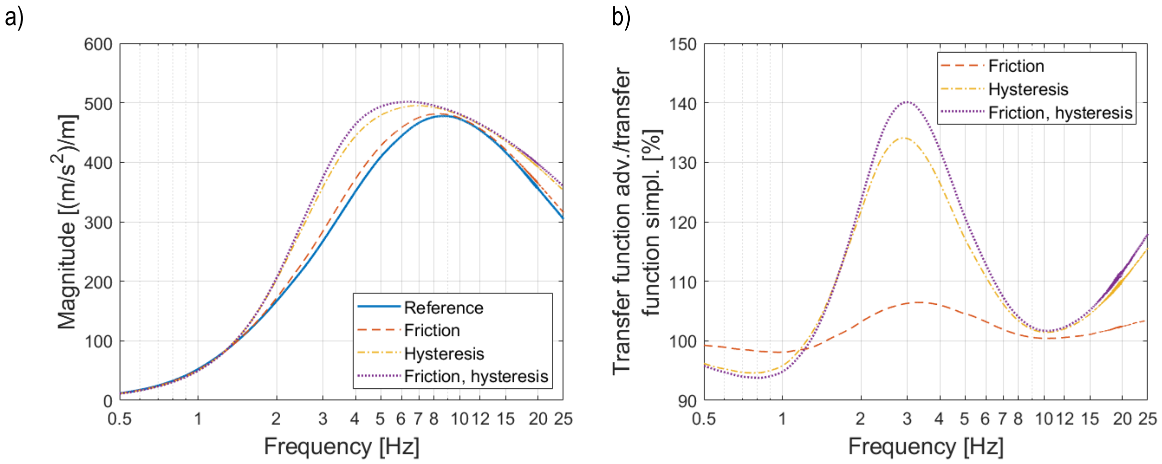


Fig. 25. Transfer function (a) between road excitation and sprung mass acceleration and ratio (b) between sprung mass acceleration transfer functions of reference (simpl.) and tested (adv.) models for maximal non-linear damping (adv., advanced; simpl., simplified)

The relative values for the maximum non-linear model's cumulative tyre force (Fig. 24) are analogous to the ones for minimum non-linear damping – the transfer function values are greater for the 3-Hz range while remaining almost unchanged for the remaining frequencies. Hysteresis once again has a bigger impact than friction.

The transfer functions for sprung mass accelerations show a slightly bigger influence of friction and hysteresis (Fig. 25). The difference relative to the cumulative tyre force is that the influence of friction and hysteresis is similar to and slightly bigger for the ranges of approximately1 Hz and 3 Hz, bigger for approximately 10 Hz and much bigger for approximately 25 Hz.

5. SUMMARY RESULTS

After the analysis of each of the four cases (minimum linear, maximum linear, minimum non-linear and maximum non-linear), a synthesis of the influence was done. For better visualisation of the influence of friction, hysteresis and both combined on different indicators of suspension performance, summary charts were prepared (Figs. 26–28). These charts present the results of the relative values between a given case and the reference model for four chosen frequencies – near the first resonant frequency (approximately 1 Hz), 3 Hz, the second resonant frequency (around 10 Hz) and the maximum tested frequency (25 Hz).

Suspension deflection transfer functions are less sensitive to added friction or hysteresis in the range between the first and second resonances. The biggest influence is visible for the first resonance (approximately 1 Hz), whereby adding friction decreases the transfer function values from 25% to 48% compared to the linear static model. The influence for the non-linear minimum damping is smaller – the decrease is about 13%–29%.

The second resonance shows a smaller influence – the decrease is from 7% to 15% for linear models and from 6% to 26.33% for non-linear models. Changes in the values of the transfer functions for the frequencies about 25 Hz are similar to those for the second resonance, but the influence is half the value (3%–15%) for non-linear models and much smaller compared to linear models (1%–6%).

The transfer functions for the cumulative tyre forces are most sensitive to added friction or hysteresis in the range between the first and second resonances. Added friction or hysteresis increases the transfer function values from 7% to 92% compared to the linear static model. The biggest influence is observed for the minimal damping models – especially the minimum damping linear model.

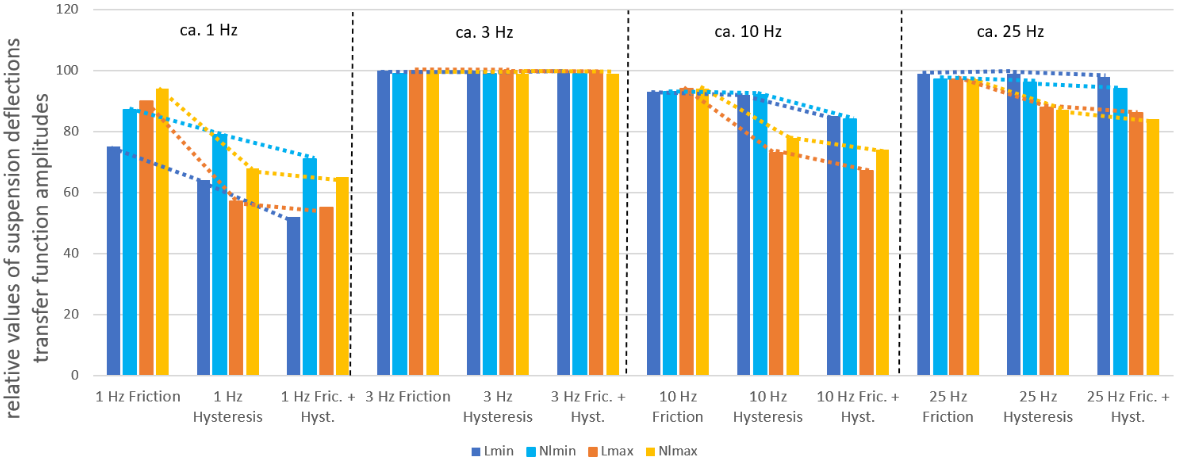


Fig. 26. Relations of the suspension deflection transfer function values of the tested (adv.) model to the values of the reference (simpl.) models for different damping characteristics and models at selected frequency regions (adv., advanced; simpl., simplified)

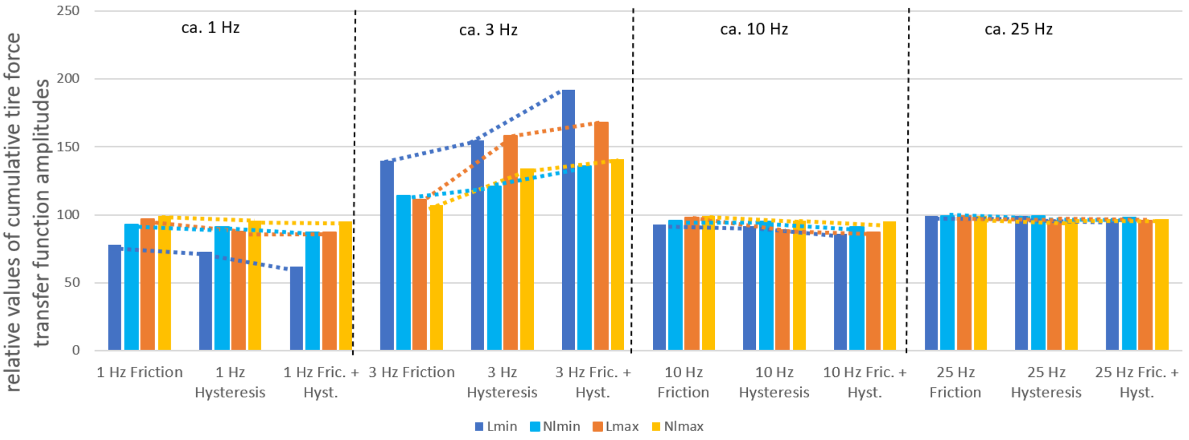


Fig. 27. Relations of the cumulative tyre force transfer values of the tested (adv.) model to the values of the reference (simpl.) models for different damping characteristics and models at selected frequency regions (adv., advanced; simpl., simplified)

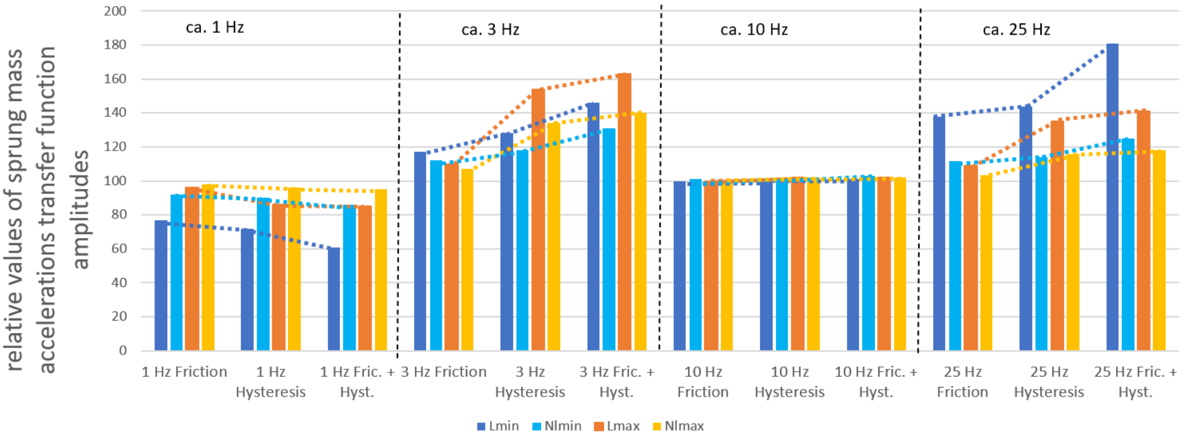


Fig. 28. Relations of the sprung mass acceleration transfer function values of the tested (adv.) model to the values of the reference (simpl.) models for different damping characteristics and models at selected frequency regions (adv., advanced; simpl., simplified)

The next level of influence is in the first resonance – it results in a decrease of the transfer function values from 22% to 38% for the minimum linear model and from 1% to 23% for the other models. It is less visible for the maximum damping non-linear model.

In the range of the second resonance (unsprung mass resonance), adding friction and hysteresis also decreases the transfer function values but by a lesser extent – from 4% to 14% for the minimal damping linear models and from 1% to 9% for the maximal damping linear models.

In the case of sprung mass accelerations used to assess the ride comfort, the increase in transfer function values are visible in the range between the first and second resonances and around the 25 Hz range. This increase is bigger for the minimum damping models – from 12% to 63% for the first resonance range and from 9% to 81% for the range of approximately 25 Hz.

There is almost no influence on the sprung mass transfer function values in the range of the second resonance – only a small increase of about 1%–3% is seen. For the range of the first resonance, a small decrease is visible.

6. CONCLUSION

During the research, the influence of inclusion of friction, hysteresis and combined friction and hysteresis into models of linear and non-linear damping forces on the three most important dynamic responses of a passive quarter-car suspension was tested.

The obtained results allow the assessment of this influence in both qualitative and quantitative ways. The quantitative influence is described in more detail in the "Results" and "Summary results" sections.

In general, it can be assumed that the effect of including friction and hysteresis gives an effect similar to that found on increasing the damping force; however, this effect is not linear, which results from the different influence of additional forces in the cases of low and higher operating speeds of the shock absorber (Figs. 11 and 12).

The differences in the influence of added friction and hysteresis are, in general, similar for both the minimum and maximal damping characteristics of the tested models. Bigger changes are visible between the minimum linear and non-linear models due to the fact that the linear model has a low damping coefficient for all working speeds. In the case of the non-linear model, for the lowest speed, the damping coefficient is higher than for the higher speed (Fig. 4).

In the case of suspension deflection transfer functions, the addition of friction and hysteresis for all frequencies decreases or does not change (in the range of approximately 3 Hz) the transfer function values. The biggest decrease (up to almost 48% compared to the reference linear model) is observed for the first resonance, then for the second resonance and then the range of approximately 25 Hz.

In the case of the cumulative tyre force, which is used to assess the safety potential, the biggest negative impact on safety is visible in the range of approximately 3 Hz. However, in that range, only the relative gain is large. The absolute values of this transfer function for the said frequency range are small in general, and the only important one is the second resonance range (approximately 10.12 Hz) – the decrease in the amplitudes of the transfer function is from 2% to 14% for linear and from 1% to 9% for non-linear damping characteristics.

For assessing comfort, the sprung mass acceleration transfer function is used, and the range of the second resonance is the most important as the absolute values of this function are the biggest. Added friction and hysteresis cause only a small (1%–3%: Fig. 28) increase. A much bigger influence (increase by 63% to 81%) is found for ranges of approximately 3 Hz and approximately 25 Hz. Yet, the absolute values are much smaller than for the second resonant frequency.

The research shows the importance of including the proposed modules (friction and hysteresis) in the damper model analysing the transfer functions for passive dampers. It can explain some differences observed between the dynamic responses of suspension models with only static characteristics and responses of a real suspension. This is all the more important because, currently, hydraulic gas-pressure shock absorbers with non-linear static characteristics are used in modern cars.

REFERENCES

1. Ślaski G. Studium projektowania zawieszeń samochodowych o zmiennym tłumieniu. Wydawnictwo Politechniki Poznańskiej. Poznań: Wydawnictwo Politechniki Poznańskiej; 2012. 399–404 p.
2. Grajnert J. Izolacja drgań w maszynach i pojazdach. Wydawnictwo Politechniki Wrocławskiej; 1997.

3. Mitschke M. Dynamika samochodu t.2 Drgania. Warszawa: Wydawnictwo Komunikacji i Łączności; 1989.
4. Reński A. Bezpieczeństwo czynne samochodu. Zawieszenia oraz układy hamulcowe i kierownicze. Oficyna Wydawnicza Politechniki Warszawskiej; 2011.
5. Zdanowicz P, Lozia Z. Wyznaczenie optymalnej wartości współczynnika asymetrii amortyzatora pasywnego zawieszenia samochodu z wykorzystaniem modelu „ćwiartki samochodu". Pr Nauk Politech Warsz Transp. 2017;(119):249–65.
6. Zdanowicz P. Ocena stanu amortyzatorów pojazdu z uwzględnieniem tarcia suchego w zawieszeniu. Politechnika Warszawska; 2012.
7. Lyu D, Zhang Q, Lyu K, Liu J, Li Y. Influence of the Dry Friction Suspension System Characteristics on the Stick-Slip of Vertical Vibration of a Three-Piece Bogie. Shock Vib. 2021;2021.
8. Ślaski G, Klockiewicz Z. the Influence of Shock Absorber Characteristics' Nonlinearities on Suspension Frequency Response Function Estimation and Possibilities of Simplified Characteristics Modelling. Arch Automot Eng. 2022;96(2):77–95.
9. Dąbrowski K. Algorytmizacja adaptacyjnego sterowania tłumieniem zawieszenia samochodu dla uwzględnienia zmienności warunków eksploatacji. 2018.
10. Kwok NM, Ha QP, Nguyen TH, Li J, Samali B. A novel hysteretic model for magnetorheological fluid dampers and parameter identification using particle swarm optimization. Sensors Actuators, A Phys. 2006;132(2):441–51.
11. Więckowski D, Dąbrowski K, Ślaski G. Adjustable shock absorber characteristics testing and modelling. IOP Conf Ser Mater Sci Eng. 2018;421(2).
12. Savaresi SM, Poussot-Vassal C, Spelta C, Sename O, Dugard L. Semi-Active Suspension Control Design for Vehicles. 2010.

Zbyszko Klockiewicz: https://orcid.org/0000-0003-4353-550X

Grzegorz Ślaski: https://orcid.org/0000-0002-6011-6625

sciendo

DOI 10.2478/ama-2023-0025

FINITE LENGTH TRIPLE ESTIMATION ALGORITHM AND ITS APPLICATION TO GYROSCOPE MEMS NOISE IDENTIFICATION

Michal MACIAS*, Dominik SIEROCIUK*

*Institute of Control and Industrial Electronics, Warsaw University of Technology, ul. Koszykowa 75, 00–662 Warsaw, Poland

michal.macias@pw.edu.pl, dominik.sierociuk@pw.edu.pl

received 31 October 2022, revised 26 December 2022, accepted 4 January 2023

Abstract: The noises associated with MEMS measurements can significantly impact their accuracy. The noises characterised by random walk and bias instability errors strictly depend on temperature effects that are difficult to specify during direct measurements. Therefore, the paper aims to estimate the fractional noise dynamics of the stationary MEMS gyroscope based on finite length triple estimation algorithm (FLTEA). The paper deals with the state, order and parameter estimation of fractional order noises originating from the MEMS gyroscope, being part of the popular Inertial Measurement Unit denoted as SparkFun MPU9250. The noise measurements from x, y and z gyroscope axes are identified using a modified triple estimation algorithm (TEA) with finite approximation length. The TEA allows a simultaneous estimation of the state, order and parameter of fractional order systems. Moreover, as it is well-known that the number of samples in fractional difference approximations plays a key role, we try to show the influence of applying the TEA with various approximation length constraints on final estimation results. The validation of finite length TEA in the noise estimation process coming from MEMS gyroscope has been conducted for implementation length reduction achieving 50% of samples needed to estimate the noise with no implementation losses. Additionally, the capabilities of modified TEA in the analysis of fractional constant and variable order systems are confirmed in several numerical examples.

Key words: fractional calculus, fractional Kalman filter, estimation of fractional order systems, fractional order noise

1. INTRODUCTION

The fractional calculus (FC) is, in itself, an extension of traditional differential and integral calculus. The differential orders in FC can be real or even complex numbers. The fractional derivative appeared for the first time in the correspondence between Leibniz and l'Hôpital in 1695, and thereby, it appeared almost simultaneously with the integer order calculus. The theoretical background for this calculus can be found in several already classic works in the literature [1, 2, 3, 4, 6]; additionally, in multiple relatively recently published books [7,8], some applications of this calculus have been explored.

The main advantage of fractional order operators in comparison to the integer order case is that the fractional order derivatives depend not only on local time conditions but also on the whole past of the function [10]. This property can be especially useful for the description of dynamics possess with a long-term memory nature. The FC was found to be especially efficient in modelling diffusive systems [11, 12, 13, 14]. For example, in the heat transfer process of the solid beam, it is possible to describe dynamics between temperature and heat flux at the desired point as a half-order integral. When the heated material is not solid (heterogeneous), the order of the integration can be different by half, as was presented in the study of Sierociuk et al. [11].

The FC also allows the construction of new types of filters and new tools for signal analysis. Some applications of fractional order calculus in signal processing have been presented in the literature [5, 15, 16, 17]. For constant and variable fractional order systems, some generalisations of the Kalman filter have been presented in the studies of Sierociuk et al. [18], Sierociuk [19] and Sierociuk and Ziubinski [20]. When the uncorrelated noise (such as white noise) passes through a dynamical system, the dynamically correlated noise (coloured noise) is obtained. When the dynamics contain fractional order, a fractional noise is obtained. In the study of Wyss [21], an introduction is presented to fractional order noises (the noises obtained by applying uncorrelated white noise to fractional order dynamics). In the study of Sierociuk and Ziubinski [22], estimation schemes are presented for discrete fractional and integer order state-space systems with fractional order coloured noise. In the latter of these studies, owing to the additional information about noise dynamic used by the estimation algorithm proposed therein, better estimates of the state vector could be obtained.

The MEMS gyroscope sensors are quite complex dynamical systems encompassing non-linear dynamics, external disturbances and thermal noises, especially in high acceleration and high-velocity environments such as space crafts [34], hypersonic vehicles [35], missiles or munition [36]. The use of advanced control algorithms is necessary for a study of the application of nonlinearities in MEMS gyroscopes, especially those involving a hysteresis of quantisation levels.. For example, neural network approaches were used in the studies of Shao et al. [37, 38] and Shao and Shi [39]; additionally, in the study of Shao et al. [40], a fuzzy wavelet neural control was applied. In this paper, only static case noise analysis will be considered, and accordingly the research approach used in the present study would be a special case that omits the influence of sensor externally driven dynamics and focusses only on modelling thermal and other noises that can be measured when the sensor does not move. Especially, thermal-like noises can be efficiently modelled using fractional order models (fractional noises), which is the main motivation for applying fractional order estimation tools and undertaking investigation of static noise cases.

In practical fractional noise estimation, there is a problem with properly determining the parameters and order of the noise. In the study of Sierociuk and Macias [23], the triple estimation algorithm (TEA) for state vector, order and system parameters' estimation was proposed and described in detail. In the study of Macias et al. [24], a triple estimation algorithm was used to carry out identification of fractional order noise in MEMS accelerometer measurements. In the practical application of the triple estimation algorithm, it was found that the algorithm requires quite a high numerical power. The principal difficulty lies in the realisation of fractional order differences when the full number of samples is used.The problem of direct application of fractional order difference resulting in a high numerical power consumption is a well-recognised one in the literature, and there have been multiple studies proposing other methods characterised by much lower levels of numerical power consumption. In Stanisławski et al. [25,26], some types of approximations including Laguerre-based differences were proposed and analysed. The most typical method of approximation is to reduce the number of samples that are considered, which has an influence on derivative accuracy.

The paper's novelty lies in its modification of the triple estimation algorithm by introducing a limited number of samples during computation. Then, the new algorithm will be applied in the identification of the state, order and parameter of the gyroscope's x, y and z axes' noises as part of the Inertial Measuring Unit denoted by MPU9250. Moreover, the results will be compared to each other considering the various implementation lengths of FLTEA. So, in contrast with the approaches used in the studies of Sierociuk and Ziubinski [22] and Sierociuk and Macias [23], we present the triple estimation algorithm with finite length approximation and its application to the gyroscope's noises' estimation. The numerical power consumption of the proposed algorithm has also been analysed.

The remainder of the paper is organised as follows: Section 2 recalls the fractional noises' definition and particular fractional order definition with finite length approximation. In Section 3, the modified TEA with approximation length constraints is presented. Finally, Sections 4 and 5 show the possibilities of TEA in several numerical examples and during the estimation process of noise data originating from the gyroscope's built-in MEMS technology.

2. FRACTIONAL CALCULUS AND FRACTIONAL NOISES

In FC, the three most popular definitions of fractional constant order integral and derivative are used, namely, Grünwald–Letnikov, Riemann–Liouville and Caputo. These definitions possess different properties and may be applied in various areas of engineering.

In this paper, we use the Grünwald–Letnikov definition, which is usually used in discrete systems, as a base for fractional variable order (FVO) difference definition. Due to the applied nature of this work, we will use a discrete approximation of the Grünwald–Letnikov derivative with a finite (not going to 0) sampling time h. Hence, we have the constant order difference definition, which is formulated as the following:

$${}_0\Delta_k^{\alpha} x_k \equiv \sum_{j=0}^{k} \frac{1}{h^{\alpha}} (-1)^j \binom{\alpha}{j} x_{k-j},$$

where

$$\binom{\alpha}{j} \equiv \begin{cases} 1 & \text{for} \quad j = 0, \\ \dfrac{\alpha(\alpha-1)\ldots(\alpha-j+1)}{j!} & \text{for} \quad j > 0, \end{cases}$$

$\alpha \in \mathbb{R}$ is a fractional order and h is a time sampling.

Since the estimation of the order will be processed in time, this leads us to variable order operators. Four switching schemes and their equivalence to four definitions of FVO derivatives are presented in the literature [27, 28, 29]. In our paper, we will use the following FVO type of difference:

$${}_0^{\mathcal{A}}\Delta_k^{\alpha_k} x_k \equiv \sum_{j=0}^{k} \frac{(-1)^j}{h^{\alpha_k}} \binom{\alpha_k}{j} x_{k-j}$$

where $\alpha_k \in \mathbb{R}$ is FVO.

3. FRACTIONAL NOISE

The time-correlated (coloured) noises are the noises that contain a dynamic correlation between the noise samples. Such noises can be obtained when some noise (uncorrelated) is passed through dynamical systems. For example, electromagnetic field noise can induct some current in an electronic circuit, leading to some dynamically correlated noise in voltage because of some dynamic between current and voltage. When the order of the dynamics is an integer, we will have a dynamically correlated integer order noise, which the following relation can describe:

$$x_{k+1} = f x_k + \omega_k,$$

where x_k is a time-correlated noise, and ω_k is an uncorrelated noise, for example, white Gaussian noise.

When the dynamics of the system are fractional, for example, in temperature transport (for ideal beam temperature is half order integral of heat flux [11]), the uncorrelated heat flux noise can lead to fractional order dynamically correlated noise in temperature. The coloured fractional order noise is given as follows:

$$\begin{aligned} {}_0\Delta_{k+1}^{\alpha} x_{k+1} &= f x_k + \omega_k \\ x_{k+1} &= h^{\alpha}\, {}_0\Delta_{k+1}^{\alpha} x_{k+1} \\ &\quad - \sum_{j=1}^{k+1} (-1)^j \binom{\alpha}{j} x_{k-j+1}, \end{aligned}$$

where x_k is a fractional coloured noise, α is an order of the noise and ω_k is an uncorrelated noise.

The appearance of FVO noise can be observed in the case wherein the fractional order of the dynamical system is characterised by changes with time (e.g. when the structure of the heated medium changes over time [12]). Depending on the order-switching manner, different definitions can describe such dynamics. For example, for $\mathcal{A}$-type definition, we will have the following FVO noise dynamics:

$$\begin{aligned} {}_{0}^{\mathcal{A}}\Delta_k^{\alpha_{k+1}} x_{k+1} &= f x_k + \omega_k \\ x_{k+1} &= h^{\alpha_{k+1}}\, {}_{0}^{\mathcal{A}}\Delta_k^{\alpha_{k+1}} x_{k+1} \\ &-\sum_{j=1}^{k+1} (-1)^j \binom{\alpha_{k+1}}{j} x_{k-j+1}. \end{aligned}$$

Identification of the fractional noise in a real application is a complex process because we do not know the order and system parameters of the noise. We also do not have information about dynamically uncorrelated source noise. In the study of Ziubinski and Sierociuk [30], an identification algorithm for fractional noise was presented, but under the assumption that output noise is the only evident fractional order noise. In experimentally obtained noises, we would instead acquire a combination of dynamically correlated and uncorrelated noises, as the following expression describes:

$$y_k = x_k + \nu_k.$$

That is why, in this article, we use a triple estimation algorithm to identify parameters of fractional order noises.

4. FINITE LENGTH APPROXIMATION

The definition given by Eq. (2) leads to some implementation problems because of the very long tail of samples used for obtaining the value of difference. This resulted in the problems' characterisation by a high number of numerical operations as well as a high degree of memory consumption. In the literature there exist several algorithms using which to arrive at a more numerically efficient fractional difference approximation, among which the studies of Stanisławski et al. [25,26] can be mentioned as prime examples. In our paper, we will use the most popular method, which involves restricting the number of samples considered to some predefined value L, which will be known as the length of implementation. The finite length approximation will have the following form:

$${}_{0}^{\mathcal{A},L}\Delta_k^{\alpha_k} x_k = \sum_{j=0}^{L(k)} \frac{(-1)^j}{h^{\alpha_k}} \binom{\alpha_k}{j} x_{k-j},$$

where

$$L(k) = \begin{cases} k & \text{if } k < L \\ L & \text{if } k \geq L \end{cases}$$

Naturally, this approximation will have an influence in determining the accuracy of the obtained results, which will depend on used sampling time, time constants of the object and used input signals.

5. FINITE LENGTH TRIPLE ESTIMATION ALGORITHM

The triple estimation algorithm (TEA) allows estimating state vector, system parameters and fractional order simultaneously. In deploying this algorithm, our main idea is to separate the estimation processes used for states, parameters and orders.This separation allows a better adjustment of used filters, making it possible to obtain better estimation results. A detailed introduction of TEA was presented in the study of Sierociuk and Macias [23]. Here, a modification of that algorithm including finite length approximation of fractional order differences will be proposed (FLTEA).

The FLTEA will be defined for the following linear discrete fractional variable order state-space (DFVOSS) $\mathcal{A}$-type system [31] with finite length implementation of difference:

$$\begin{aligned} &{}_{0}^{\mathcal{A},L}\Delta_{k+1}^{\alpha_{k+1}} x_{k+1} = A x_k + B u_k + \omega_k, \\ &x_{k+1} = h^{\alpha_{k+1}}\, {}_{0}^{\mathcal{A},L}\Delta_{k+1}^{\alpha_{k+1}} x_{k+1} \\ &-\sum_{j=1}^{L(k)+1} (-1)^j \binom{\alpha_{k+1}}{j} x_{k-j+} \\ &y_k = C x_k + \nu_k, \end{aligned}$$

where $u_k \in \mathbb{R}^d$ is a system input; $y_k \in \mathbb{R}^p$ is a system output; $A \in \mathbb{R}^{N\times N}, B \in \mathbb{R}^{N\times d}$ and $C \in \mathbb{R}^{p\times N}$ are the state system, input and output matrices, respectively; $x_k \in \mathbb{R}^N$ is a state vector; and N is a number of state equations.

In the TEA process, the estimations of the FVO, state variables and parameters are divided into three estimation actions (filters). The first filter, $\mathrm{KF}x$, estimates the state variables' vector $\hat{x}_k$ based on estimates of order and system parameters corresponding to the other filters, namely KFo and $\mathrm{KF}w$, respectively. The second, $\mathrm{KF}w$, estimates the vector of system parameters $\hat{w}_k$ based on state variable and order estimates obtained in the remaining two filters $\mathrm{KF}x$ and KFo, respectively. The third filter, KFo, estimates the FVO with the knowledge of state variable and system parameters from filters $\mathrm{KF}x$ and $\mathrm{KF}w$, respectively. The scheme of the TEA is given in Fig. 1.

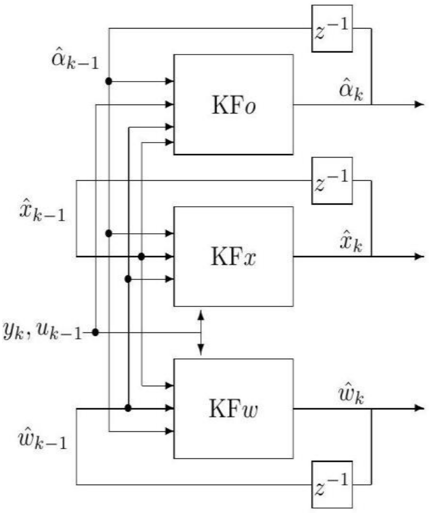


Fig. 1. The triple estimation algorithm scheme

5.1. Order estimation filter KFo

For the order estimation problem, the unscented fractional variable order Kalman filter with finite length differences implementation is used. The order changing dynamics is assumed to be a constant, given by

$$\alpha_{k+1} = \alpha_k + \omega_k^o,$$

where ω_k^o is a noise with variance given by matrix Q_k^o. The matrix Q_k^o represents our knowledge over how big fluctuations in time actually are vis-à-vis those that were assumed by us.

The KFo algorithm equations are given as follows:

$$
\begin{aligned}
\tilde{\alpha}_k &= \hat{\alpha}_{k-1},\\
\tilde{P}_k^o &= \hat{P}_{k-1}^o + Q_{k-1}^o,\\
\tilde{\boldsymbol{\alpha}}_k &= \left[\tilde{\alpha}_k \quad \tilde{\alpha}_k \pm \left(\sqrt{(L+\lambda)\tilde{P}_k^o}\right)_i\right],\\
{}^{\mathcal{A},L}\Delta^{\tilde{\boldsymbol{\alpha}}_{k,i}}\tilde{\chi}_{k,i}^o &= A(\hat{w}_{k-1})\hat{x}_{k-1} + Bu_{k-1},\\
\tilde{\chi}_{k,i}^o &= h^{\tilde{\boldsymbol{\alpha}}_{k,i}L}\Delta^{\tilde{\boldsymbol{\alpha}}_{k,i}}\tilde{\chi}_{k,i}^o\\
&\quad -\sum_{j=1}^{L(k)} (-1)^j \binom{\tilde{\boldsymbol{\alpha}}_{k,i}}{j}\hat{x}_{k-j},\\
\tilde{Y}_{k,i}^o &= C\tilde{\chi}_{k,i}^o,\\
\tilde{y}_k^o &= \sum_{i=0}^{2L} W^{(m)}\tilde{Y}_{k,i},\\
P_{y_k y_k}^o &= \sum_{i=1}^{2L} W_i^{(c)}\left[\tilde{\mathcal{Y}}_{i,k} - \tilde{y}_k\right]\left[\tilde{\mathcal{Y}}_{i,k} - \tilde{y}_k\right]^T\\
&\quad +R^o,\\
P_{\alpha_k y_k}^o &= \sum_{i=1}^{2L} W_i^{(c)}\left[\tilde{\alpha}_{i,k} - \tilde{\alpha}_k\right]\left[\tilde{Y}_{i,k} - \tilde{y}_k\right]^T,\\
\mathcal{K}_k^o &= P_{\alpha_k y_k}^o\left(P_{y_k}^o y_k\right)^{-1},\\
\hat{\alpha}_k &= \tilde{\alpha}_k + \mathcal{K}_k^o(y_k - \tilde{y}_k^o),\\
P_k^o &= \hat{P}_k^o - \mathcal{K}_k^o P_{y_k}^o y_k \mathcal{K}_k^o,\\
Q_k^o &= (1-\delta^o)Q_{k-1}^o\\
&\quad +\delta^o(\mathcal{K}_k^o)(y_k - \tilde{y}_k^o)(y_k - \tilde{y}_k^o)^T(\mathcal{K}_k^o)^T,
\end{aligned}
$$

where $\left(\sqrt{(L+\lambda)P_k}\right)_i$ is i-th column of matrix square root (e.g. Cholesky factorisation), L is a dimension of estimated state vector ($2L+1$ is a number of sigma points) and coefficients of unscented transformation W are given by

$$
\begin{aligned}
W_0^{(m)} &= \lambda/(L+\lambda),\\
W_0^{(c)} &= \lambda/(L+\lambda) + (1 - \mathfrak{A}^2 + \mathfrak{B}),\\
W_i^{(m)} &= W_i^{(c)} = 1/(2(L+\lambda)),
\end{aligned}
$$

where $\lambda = \mathfrak{A}^2(L+\kappa) - L, \mathfrak{A}$ is a coefficient describing the width of point expansion during the transformation (in the literature, this is obtained in the range $1 \le \mathfrak{A} \le 1e-4$, and is usually denoted as α, but in the present article, since we are using order α, this notation has been changed); κ is an additional scaling coefficient usually chosen as 3 – L; and $\mathfrak{B}$ is a coefficient that corresponds with our knowledge about type of noise, for Gaussian noise is chosen as $\mathfrak{B} = 2$ (in the literature this is usually denoted as β). The δ coefficient is a 'forgetting factor' according to the Robbins–Monro stochastic approximation scheme for estimating the innovations (see Haykin's study [32], p. 240). The initial values of matrix P_0^o represent our a priori knowledge about error in choosing the initial value of order α_0 (we assume that the initial value is different from the original).

5.2. State estimation filter KFx

As the KFx filter, the fractional variable order Kalman filter algorithm with finite length differences implementation is used, and has the following form:

$$
\begin{aligned}
{}_0^{\mathcal{A},L}\Delta_{k+1}^{\hat{\alpha}_k}\tilde{x}_{k+1} &= A(\hat{w}_{k-1})\hat{x}_k + Bu_k,\\
\tilde{x}_{k+1} &= h^{\hat{\alpha}_k}\,{}_0^{\mathcal{A},L}\Delta_{k+1}^{\hat{\alpha}_k}\tilde{x}_{k+1}\\
-\quad &\sum_{j=1}^{L(k)+1} (-1)^j\binom{\hat{\alpha}_k}{j}\hat{x}_{k+1-j},\\
\tilde{P}_k = \quad &\left(h^{\hat{\alpha}_k}A(\hat{w}_{k-1}) + \hat{\alpha}_k\right)P_{k-1}\\
&\left(h^{\hat{\alpha}_k}A(\hat{w}_{k-1}) + \hat{\alpha}_k\right)^T\\
&+Q_{k-1} + \sum_{j=2}^{L(k)}\binom{\hat{\alpha}_k}{j}P_{k-j}\binom{\hat{\alpha}_k}{j}^T,\\
K_k &= \tilde{P}_k C^T\left(C\tilde{P}_k C^T + R_k\right)^{-1},\\
\hat{x}_k &= \tilde{x}_k + K_k(y_k - C\tilde{x}_k),\\
P_k &= (I - K_k C)\tilde{P}_k,
\end{aligned}
$$

where initial conditions are

$$x_0 \in \mathbb{R}^N,\ P_0 = \mathrm{E}[(\tilde{x}_0 - x_0)(\tilde{x}_0 - x_0)^T],$$

and ν_k and ω_k are assumed to be independent with zero expected value.

5.3. Parameters estimation filter $\mathrm{KF}w$

For KFw filter, another unscented fractional variable order Kalman filter with finite implementation of differences is used. The dynamics of parameter-change are also assumed to be constant, given by

$$w_{k+1} = w_k + \omega_k^w,$$

where ω_k^w is a noise with variance given by matrix Q_k^w. The equations of the filter $\mathrm{KF}w$ are very similar to those for filter KFo, and the difference is only in the model replica part:

$$
\begin{aligned}
\tilde{w}_k &= \hat{w}_{k-1}\\
\tilde{P}_k^w &= \hat{P}_{k-1}^w + Q_{k-1}^w,\\
\tilde{\mathcal{W}}_k &= \left[\tilde{w}_k \quad \tilde{w}_k \pm \left(\sqrt{(L+\lambda)\tilde{P}_k^w}\right)_i\right],\\
{}^{\mathcal{A},L}\Delta^{\hat{\alpha}_{k-1}}\tilde{\chi}_{k,i}^w &= A\left(\tilde{\mathcal{W}}_{k,i}\right)\hat{x}_{k-1} + Bu_{k-1},\\
\tilde{\chi}_{k,i}^w &= h^{\hat{\alpha}_{k-1}}\,{}_0^{\mathcal{A},L}\Delta^{\hat{\alpha}_{k-1}}\tilde{\chi}_{k,i}^w\\
&\quad -\sum_{j=1}^{L(k)} (-1)^j\binom{\hat{\alpha}_{k-1}}{j}\hat{x}_{k-j}.
\end{aligned}
$$

Resuming the explanation for TEA, it consists of three subfilters requiring separate sets of parameters and initial conditions.Parameters of the order estimation filter KFo are denoted with the upper index ^{o} (e.g. $\tilde{P}_k^o, Q_{k-1}^o$), whereas parameters of $\mathrm{KF}w$ are denoted with the upper index ^{w} (e.g., $\tilde{P}_k^w, Q_{k-1}^w$) and parameters of KFx are rendered without an upper index.

6. IDENTIFICATION AND ANALYSIS OF FRACTIONAL VARIABLE ORDER SYSTEM PARAMETERS

Before we apply the finite length triple estimation algorithm to real plant data (noises' estimation of MEMS sensor), we will present the results of some numerical experiments for constant and

FVO systems. The one state variable discrete state-space system, used in numerical experiments, is given as follows:

$$ {}_{0}^{\mathcal{A},L}\Delta_{k+1}^{\alpha_{k+1}} x_{k+1} = f x_k + u_k + \omega_k \tag{1} $$

$$ x_{k+1} = h^{\alpha_{k+1}}\, {}_{0}^{\mathcal{A},L}\Delta_{k+1}^{\alpha_{k+1}} x_{k+1} - \sum_{j=1}^{L(k)+1} (-1)^j \binom{\alpha_{k+1}}{j} x_{k-j+1} \tag{2} $$

$$ y_k = x_k + v_k \tag{3} $$

6.1. Analysis of fractional constant and variable order system with input signal known

The number of samples in the numerical implementation of fractional order differences plays a significant role in determining the accuracy of the results. Therefore, in this section, we try to validate the influence of TEA with finite length approximation on final estimation results. We present the TEA with length constraints to make a numerical validation of its capabilities in the analysis of fractional constant and variable order systems. The estimation results were shown for various input signals and order functions. So, the problem in this section is formulated as follows: Estimate the state, order and parameter of the fractional order system with known input signal based on TEA with length constraints. The numerical tests were conducted in the Matlab/Simulink environment based on the Fractional Variable Order Derivative Toolkit [33], with a sample time given as $h = 0.001\ s$. The possibilities of the TEA designed with limited length were shown in the Examples (1)–(3). The aim of the examples is to assess the accuracy of the results derived from estimation of the state, order and parameter under various scenarios. The Example (1) deals with the fractional constant order system with input signal being the sawtooth wave, while the Examples (2) and (3) show the behaviour of TEA applied to fractional variable order systems for input signal being the sawtooth wave and Gaussian noise, respectively.

To compare the estimation results for various lengths of TEA implementation, the Examples (1)–(3) were run with the following, individually adjusted parameters:

– Noises parameters

$$ \mathrm{E}[\omega\omega^T] = 10^{-5}, $$
$$ \mathrm{E}[\nu\nu^T] = 10^{-3}, $$

– Parameters of KFx filter

$$ P_0 = [1], Q_0 = [10^{-5}], $$
$$ x_0 = [0], R = [10^{-3}], $$

– Parameters of KFo filter

$$ P_0^o = [0.05], Q_0^o = [0.005], $$
$$ \alpha_0 = [1], R^o = [10^{-3}], $$
$$ \mathfrak{A} = 1, \mathfrak{B} = 2, \delta^o = 0.5, $$

– Parameters of KFw filter

$$ P_0^w = [0.001], Q_0^w = [0.01], $$
$$ w_0 = [0], R^w = [10^{-3}], $$
$$ \mathfrak{A} = 1, \mathfrak{B} = 2, \delta^w = 0.5. $$

An identification of the fractional constant order system for various lengths of TEA is presented in Example 1. In this example, the implementation length of TEA is reduced to 50% of the original ones needed to cover a full range of consideration of fractional order system, with no losses in number of samples.

Example 1. Let us consider the DFVOSS $\mathcal{A}$-type system given by Eqs (1)–(3), where

$$ A = f = -0.3, B = 1, C = 1, \alpha_k = 0.6. $$

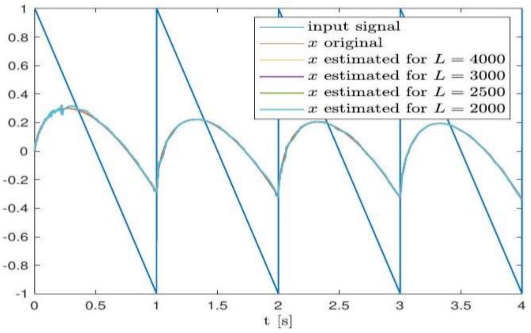


Fig. 2. Original and estimated state variable from Example 1 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 3{,}000,$ $\mathrm{L} = 2{,}500$ and $\mathrm{L} = 2{,}000)$ approximation lengths

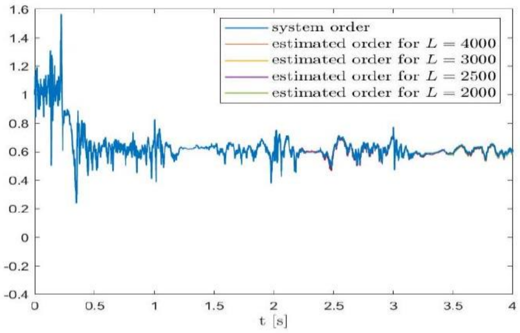


Fig. 3. Original and estimated order from Example 1 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 3{,}000,\ \mathrm{L} = 2{,}500$ and $\mathrm{L} = 2{,}000)$ approximation lengths

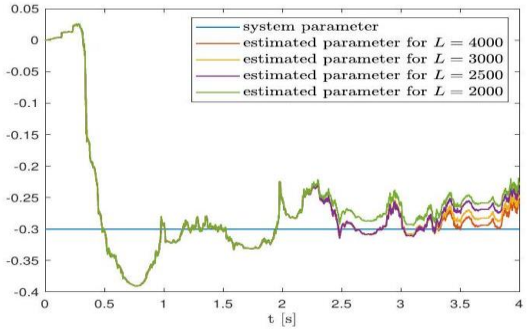


Fig. 4. Original and estimated parameter from Example 1 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 3{,}000, \mathrm{L} = 2{,}500$ and $\mathrm{L} = 2{,}000)$ approximation lengths

The state, order and parameter estimation of the system described in Example 1 are presented in Figs. 2, 3 and 4, respectively. In this example, for all implementation lengths of TEA, the state estimation converges to the original one with high accuracy. Moreover, despite the 50% length reduction of TEA, the order estimations overlap each other and this overlap does not influence the final order results. However, from Fig. 4, it is possible to note some amount of discrepancy between various tails' implementations of TEA. The plots of the estimated parameters are near the original one, but the differences between them are, in this case, noticeable. All these observations indicate that order estimation is much more robust in reducing samples of number than parameter estimation. We can infer that in pursuance of maintaining the high values of order estimation, even small differences in parameter estimation lead to high accuracy in state estimation.

We achieved satisfying estimation results even with a high percentage length reduction of TEA. Thus, it can also be interesting to show the time execution of TEA depending on its implementation length. The time consumption of the TEA algorithm, depending on its length implementation, is presented in Tab. 1. As we can see, the average execution time of TEA equals around $230\ s$ for a set of 4,000 combined samples of state, order and parameter estimation. On the other hand, for 2,000 samples, the same estimation process took around $184\ s$. The time-consuming tests were conducted on a PC with an Intel Core i7-5500U CPU, 2.4 GHz, 8 GB RAM and Matlab version 2021b 64 bit. The sum squared error (SSE) of state, order and parameter for various implementation lengths of TEA corresponding to Example 1 is given in Tab. 2. The SSE is calculated as a sum of squares' differences between original data and the corresponding estimates.

Tab. 1. Time execution of TEA depending on its implementation length (L)

Implementation length ($\boldsymbol{L}$)	Time execution ($\boldsymbol{s}$)
$L = 4{,}000$	230.75
$L = 3{,}000$	218.76
$L = 2{,}500$	195.30
$L = 2{,}000$	183.55

Tab. 2. The sum squared error of state, order and parameter corresponding to Example 1; for various lengths of TEA

Length	State	Parameter	Order
L=4,000	0.2179	37.8467	62.7249
L=3,000	0.2175	38.1044	62.7476
L=2,500	0.2180	38.6321	62.8327
L=2,000	0.2170	39.8187	62.9066

An estimation of the fractional variable order system is presented in Example 2. It is an extension of Example 1 while replacing the constant value of the order with a time-varying function. In this example, we show the results of applying the TEA with its full implementation range and reduced to 50%.

Example 2. Let us consider the DFVOSS $\mathcal{A}$-type system given by Eqs (1)–(3), where

$$A = f = -0.3, B = 1, C = 1,$$
$$\alpha_k = 0.2 + 0.1kh \text{ for } k = 1,2,3,\ldots$$

In this case, we can see that despite reduced estimation length in TEA, the results tend to be the original values of the desired system. The state, order and parameter estimation are shown in Figs. 5, 6 and 7, respectively. The state estimation is reflected with high accuracy for constraint and unconstraint TEA implementation. The original and both estimated states overlap with high accuracy. Compared to Example 1, the order and parameter estimations achieved the original values starting with $1\ s$, and maintained these during the whole simulation process. In Figs. 6 and 7, we can observe the only slight difference between the estimation of order and parameter for TEA considering 4,000 and 2,000 historical samples. These insignificant changes and time-consumption reduction show the advantage of the proposed TEA with implementation length constraints. The sum squared error (SSE) of state, order and parameter for various implementation lengths of TEA corresponding to Example 2 is given in Tab. 3.

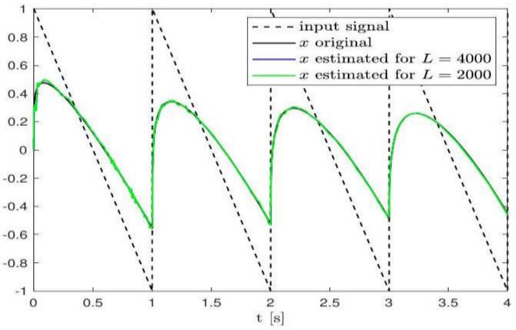


Fig. 5. Original and estimated state variable from Example 2 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 2{,}000)$ approximation lengths

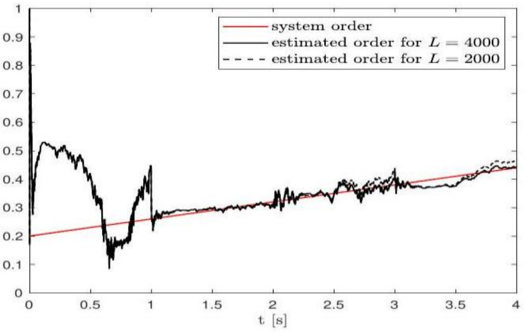


Fig. 6. Original and estimated order from Example 2 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 2{,}000)$ approximation lengths

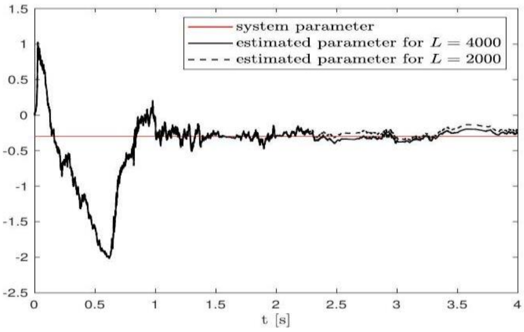


Fig. 7. Original and estimated parameter from Example 2 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 2{,}000)$ approximation lengths

Tab. 3. The sum squared error of state, order and parameter corresponding to Example 2; for various length of TEA

Length	State	Parameter	Order
L=4,000	0.1635	34.8872	1663.2
L=2,000	0.1634	35.0478	1663.1

Usage of the TEA in the estimation process for fractional variable order system with input signal being the Gaussian noise is shown in Example 3. The noise parameters are 0 mean value and 0.1 variance.

Example 3. Let us consider the DFVOSS $\mathcal{A}$-type system given by Eqs (1)–(3), where

$$A = f = -0.3, B = 1, C = 1, u_k \sim \mathcal{N}(0,0.1)$$
$$\alpha_k = 0.2 + 0.1kh \text{ for } k = 1,2,3, \dots .$$

The estimation results achieved in Example 3 are shown in Figs. 8–10. In contrast with both of the previous examples, there is no difference between state estimation for TEA with length $L = 4{,}000$ vis-à-vis that with length $L = 2{,}000$. Applying the noisy input signal and time-varying order to the TEA lead to the same results for both algorithms' lengths. It is an interesting issue that decreasing the number of samples in TEA reduces its computation time and does not influence results. Fig. 9 shows that the order estimations overlap the original one up to 2.5 s., and starting with this time, the small difference occurs. A similar situation can be observed on the plot with parameter estimation (see Fig. 10), where the curves are near the original ones. The obtained results demonstrate the high precision of TEA for the reduced number of samples applied to fractional variable order systems with Gaussian noise as an input signal. The sum squared error (SSE) of state, order and parameter for various implementation lengths of TEA corresponding to Example 3 is given in Tab. 4.

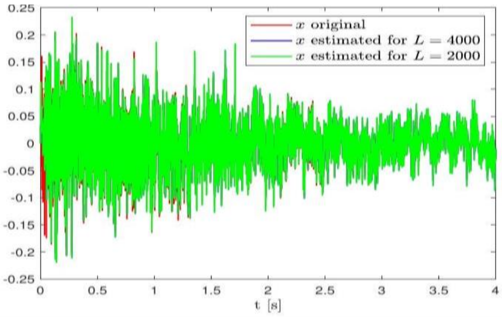


Fig. 8. Original and estimated state variable from Example 3 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 2{,}000)$ approximation lengths

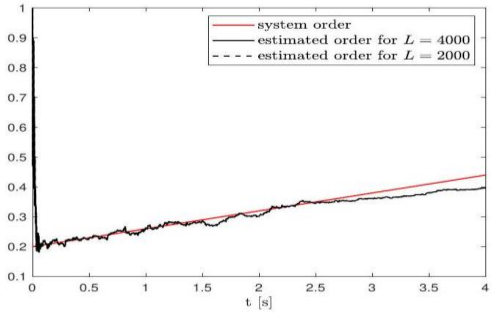


Fig. 9. Original and estimated order from Example 3 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 2{,}000)$ approximation lengths

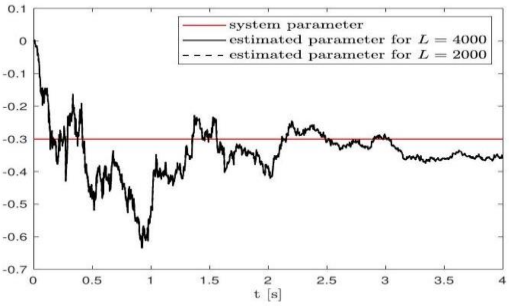


Fig. 10. Original and estimated parameter from Example 3 given for full $(\mathrm{L} = 4{,}000)$ and finite $(\mathrm{L} = 2{,}000)$ approximation lengths

Tab. 4. The sum squared error of state, order and parameter corresponding to Example 3; for various length of TEA

Length	State	Parameter	Order
L=4,000	0.4711	738.2196	1875.8
L=2,000	0.4831	744.9644	1902.9

6.2. Identification without input signal knowledge

When an input signal is not measured, the identification process can differ from desired values or order and system parameter. This can be explained by the fact that in practice, the system noise can have some unknown dynamical correlation of some order and parameter. Let us assume the fractional noise system equation in the following form:

$$\Delta_1^{\alpha} x_{k+1} = f_1 x_k + \omega'_k$$

where ω'_k is a system noise also containing the fractional order dynamical correlation described by the following relation:

$$\Delta_2^{\alpha} \omega'_{k+1} = f_2 \omega'_k + \omega_k$$

where ω_k is assumed to be noise without dynamical correlation.

By combining both equations, we obtain

$$\Delta_1^{\alpha} x_{k+1} - \frac{1}{f_2} \Delta_2^{\alpha} \omega'_{k+1} = f_1 x_k - \frac{1}{f_2} \omega_k .$$

As we can see, this dynamical correlation can have a direct effect on estimated order and system parameter in the estimation process, which can make the obtained estimation results different from those assumed in numerical models, because they consider also the dynamical correlation of the source noise. However, this will not pose a problem in estimation of real plant noise because the aim of estimation is to find the most appropriate model with the assumption that the source noise is without dynamical correlation.

7. IDENTIFICATION AND ANALYSIS OF MEMS GYROSCOPE'S NOISES

Motivated by the study of Macias et al. [24], where it was shown that the accelerometer's noise of MPU9250 contains the fractional order behaviour, we decided to make a noise analysis of its three-axes gyroscope. We utilise the TEA with various approximation lengths during the estimation process. The reduced implementation length of TEA by up to 50% decreases time execution and has an insignificant impact on final estimation results.

This section provides the experimental results pertaining to the modelling of noises for the three-axes gyroscope that forms part of the SparkFun MPU9250 Inertial Measurement Unit (IMU) built-in MEMS technology. We assume unknown input signal knowledge during the entirety of the estimation process.

The MPU9250 unit is a nine degree of freedom MEMS with three accelerometer's axes, three gyroscope's axes and three magnetometer's axes. It breakout board runs on 3.3 VDC and contains $\mathrm{I}^2\mathrm{C}$ and SPI communication protocols.

7.1. Experimental setup

The gyroscope's data were collected based on an experimental setup that is presented in Fig. 11. The Inertial Measurement Unit (IMU), denoted as MPU9250 in a stationary position, was connected to an Arduino Due development board using the $\mathrm{I}^2\mathrm{C}$ protocol. Its operating range was configured to $+/-2{,}000\ \mathrm{dps}$, and the step time for data gathering equals $0.01\ \mathrm{s}$. Then, the measurement noises from the three axes of the gyroscope were transferred to the Matlab/Simulink environment and analysed using the triple estimation algorithm. The estimations of state, order and parameter of noises corresponding to the gyroscope axes x, y and z were conducted based on TEA with full implementation length ($L = 3{,}000$) and constrained to 50% by setting L to 1,500 considering samples. The TEA was separately applied to noise estimations carried out for the x, y and z axes under, respectively, the following $\mathrm{KF}x$, $\mathrm{KF}o$ and $\mathrm{KF}w$ parameters:

– Parameters of KFx filter:

$$P_0 = [0.01], Q_0 = [0.15],$$
$$x_0 = [0], R = \begin{cases} 0.0235 & \text{for } x\text{-axis noise} \\ 0.0178 & \text{for } y\text{-axis noise} \\ 0.0191 & \text{for } z\text{-axis noise} \end{cases}$$

– Parameters of KFo filter:

$$P_0^o = [0.01], Q_0^o = [0.1],$$
$$\alpha_0 = [1], \mathfrak{A} = 1, \mathfrak{B} = 2, \delta^o = 0.5,$$
$$R^o = \begin{cases} 0.0235 & \text{for } x\text{-axis noise} \\ 0.0178 & \text{for } y\text{-axis noise} \\ 0.0191 & \text{for } z\text{-axis noise} \end{cases}$$

– Parameters of KFw filter:

$$P_0^w = [0.01], Q_0^w = [0.1],$$
$$w_0 = [0], \mathfrak{A} = 1, \mathfrak{B} = 2, \delta^w = 0.5,$$
$$R^w = \begin{cases} 0.0235 & \text{for } x\text{-axis noise} \\ 0.0178 & \text{for } y\text{-axis noise} \\ 0.0191 & \text{for } z\text{-axis noise} \end{cases}$$

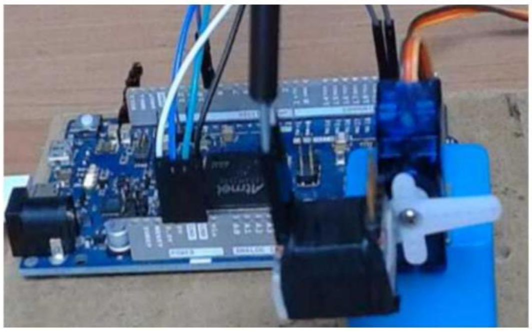


Fig. 11. The real view of experimental setup with an Arduino Due development board and MPU9252 IMU mounted on the shaft of a servo motor in a fixed position

7.2. Experimental results

The estimation of x-axis noise is presented in Fig. 12. It is worth noting that the results overlap with no differences for both lengths of TEA. The state, order and parameter plots are the same despite significant TEA length reduction. Moreover, Fig. 13 confirms, in this case, the fractional order noise, which tends to 0.3. The parameter estimation stabilises its value around -1.6 (see Fig. 14). Collectively, all these observations suggest that the estimation results of x-axis noise are characterised by a high degree of precision.

Analysis of y-axis noise is presented in Figs. 15-17. As shown in Fig.15, the state estimation is well-reflected for both lengths of TEA. In this case, the noise also exhibits the fractional order dynamic, and its order value goes to 0.3 rapidly (see Fig. 16). This order value was maintained until the final estimation process. Additionally, it can be noted in Fig. 17 that the estimated parameter tends to the value -1.5 and follows it with minor fluctuations.

The identification of z-axis noise is shown in Fig. 18. As in previous cases, the constraints of TEA implementation length do not influence estimation results. The order estimation shown in Fig. 19 reveals its fractional behaviour. We can see that this time also, the order value achieved the central value very quickly and stabilises itself around the value 0.3. The parameter estimation of y-axis noise presented in Fig. 20 attains the value -2 in approximately 5 s.

To summarise, using TEA during the noise estimation process allows us to obtain high-accuracy noise models for the x, y and z axes of the gyroscope that forms part of the MPU9250 sensor. Moreover, all the investigated data highlight its fractional order dynamic correlation and robustness for the length constraint of TEA up to 50%. This fact can significantly reduce the time consumption for TEA execution in the absence of estimation of precision losses. The sum squared error (SSE) of state, order and parameter for various implementation lengths of TEA corresponding to experimental data are given in Tab. 5. In this case, the SSE is calculated directly between their estimates for lengths L = 3,000 and L = 1,500.

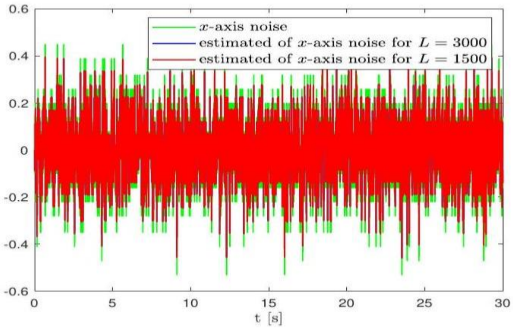


Fig. 12. Original and estimation of x-axis noise given for full ($\mathrm{L} = 3{,}000$) and finite ($\mathrm{L} = 1{,}500$) approximation lengths

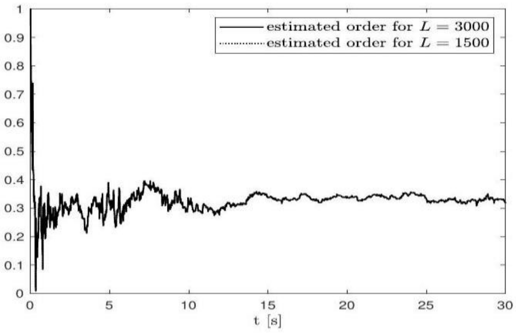


Fig. 13. Order estimation for x-axis noise given for full ($\mathrm{L} = 3{,}000$) and finite ($\mathrm{L} = 1{,}500$) approximation lengths

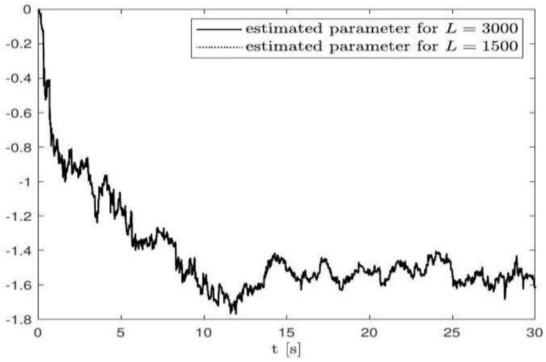


Fig. 14. Parameter estimation for x-axis noise given for full $(L = 3,000)$ and finite $(L = 1,500)$ approximation lengths

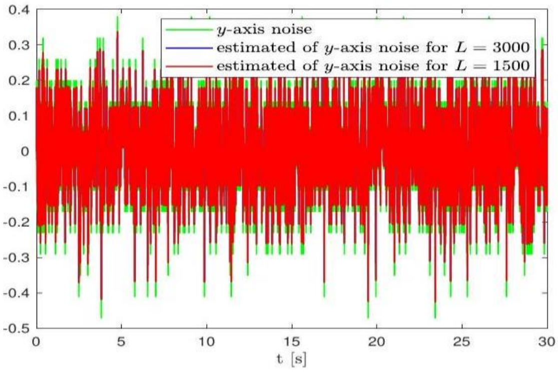


Fig. 15. Original and estimation of y-axis noise given for full $(L = 3,000)$ and finite $(L = 1,500)$ approximation lengths

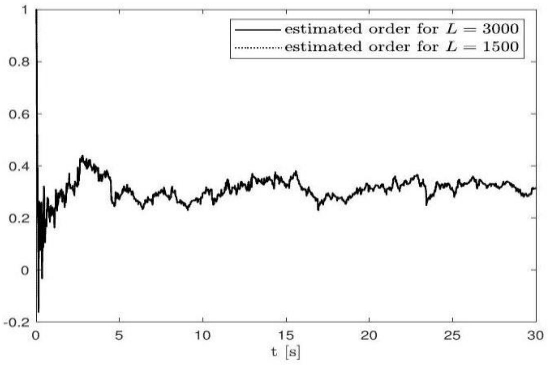


Fig. 16. Order estimation for y-axis noise given for full $(L = 3,000)$ and finite $(L = 1,500)$ approximation lengths

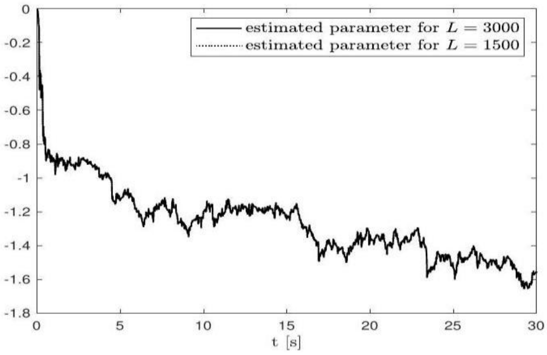


Fig. 17. Parameter estimation for y-axis noise given for full $(L = 3,000)$ and finite $(L = 1,500)$ approximation lengths

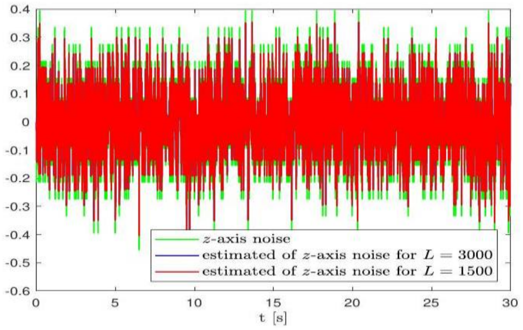


Fig. 18. Original and estimation of z-axis noise given for full $(L = 3,000)$ and finite $(L = 1,500)$ approximation lengths

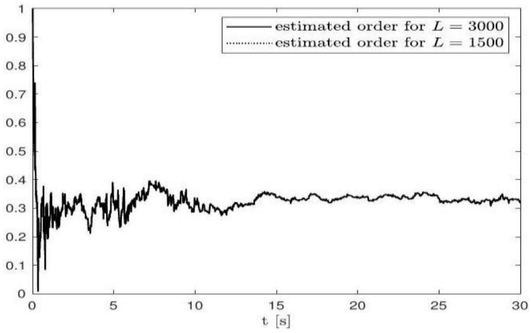


Fig. 19. Order estimation for z-axis noise given for full $(L = 3,000)$ and finite $(L = 1,500)$ approximation lengths

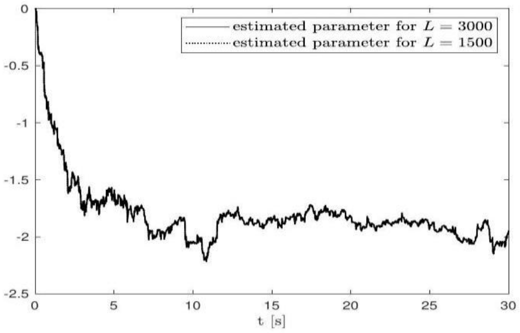


Fig 20. Parameter estimation for z-axis noise given for full $(L = 3,000)$ and finite $(L = 1,500)$ approximation lengths

Tab. 5. The sum squared error of state, order and parameter corresponding to experimental noise data

Noise	State	Parameter	Order
x-axis	$1.79 \cdot 10^{-8}$	$2.16 \cdot 10^{-7}$	$5.09 \cdot 10^{-8}$
y-axis	$3.29 \cdot 10^{-8}$	$2.52 \cdot 10^{-5}$	$1.08 \cdot 10^{-6}$
z-axis	$6.52 \cdot 10^{-9}$	$1.62 \cdot 10^{-6}$	$7.36 \cdot 10^{-8}$

8. CONCLUSIONS

The paper presents the experimental and numerical results of applying the triple estimation algorithm with approximation length constraints. The possibilities of such algorithms have been revealed during the state, order and parameter estimation of fractional constant and variable order systems in several numerical examples. In sets of numerical examples, the implementation length of TEA was reduced to 50% of the number of samples needed to cover the whole computing range with no implementation losses. The estimated state and order plots for the fractional constant order system tend to the original values. Only a tiny discrepancy occurs during parameter estimation for selective implementation length reduction. However, it has not influenced the final state estimation results. When identifying the fractional variable order system with different lengths of TEA, we can notice only a slight difference between the order estimations. Despite a length reduction to 50% of the original size, the estimated state, order and parameter curves overlap and are near the original ones. The numerical tests confirm the high accuracy of the achieved estimation results even for a finite length of the triple estimation algorithm. Combining numerical results leads to the conclusion that a reduction of samples by up to 50% does not significantly affect the state estimation results of considering fractional order systems and substantially decreases its time duration.

The triple estimation algorithm was also successfully used for fractional noise estimation of the gyroscope, part of a popular Inertial Measurement Unit known as MPU9250. The noise analysis was conducted for x, y and z axes of the gyroscope. The state, order and parameter estimation results for each axis of the gyroscope are similar. The conducted experiments show that the order of noises for the three gyroscopes' axes equals approximately 0.3, and the estimated parameters achieve a value of around –1.8. At this time, the approximation length also does not influence the final estimation results. Moreover, the experiments show the fractional order correlation dynamics of the investigated noises.

The plethora of numerical examples and experiments allow us to ascertain that the triple estimation algorithm with finite length approximation becomes a convenient tool during the analysis, identification and estimation of fractional variable order systems.

REFERENCES

1. Samko SG, Kilbas AA, Maritchev OI. Fractional Integrals and Derivative. Theory and Applications. Gordon & Breach Sci. Publishers; 1987.
2. Miller KS, Ross B. An Introduction to the Fractional Calculus and Fractional Differenctial Equations. New York, USA: John Wiley & Sons Inc.; 1993.
3. Monje CA, Chen Y, Vinagre BM, Xue D, Fe-liu V. Fractional-order Systems and Controls.London. UK: Springer; 2010.
4. Podlubny I. Fractional Differential Equations. Academic Press; 1999.
5. Magin R, Ortigueira MD, Podlubny I, Trujillo J. On the fractional signals and systems. Signal Processing. 2011;91(3):350 371. Advances in Fractional Signals and Systems.
6. Kilbas AA, Srivastava HM, Trujillo JJ. Theory and Applications of Fractional Differential Equations, Volume 204 (North-Holland Mathematics Studies). USA: Elsevier Science Inc. 2006.
7. West BJ. Fractional Calculus and the Future of Science. Entropy. 2021;23(12). https://www.mdpi.com/1099-4300/23/12/1566
8. Anastassiou GA. Generalized Fractional Calculus. Springer. Cham. 2021.
9. Yang XJ. General Fractional Derivatives: Theory, Methods and Applications. Chapman and Hall/CRC; 2019.
10. Tarasov VE. Generalized Memory: Fractional Calculus Approach. Fractal and Fractional. 2018;2(4).https://www.mdpi.com/2504-3110/2/4/23.
11. Sierociuk D, Dzielinski A, Sarwas G, Petras I, Podlubny I, Skovranek T. Modelling heat transfer in heterogeneous media using fractional calculus. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2013;371(1990).
12. Sakrajda P, Sierociuk D. Modeling Heat Transfer Process in Grid-Holes Structure Changed in Time Using Fractional Variable Order Calculus. In: Babiarz A, Czornik A, Klamka J, Niezabitowski M, editors. Theory and Applications of Non-integer Order Systems. Cham: Springer International Publishing. 2017: 297-306.
13. Reyes-Melo ME, Martinez-Vega JJ, Guerrero-Salazar CA, Ortiz-Mendez U. Application of fractional calculus to modelling of relaxation phenomena of organic dielectric materials. In: Proceedings of International Conference on Solid Dielectrics. Toulouse. France: 2004.
14. Dzielinski A, Sierociuk D, Sarwas G. Some applications of fractional order calculus. Bulletin of The Polish Academy of Sciences – Technical Sciences. 2010;58(4):583-92.
15. Ortigueira MD, Val ́erio D. Fractional Signals and Systems. De Gruyter; 2020. https://doi.org/10.1515/9783110624588.
16. Sheng H, Chen Y, Qiu T. Fractional Processes and Fractional-Order Signal Processing. Springer. London; 2012.
17. Muresan CI, Birs IR, Dulf EH, Copot D, Miclea L. A Review of Recent Advances in Fractional-Order Sensing and Filtering Techniques. Sensors. 2021;21(17). Available from: https://www.mdpi.com/1424-8220/21/17/5920.
18. Sierociuk D, Tejado I, Vinagre BM. Improved fractional Kalman filter and its application to estimation over lossy networks. Signal Processing. 2011 MAR;91(3, SI):542-52.
19. Sierociuk D. Fractional Kalman Filter algorithms for correlated system and measurement noises. Control and Cybernetics. 2013;42(2):471-90.
20. Sierociuk D, Ziubinski P. Variable order fractional Kalman filters for estimation over lossy network. Lecture Notes in Electrical Engineering. 2015;320:285-94.
21. Wyss W. Fractional noise. Foundations of Physics Letters. 1991;4: 235–246.
22. Sierociuk D, Ziubinski P. Fractional order estimation schemes for fractional and integer order systems with constant and variable fractional order colored noise. Circuits, Systems, and Signal Processing. 2014;33(12):3861-82. DOI:10.1007/s00034-014-9835-0.
23. Sierociuk D, Macias M. Triple Estimation of Fractional Variable Order, Parameters, and State Variables Based on the Unscented Fractional Order Kalman Filter. Sensors. 2021;21(23). Available from: https://www.mdpi.com/1424-8220/21/23/8159.
24. Macias M, Sierociuk D, Malesza W. MEMS Accelerometer Noises Analysis Based on Triple Estimation Fractional Order Algorithm. Sensors. 2022;22(2).
25. Stanislawski R, Latawiec KJ, Lukaniszyn M, Galek M. Time-domain approximations to the Grunwald-Letnikov difference with application to modeling of fractional-order state space systems. In: 2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR). 2015: 579-84.
26. Stanislawski R, Hunek WP, Latawiec KJ. Finite approximations of a discrete-time fractional derivative. In: 2011 16th International Conference on Methods Models in Automation Robotics. 2011:142-5.
27. Sierociuk D, Malesza W, Macias M. Derivation, interpretation, and analog modelling of fractional variable order derivative definition. Applied Mathematical Modelling. 2015;39(13):3876-88. http://dx.doi.org/10.1016/j.apm.2014.12.009.
28. Sierociuk D, Malesza W, Macias M. On the Recursive Fractional Variable-Order Derivative: Equivalent Switching Strategy, Duality, and Analog Modeling. Circuits, Systems, and Signal Processing. 2015;34(4):1077-113.
29. Macias M, Sierociuk D. An alternative recursive fractional variable-order derivative definition and its analog validation. In: Proceedings of International Conference on Fractional Differentiation and its Applications. Catania, Itally; 2014.

30. Ziubinski P, Sierociuk D. Fractional order noise identification with application to temperature sensor data. In: Circuits and Systems (ISCAS), 2015 IEEE International Symposium on; 2015: 2333-6.
31. Sierociuk D, Malesza W. Fractional variable order discrete-time systems, their solutions and properties. International Journal of Systems Science. 2017;48(12):3098-105.
32. Haykin S. Kalman Filtering and Neural Networks. John Wiley & Sons, Inc.: New York, USA; 2001.
33. Sierociuk D. Fractional Variable Order Derivative Simulink Toolkit; 2012. http://www.mathworks.com/matlabcentral/fileexchange/38801-fractional-variable-order-derivative-simulink-toolkit
34. Wu B, Cao X. Robust Attitude Tracking Control for Spacecraft with Quantized Torques. IEEE Transactions on Aerospace and Electronic Systems. 2017;11:1-1.
35. Wang Y, Yang X, Yan H. Reliable Fuzzy Tracking Control of Near-Space Hypersonic Vehicle Using Aperiodic Measurement Information. IEEE Transactions on Industrial Electronics. 2019;66(12):9439-47.
36. Sulochana S, Hablani H. Precision Munition Guidance and Moving-Target Estimation. Journal of Guidance, Control, and Dynamics. 2016;39:1-12.
37. Xingling S, Shi Y, Wendong Z. Input-and-Measurement Event-Triggered Output Feedback Chattering Reduction Control for MEMS Gyroscopes. IEEE Transactions on Systems, Man,and Cybernetics: Systems. 2021.
38. Xingling S, Shi Y, Wendong Z, Cao H, Jiawei L. Neurodynamic Approximation-Based Quantized Control with Improved Transient Performances for MEMS Gyroscopes: Theory and Experimental Results. IEEE Transactions on Industrial Electronics. 2020.
39. Xingling S, Shi Y. Neural-Network-Based Constrained Output-Feedback Control for MEMS Gyroscopes Considering Scarce Transmission Bandwidth. IEEE Transactions on Cybernetics. 2022.
40. Xingling S, Si H, Zhang W. Fuzzy wavelet neural control with improved prescribed performance for MEMS gyroscope subject to input quantization. Fuzzy Sets and Systems. 2020;411.

Michał Macias: https://orcid.org/0000-0001-7123-7708

Dominik Sierociuk: https://orcid.org/0000-0002-3700-3665

DOI 10.2478/ama-2023-0026

AN ITAE OPTIMAL SLIDING MODE CONTROLLER FOR SYSTEMS WITH CONTROL SIGNAL AND VELOCITY LIMITATIONS

Mateusz PIETRALA*, Piotr LEŚNIEWSKI*, Andrzej BARTOSZEWICZ*

*Institute of Automatic Control, Łódź University of Technology,
ul. Stefanowskiego 18., 90-537 Łódź, Poland

mateusz.pietrala@gmail.com, piotr.lesniewski@p.lodz.pl, andrzej.bartoszewicz@p.lodz.pl

received 2 November 2022, revised 5 January 2023, accepted 5 January 2023

Abstract: In this paper, a sliding mode controller, which can be applied for second-order systems, is designed. Robustness to external disturbances, finite regulation time and a good system's behaviour are required for a sliding mode controller. In order to achieve the first two of these three goals, a non-linear, time-varying switching curve is introduced. The representative point (state vector) belongs to this line from the very beginning of the control process, which results in elimination of the reaching phase. The stable sliding motion along the switching curve is provided. Natural limitations such as control signal and system's velocity constraints will be taken into account. In order to satisfy them, the sliding line parameters will be properly selected. However, a good dynamical behaviour of the system has to be provided. In order to achieve that, the integral time absolute error (ITAE) quality index will be introduced and minimised. The simulation example will verify theoretical considerations.

Key words: sliding mode control, optimal control, state constraints

1. INTRODUCTION

The sliding mode approach belongs to the class of variable structure control methodologies. The main idea is that the trajectory of the system state is constrained to a sub-space of the state space. This simplifies the dynamical behaviour of the system. Moreover, it allows us to obtain exceptional robustness with respect to external perturbations [6] and requires small computational effort. These advantages made sliding mode control profuse in the field of mechanical systems, electric drives and similar industrial applications. The fundamental concepts of this method were introduced in the previous century [20]; however, it remains a lively field of inquiry both from the practical [21] as well as theoretical [9] point of view.

Due to the discontinuous control signal, sliding mode control is especially useful in electric devices and power electronic control; as in these applications, the control signal (semiconductor switch state) is by its nature not continuous. The paper [7] presents a super-twisting sliding mode controller for speed control of permanent magnet synchronous machine (PMSM). The parameters of the controller are tuned on-line using a neural network to adapt to the varying, unknown disturbance values. The computer simulations as well as experimental tests confirm marked improvements over traditional sliding mode control as well as a PID (Proportional Integral Derivative) controller. The authors of [23] develop a sliding mode controller for a similar problem of linear PMSM control. In order to minimise chattering, it is enhanced by a neural network, which compensates the unknown impact of perturbations. To reduce chattering even further, the sign function in the controller is replaced by a time-varying saturation function, which determines the size of the boundary region to guarantee small error and minimum chattering. The control methodology is verified on a laboratory stand and put alongside a sliding mode controller without the fuzzy logic extension. The results plainly demonstrate an improvement in control precision and oscillation reduction. In contemporary applications, it becomes increasingly common to control the plant through a network system, instead of a direct connection between each sensor/actuator and the controller. Such an approach can lower the costs and increase the modularity of the system. Unfortunately, it also presents some new difficulties, namely packet losses and transmission delays. In the work [8], an extensive review of modern sliding mode control algorithms that utilise networked control is presented.

In [10], a new sliding mode control approach was used in uninterruptible power supply (UPS) systems. The slope of the sliding line depends on the output voltage error via fuzzy logic. A simulation example demonstrated that the total harmonic distortion is lower than that in the strategy with the fixed sliding line. Furthermore, in [11], a laboratory experiment for the proposed strategy was performed. Terminal sliding mode control for two classes of non-linear systems was implemented in [16]. The first one comprises systems with the first- and second-order derivatives, and the second one included only first-order derivatives. The controller provided a finite reaching time and the stable sliding motion. Computer simulations included RLC (Resistor, Inductor, Capacitor) and RC (Resistor, Capacitor) plants with a non-linear resistor and capacitor. Terminal sliding mode control was also used for multi input multi output (MIMO) systems in [26]. The non-linear switching curve was selected in order to achieve a finite time convergence to the demand state.

The sliding mode approach is regularly applied not only in controllers but also in observers. One of its advantages is forcing the estimation error to zero in finite time [1, 17], in contrast to asymptotic convergence in “traditional” observers. In the paper [5], the state of charge (SoC) estimation in a vanadium redox battery is considered. The authors begin by deriving a concentration model and tuning it using particle swarm optimisation. Next, it is transformed to the canonical control form, and a sliding mode observer is proposed. Since SoC in an actual battery is hard to measure directly, the battery voltage discrepancy between the observer and the real plant is used to assess the observation precision. What is more is the colour of the electrolytes in the full charge and discharge state verifies the observer performance. In the paper [4], the stator current and rotor flux linkage in a bearingless induction machine is estimated via a sliding mode observer. Such motors have important advantages, resulting from replacing mechanical bearings by additional windings in the stator. With proper control of the current in the windings, radial forces can be developed, which make the rotor “levitate” inside of the stator. This allows the motor to attain significantly higher speeds and lowers friction force. A saturation function was used to reduce the chattering of the observer. The results of computer simulations show that the presented observer has faster convergence and lesser steady state error than the model reference adaptive system (MRAS) speed identification, which is typically used in its place. The sliding mode paradigm was also used in [12] both for observer and controller design for a gimbal control system for use in satellite orientation control. The performance is verified in simulations and in tests on a laboratory stand.

In the article [2], a sliding mode controller was implemented on a robotic manipulator. The reaching phase was eliminated by selecting the switching curve parameters in such a way that the initial state belongs to it. In the first example, the proposed strategy was applied to control the two-joint manipulator. The stable sliding motion was provided, and the convergence to the demand state is faster than that in previous methods. Moreover, the robustness to the external disturbances and modelling uncertainties was achieved. In the second example, the five-joint, bipedal manipulator was considered. A similar approach was implemented in [3], where the example with five-joint, bipedal manipulator was studied in greater detail. Moreover, the strategy was compared with a classical robot controller.

The control of unmanned aerial vehicles (UAVs) is recently very active and often utilises the sliding mode methodology. In the paper [24], quadrotor control is analysed. A reaching law comprising two hyperbolic functions is introduced. It enables on the one hand rapid convergence at a large distance from the switching hyperplane, while on the other hand it limits the risk of exciting oscillations in its vicinity. Then, a “system dynamics estimator” is derived in order to assess the wind impact. In general, using disturbance observers allows us to reduce chattering since it enables a reduction of the discontinuous portion of the control signal. The estimator design is grounded on the assumption of bounded variation of wind speed. Since in reality, the wind can come in gusts, it is not clear if this premise is realistic. Although the approach is tested on a real UAV, the tests are performed inside a laboratory, with constant wind speed simulated by a fan. Thus, it is unclear, how the proposed controller would perform in real conditions. A similar problem of UAV formation control was tackled in [19], where an adaptive non-singular terminal super-twisting sliding mode controller was proposed. One of the uses of UAV formation is inspection works, such as monitoring electric cables or solar panels. In the proposed approach, a sliding mode trajectory tracking controller for the formation leader is derived. Based on this trajectory, a formation controller produces trajectories for all of the following UAVs. Then, the same controller which was used for the leader is implemented in every follower. The robustness to wind perturbations is verified by computer simulations. The formation control problem was also analysed in [15] where a non-singular terminal sliding mode controller was designed. The major difference is that in [15] the authors assume that only local distance information is available, as not all followers are able to directly communicate with the leader. Furthermore, a collision avoidance mechanism was introduced, which is based on an artificial potential field. The concept was analysed in theory, as well as tested in computer simulations.

The topic of the paper [22] is control of a bridge crane. The authors propose a time-varying sliding mode controller. Similarly, as in this work, the sliding hyperplane initially passes through the starting state, which ensures robustness from the beginning of the control process. In the selected model, the movement of all the masses (the load, the hook and the trolley) are taken into account, while the main goal is to move the suspended load rapidly but not induce large oscillations in the system. The authors compare their approach with two methods with constant sliding hyperplanes in computer simulations. The advantages, namely improved robustness and smaller oscillations, are evident. The task of oscillation minimisation of moving masses was also considered in [25] where a benchmark problem of balancing a ball on a beam was analysed. The reaching phase was removed by using an integral sliding mode controller, while the sign function was smoothed out to reduce chattering. The laboratory tests confirm the impressive position control precision. In the paper [18], a second-order sliding mode controller for frequency control in a multi-area power system was derived. The unknown system states are first estimated using a linear observer. Then, a sliding mode controller utilises this information to reduce the frequency deviations.

While designing any practical controller, it is necessary to consider the constraints of the states and control action. One example is limiting the velocity in order to prevent mechanical damage. In our research, we came upon several methods which combine these bounds with sliding mode control. This motivated us to design such a sliding mode controller, which achieves good performance, despite limitations. This paper builds upon the previous work [13], by taking into account the minimisation of a different quality criterion. The method allows the designer to ensure a priori known bound on the system velocity and/or the control signal value.

2. SLIDING MODE CONTROLLER DESIGN

This section will present the sliding mode controller for the second-order system. During the design, we will focus on the following several main goals:

1. The robustness for the whole control process has to be obtained. It will be achieved by using a time-varying sliding curve, and as the result of a consequence, the reaching phase will be eliminated.
2. The state vector (representative point) has to reach the predetermined demand state in finite time.
3. External disturbances and natural restrictions such as control signal or velocity limitations have to be considered.
4. The integral time absolute error (ITAE) quality index has to be minimised in order to evaluate the performance of the controller.

Let us consider the following system:

$$\dot{\Theta}_1(\tau) = \Theta_2(\tau) \tag{1}$$

$$\dot{\Theta}_2(\tau) = \Psi(\Theta_1(\tau), \Theta_2(\tau), \tau) + \Upsilon(\tau) + \xi\nu(\tau)$$

where $\Theta_1(\tau)$ describes the position of the system and $\Theta_2(\tau)$ is its velocity. Two functions Ψ (function of state vector and time τ) and Υ (external disturbances) vary in time. Nevertheless, due to the practical considerations, the absolute value of the sum of these functions has to be limited from above by a known, positive parameter denoted by Ω. This inequality can be written as follows:

$$|\Psi(\Theta_1(\tau), \Theta_2(\tau), \tau) + \Upsilon(\tau)| \leq \Omega. \tag{2}$$

The main advantage of using inequality (2) is that it allows to describe any bounded time-varying uncertainties. One does not have to assume any frequency distribution of the disturbances, which can be hard to obtain in practice. Positive scalar ξ affects the control signal ν. The representative point starts moving from any position different from the desired one which is equal to 0, i.e $\Theta_1(0) \neq 0$ and $\Theta_2(0) = 0$. Moreover, it has to stop in the desired state. Therefore, the demand point is equal to $(0,0)$. The first goal that we mentioned in this section was the robustness for the whole control process and the elimination of the reaching phase. In order to achieve that goal, we introduce a time-varying, non-linear switching curve described by the following equation:

$$s(\tau) = \Theta_2(\tau) + \kappa(\tau)\operatorname{sgn}\left(\Theta_1(\tau)\right)\sqrt{|\Theta_1(\tau)|}. \tag{3}$$

The function κ corresponds to the variation rate of the sliding curve. The faster κ changes in time, the faster the sliding line evolves. Once κ ceases to change, the sliding line remains stationary. It is given as follows:

$$\kappa(\tau) = \begin{cases} \gamma\tau & for \quad \tau \leq \tau_0 \\ \gamma\tau_0 & for \quad \tau > \tau_0 \end{cases}. \tag{4}$$

Parameter γ is positive and denotes the movement speed of the switching curve. At the time τ_0, that curve stops and remains fixed. Sign function sgn is equal to 1 for positive arguments, is equal to 0 for 0 and is equal to –1 for negative arguments. Substituting $\tau = 0$ one can see that

$$s(0) = \Theta_2(0) + \kappa(0)\operatorname{sgn}\left(\Theta_1(0)\right)\sqrt{|\Theta_1(0)|} = 0, \tag{5}$$

which means that the representative point is on the sliding line at the initial state and the reaching phase is eliminated. It results in robustness to the external disturbances for the whole control process. In order to provide the stable sliding motion, we propose the following control signal:

$$\nu(\tau) = -\frac{1}{\xi}\operatorname{sgn}\left(\Theta_1(\tau)\right)\sqrt{|\Theta_1(\tau)|}\frac{d}{d\tau}\kappa(\tau)$$
$$-\frac{\Theta_2(\tau)\kappa(\tau)}{2\xi\sqrt{|\Theta_1(\tau)|}}\frac{\Omega}{\xi}\operatorname{sgn}\left(s(\tau)\right). \tag{6}$$

Theorem 2.1 Control signal Eq. (6) provides the stable sliding motion for the whole control process. To prove the stability of the sliding motion, we have to guarantee that the following inequality

$$s(\tau)\dot{s}(\tau) \leq 0 \tag{7}$$

is fulfilled for any $\tau > 0$. The strict inequality is not necessary because the representative point is on the switching curve at the initial state. We determine the derivative of the sliding variable as follows:

$$\dot{s}(\tau) = \frac{d}{d\tau}\Theta_2(\tau) + \operatorname{sgn}\left(\Theta_1(\tau)\right)\sqrt{|\Theta_1(\tau)|\dot{\kappa}}(\tau) + \frac{\Theta_2(\tau)\kappa(\tau)}{2\sqrt{|\Theta_1(\tau)|}} \tag{8}$$

Using Eq. (1), we obtain the following:

$$\dot{s}(\tau) = \Psi(\Theta_1(\tau), \Theta_2(\tau), \tau) + \Upsilon(\tau) - \Omega\operatorname{sgn}\left(s(\tau)\right) \tag{9}$$

From Eq. (2) and the sign function properties, we obtain that inequality (7) is true.

3. ADMISSIBLE SETS

In this section, three sets of the switching curve parameters that guarantee the control signal, system's velocity and both of these quantities limitation will be determined. These sets will be composed of two parameters: γ and τ_0. Moreover, we will cover two possible outcomes–when the sliding line stops during the control process and when the demand state is reached when it is still in motion. Let us start by deriving formulas for absolute values of both system's position and system's velocity. From Eq. (3) and the fact that the sliding variable is equal to zero, we have the following:

$$\Theta_2(\tau) = -\kappa(\tau)\operatorname{sgn}\left(\Theta_1(\tau)\right)\sqrt{|\Theta_1(\tau)|}. \tag{10}$$

Using Eqs (1) and (10), one can obtain the following:

$$\Theta_1(\tau) + \kappa(\tau)\operatorname{sgn}\left(\Theta_1(\tau)\right)\sqrt{|\Theta_1(\tau)|} = 0 \quad . \tag{11}$$

The above differential equation is fulfilled for

$$\sqrt{|\Theta_1(\tau)|} = \begin{cases} \delta_1 - \frac{\gamma\tau^2}{4} & for \quad \tau \in [0, \tau_0) \\ \delta_1 - \frac{\gamma\tau^2}{4} & for \quad \tau \in \left(\tau_0, \tau_f\right] \\ 0 & for \quad \tau \in \left(\tau_f, \infty\right) \end{cases}, \tag{12}$$

where τ_f denotes the regulation time. System's position has to be a continuous function. Therefore, in order to fulfil that property, the absolute value of Θ_1 has to be of the following form:

$$|\Theta_1(\tau)| = \begin{cases} \left(\sqrt{|\Theta_1(0)|} - \frac{\gamma\tau^2}{4}\right)^2 & for \quad \tau \in [0, \tau_0) \\ \left(\sqrt{|\Theta_1(0)|} + \frac{\gamma{\tau_0}^2}{4} - \frac{\gamma\tau_0\tau}{4}\right) & for \quad \tau \in \left(\tau_0, \tau_f\right] \\ 0 & for \quad \tau \in \left(\tau_f, \infty\right) \end{cases} \tag{13}$$

and the regulation time is given as follows:

$$\tau_f = \frac{1}{2}\tau_0 + \frac{2\sqrt{|\Theta_1(0)|}}{\gamma\tau_0}. \tag{14}$$

In order to derive the absolute value of the system's velocity, we use Eqs (10) and (13) and get:

$$|\Theta_2(\tau)| =$$
$$\begin{cases} \gamma\tau\sqrt{|\Theta_1(0)|} - \frac{\gamma^2\tau^3}{4} & for \quad \tau \in [0, \tau_0) \\ \gamma\tau_0\sqrt{|\Theta_1(0)|} + \frac{\gamma^2{\tau_0}^3}{4} - \frac{\gamma^2{\tau_0}^2\tau}{4} & for \quad \tau \in \left(\tau_0, \tau_f\right] \\ 0 & for \quad \tau \in \left(\tau_f, \infty\right) \end{cases}. \tag{15}$$

In the second possible scenario, when the switching curve moves for the whole control process, we take into account only two time intervals. In this case, the absolute values of system's position and system's velocity can be written as follows:

$$|\Theta_1(\tau)| = \begin{cases} \left(\sqrt{|\Theta_1(0)|} - \frac{\gamma\tau^2}{4}\right)^2 & for \quad \tau \in [0, \tau_f) \\ 0 & for \quad \tau \in (\tau_f, \infty) \end{cases} \tag{16}$$

$$|\Theta_2(\tau)| = \begin{cases} \gamma\tau\sqrt{|\Theta_1(0)|} - \frac{\gamma^2\tau^3}{4} & for \quad \tau \in [0, \tau_f) \\ 0 & for \quad \tau \in (\tau_f, \infty) \end{cases}. \tag{17}$$

Regulation time is derived as follows:

$$\tau_f = \frac{2\sqrt[4]{|\Theta_1(0)|}}{\sqrt{\gamma}} \quad . \tag{18}$$

3.1. Control signal admissible set

In order to derive the admissible set in which the absolute value of the control signal is bounded from above the following inequality,

$$|\nu(\tau)| \le \nu_{max} \tag{19}$$

has to be fulfilled for any $\tau \ge 0$. Parameter $\nu_{\max}$ is the maximum admissible value of the control signal. After substituting Eqs (6)–(19), we obtain:

$$\left|\mathrm{sgn}(\Theta_1(\tau))\sqrt{\Theta_1(\tau)\dot{\kappa}}(\tau) + \frac{\Theta_2(\tau)\kappa(\tau)}{2\sqrt{\Theta_1(\tau)}} + \Omega\,\mathrm{sgn}(s(\tau))\right| \le |\xi|\nu_{max}. \tag{20}$$

Using properties of the absolute value, we can simplify the above inequality to the following form:

$$\left|\sqrt{|\Theta_1(\tau)|\dot{\kappa}}(\tau) - \frac{\kappa^2(\tau)}{2}\right| \le |\xi|\nu_{max} - \Omega. \tag{21}$$

Let us consider two possible movements of the representative point. In the first one, it moves along the fixed sliding line. In the second one, the representative point will be sliding on a moving switching curve.

1. The sliding line remains fixed.

Using the form of function κ given by Eq. (4), one can see that when the switching curve stops moving, it can be rewritten as follows $\kappa(t) = \gamma\tau_0$. Hence, $\dot{\kappa}(t) = 0$ and (21) can be rewritten as follows:

$$\frac{\gamma^2\tau_0^2}{2} \le |\xi|\nu_{max} - \Omega. \tag{22}$$

Transforming the above equation one can get

$$\tau_0 \le \frac{\sqrt{2(|\xi|\nu_{max} - \Omega)}}{\gamma}. \tag{23}$$

In order to derive the admissible set, we have to analyse the second scenario, when the sliding switching curve is in motion.

2. The sliding line moves.

Now the function κ takes the following form $\kappa(\tau) = \gamma\tau$, and its derivative can be expressed as follows $\dot{\kappa}(\tau) = \gamma$.Therefore, we can rewrite our boundary in the form:

$$\left|\gamma\sqrt{|\Theta_1(\tau)|} - \frac{\gamma^2\tau^2}{2}\right| \le |\xi|\nu_{max} - \Omega. \tag{24}$$

Inequality (22) from the previous scenario shows that $\left|\frac{\gamma^2\tau^2}{2}\right| \le |\xi|\nu_{max} - \Omega$ has to be true for $\tau \le \tau_0$. Parameters γ and τ are both positive; therefore, we will demand that inequality

$$\left|\gamma\sqrt{|\Theta_1(\tau)|}\right| \le |\xi|\nu_{max} - \Omega \tag{25}$$

is always true. The above equation can be transformed into the boundary of the parameter γ:

$$\gamma \le \frac{|\xi|\nu_{max} - \Omega}{\sqrt{|\Theta_1(0)|}}. \tag{26}$$

When the sliding line changes its position for the whole regulation process, we check the edge of the admissible set and obtain that

$$\gamma \le \frac{|\xi|\nu_{max} - \Omega}{2\sqrt{|\Theta_1(0)|}}. \tag{27}$$

One can see that the second inequality is more strict; therefore, we have to demand its fulfilment.

3.2. System's velocity admissible set

This subsection comprises the condition that has to be fulfilled to constrain the velocity of the system. We have to demand that for any $\tau \ge 0$ the following condition

$$|\Theta_2(\tau)| \le \Theta_{2max} \tag{28}$$

is satisfied. Let us start by considering the scenario when the switching curve moves. Hence, we will use the Eq. (17). We have to find the solution of the equation $\dot{\Theta}_2(\tau) = 0$. One can get

$$\gamma\sqrt{|\Theta_1(0)|} - \frac{3}{4}\gamma^2\tau^2 = 0. \tag{29}$$

From the above equation, we get that the maximum absolute value of the system's position is given as follows:

$$\max_{\tau>0}|\Theta_2(\tau)| = \frac{4\sqrt{3\gamma\sqrt[4]{|\Theta_1(0)|^3}}}{9}. \tag{30}$$

and is obtained at the time

$$\tau_{max} = \frac{2\sqrt{3\sqrt[4]{|\Theta_1(0)|}}}{3\sqrt{\gamma}}. \tag{31}$$

Therefore, the limitation on parameter γ can be written as follows:

$$\gamma \le \frac{27\Theta_{2max}^2}{16\sqrt{|\Theta_1(0)|^3}}. \tag{32}$$

In the second scenario, we do not have to consider the case when the sliding line is fixed. In this case, the maximum absolute value of the systems velocity is reached at the time τ_0 due to the shape of the switching curve. So, in this scenario, the constraint on the parameter γ is given as in Eq. (32).

3.3. Control signal and system's velocity admissible set

Taking into account both previous subsections one can write that the constraint on parameter γ in the case when the sliding line becomes fixed during the control process is given as follows:

$$\gamma \le \min\left\{\frac{|\xi|\nu_{max} - \Omega}{\sqrt{|\Theta_1(0)|}}; \frac{27\Theta_{2max}^2}{16\sqrt{|\Theta_1(0)|^3}}\right\}. \tag{33}$$

In the second scenario, when the switching curve moves, we have the following:

$$\gamma \le \min\left\{\frac{|\xi|\nu_{max} - \Omega}{2\sqrt{|\Theta_1(0)|}}; \frac{27\Theta_{2max}^2}{16\sqrt{|\Theta_1(0)|^3}}\right\}. \tag{34}$$

DOI 10.2478/ama-2023-0026

4. MINIMISATION OF ITAE QUALITY INDEX

This section comprises the evaluation of dynamical performance of the system by deriving and minimising the ITAE quality index in the presence of earlier mentioned constraints. This quality index takes the form:

$$I = \int_0^{\infty} |\Theta_1(\tau)| \tau d\tau. \tag{35}$$

From the fact that we have proven that our controller provides that the representative point reaches the desired state in the finite time, the above equation is given as follows:

$$E = \int_0^{\tau_0} |\Theta_1(\tau)| \tau d\tau. \tag{36}$$

In the first scenario, when the sliding line is fixed from some time of the control process, we use Eq. (13) and rewrite ITAE as follows:

$$E = \int_0^{\tau_0} \left(\sqrt{|\Theta_1(0)|} - \frac{\gamma\tau^2}{4}\right)^2 \tau\, d\tau + \int_{\tau_0}^{\tau_f} \left(\sqrt{|\Theta_1(0)|} + \frac{\gamma\tau_0^2}{4} - \frac{\gamma\tau_0\tau}{2}\right)^2 \tau\, d\tau. \tag{37}$$

After some calculations, we can express the above equation in the form:

$$E = \frac{1}{8}|\Theta_1(0)|\tau_0^2 - \frac{1}{48}\gamma\tau_0^4\sqrt{|\Theta_1(0)|} + \frac{1}{768}\gamma^2\tau_0^6 + \frac{\sqrt{|\Theta_1(0)|^3}}{3\gamma} + \frac{|\Theta_1(0)|^2}{3\gamma^2\tau_0^2}. \tag{38}$$

Our goal is to minimise this quality index. Therefore, we will equate the partial derivatives of ITAE with respect to γ and τ_0 to zero as follows:

$$\frac{\partial E}{\partial \gamma} = -\frac{1}{48}\tau_0^4\sqrt{|\Theta_1(0)|} + \frac{1}{384}\gamma\tau_0^6 - \frac{\sqrt{|\Theta_1(0)|^3}}{3\gamma^2} - \frac{2|\Theta_1(0)|^2}{3\gamma^3\tau_0^2}, \tag{39}$$

$$\frac{\partial E}{\partial \tau_0} = \frac{1}{4}|\Theta_1(0)|\tau_0 - \frac{1}{12}\gamma\tau_0^3\sqrt{|\Theta_1(0)|} + \frac{1}{128}\gamma^2\tau_0^5 - \frac{2|\Theta_1(0)|^2}{3\gamma^2\tau_0^3}. \tag{40}$$

Equating Eq. (40) to zero one can get the real form of γ as a function of τ:

$$\gamma = -\frac{4\sqrt{|\Theta_1(0)|}}{3\tau_0^2} \lor \gamma = \frac{4\sqrt{|\Theta_1(0)|}}{\tau_0^2}. \tag{41}$$

From the fact that γ and τ have to be positive, one can get that only the second value of γ in Eq. (41) belongs to the domain. Substituting this value to Eq. (39) and equating it to zero, we have:

$$-\frac{1}{24}\tau_0^4\sqrt{|\Theta_1(0)|} = 0. \tag{42}$$

Again, τ_0 is positive; therefore, the quality index has no stationary points, and as a consequence, the minimum is obtained on a boundary of the admissible set. In the second scenario, when the sliding line moves for the whole regulation process, the ITAE takes the form:

$$E = \int_0^{\tau_f} \left(\sqrt{|\Theta_1(0)|} - \frac{\gamma\tau^2}{4}\right)^2 \tau\, d\tau = |\Theta_1(0)|\frac{\tau_f^2}{2} - \frac{\gamma\tau_f^4}{8}\sqrt{|\Theta_1(0)|} + \frac{\gamma^2\tau_f^6}{96}. \tag{43}$$

Substituting the regulation time Eq. (18) one gets that the above equation can be rewritten as follows:

$$E = \frac{2\sqrt{|\Theta_1(0)|^3}}{3\gamma} \quad . \tag{44}$$

Hence,

$$\frac{\partial E}{\partial \gamma} = -\frac{2\sqrt{|\Theta_1(0)|^3}}{3\gamma^2} \tag{45}$$

and we conclude that again ITAE has no stationary points; therefore, in both cases, the quality index is minimised on the boundary of the admissible set.

Minimisation of ITAE with control signal constraint In this subsection, the minimum value of ITAE with control signal limitation will be derived. We have already shown that ITAE is minimised on the boundary of the admissible set, i.e. on the lines $\gamma = \frac{|\xi|\nu_{\max} - \Omega}{\sqrt{|\Theta_1(0)|}}$ or $\tau_0 = \frac{\sqrt{2(|\xi|\nu_{\max} - \Omega)}}{\gamma}$. Substituting γ to the second equation in Eq. (41), we get:

$$\tau_0 = 2\sqrt{\frac{|\Theta_1(0)|}{|\xi|\nu_{max} - \Omega}}. \tag{46}$$

From the fact that the maximum value of τ_0 on this line is

$$\tau_0 = \sqrt{\frac{2|\Theta_1(0)|}{|\xi|\nu_{max} - \Omega}} \tag{47}$$

we obtain that the only stationary point does not belong to the admissible set. Therefore, the minimum value of ITAE will be obtained for the maximum possible value of τ_0 given by Eq. (47). Now let us focus on deriving the optimal parameters of the switching curve on the line $\tau_0 = \frac{\sqrt{2(|\xi|\nu_{\max} - \Omega)}}{\gamma}$. Substituting this value to Eq. (38), one gets the following:

$$E = \frac{|\Theta_1(0)|(|\xi|\nu_{max} - \Omega)}{4\gamma^2} - \frac{\sqrt{|\Theta_1(0)|}(|\xi|\nu_{max} - \Omega)^2}{12\alpha^3} + \frac{(|\xi|\nu_{max} - \Omega)^3}{96\gamma^4} + \frac{\sqrt{|\Theta_1(0)|^3}}{3\gamma} + \frac{|\Theta_1(0)|^2}{6(|\xi|\nu_{max} - \Omega)}. \tag{48}$$

Calculating the derivative of the right-hand side of the above equation and equating it to zero, we obtain the only stationary point:

$$\gamma = -\frac{\left(1 + \sqrt[3]{2} + \sqrt[3]{4}\right)(|\xi|\nu_{max} - \Omega)}{2\sqrt{|\Theta_1(0)|}}. \tag{49}$$

From the fact that both $(|\xi|\nu_{\max} - \Omega)$ and $|\Theta_1(0)|$ are positive, we get that Eq. (49) is negative. Therefore, again, minimum of ITAE is obtained on the boundary of admissible set and is equal to the following:

$$E = \frac{65|\Theta_1(0)|^2}{96(|\xi|\nu_{max} - \Omega)}. \tag{50}$$

Optimal switching curve parameters are given as follows:

$$\begin{cases} \gamma = \frac{|\xi|\nu_{max} - \Omega}{\sqrt{|\Theta_1(0)|}} \\ \tau_0 = \sqrt{\frac{2|\Theta_1(0)|}{|\xi|\nu_{max} - \Omega}} \end{cases}. \tag{51}$$

In the second case, when the sliding line moves for the whole control process, ITAE is also minimised for the maximum value of $\gamma = \frac{|\xi|\nu_{\max} - \Omega}{2\sqrt{|\Theta_1(0)|}}$ and is expressed as follows:

$$E = \frac{4|\Theta_1(0)|^2}{3(|\xi|\nu_{max} - \Omega)}. \tag{52}$$

One can observe that E determined by Eq. (50) is smaller than the value of the same parameter given by Eq. (52). Hence, we can conclude that the optimal strategy is achieved when the sliding line moves, and after that, it becomes and remains fixed to the end of the control process.

4.1. Minimisation of ITAE using system's velocity constraint

Again, we will start by considering the strategy, when the switching curve stops during the control process. We have already shown that the minimum of ITAE is obtained on the boundary of the admissible set. Hence, it is equal to the following equation:

$$E = \frac{928|\Theta_1(0)|^3}{2187\Theta_{2max}^2}. \tag{53}$$

When the sliding line moves for the whole control process, then the minimum value of ITAE is written as follows:

$$E = \frac{32|\Theta_1(0)|^3}{81\Theta_{2max}^2}. \tag{54}$$

One can easily conclude that the value of Eq. (54) is smaller than that of Eq. (53). Hence, ITAE is minimised when the sliding line moves for the whole control process.

4.2. Minimisation of ITAE by using both control signal and system's velocity constraints

In this section, when the switching curve stops during the control process, we have to consider three possible cases:

1. $\gamma = \frac{|\xi|\nu_{max}-\Omega}{\sqrt{|\Theta_1(0)|}}$ and $\tau_0 = \sqrt{\frac{2|\Theta_1(0)|}{|\xi|\nu_{max}-\Omega}}$. In this case, the minimized ITAE is written as follows:

$$E = \frac{65|\Theta_1(0)|^2}{96(|\xi|\nu_{max}-\Omega)}. \tag{55}$$

2. $\gamma = \frac{27\Theta_{2\mathrm{max}}^2}{16\sqrt{|\Theta_1(0)|^3}}$ and $\tau_0 = \frac{8\sqrt{3}|\Theta_1(0)|}{9\Theta_{2\mathrm{max}}}$. Now ITAE is given by

$$E = \frac{32|\Theta_1(0)|^3}{81\Theta_{2max}^2}. \tag{56}$$

3. $\gamma = \frac{27\Theta_{2\mathrm{max}}^2}{16\sqrt{|\Theta_1(0)|^3}}$ and $\tau_0 = \frac{16\sqrt{2(|\xi|\nu_{\mathrm{max}}-\Omega)|\Theta_1(0)|^3}}{27\Theta_{2\mathrm{max}}^2}$ In this case, the minimum value of ITAE can be expressed as follows:

$$E = \frac{64(|\xi|\nu_{max}-\Omega)|\Theta_1(0)|^4}{729\Theta_{2max}^4} - \frac{1024(|\xi|\nu_{max}-\Omega)^2|\Theta_1(0)|^5}{59049\Theta_{2max}^6} + \frac{2048(|\xi|\nu_{max}-\Omega)^3|\Theta_1(0)|^6}{1594323\Theta_{2max}^8} + \frac{16|\Theta_1(0)|^3}{81\Theta_{2max}^2} + \frac{|\Theta_1(0)|^2}{6(|\xi|\nu_{max}-\Omega)}. \tag{57}$$

Considering the second strategy, when the sliding line moves for the whole control process, we get the minimum value of ITAE as follows:

$$E = \max\left[\frac{4|\Theta_1(0)|^2}{3(|\xi|\nu_{max}-\Omega)}; \frac{32|\Theta_1(0)|^3}{81\Theta_{2max}^2}\right]. \tag{58}$$

One can observe that both Eqs (55) and (56) are always smaller than Eq. (58). Without knowing the values of initial parameters, we will not be able to state which of the three values such as Eqs (55), (56) or (57) will be the ITAE minimum. However, this value can be easily calculated when these parameters are given. The optimal strategy is the one: when the sliding line stops during the control process.

5. SIMULATION EXAMPLE

In this section, we will verify theoretical considerations by introducing the simulation. We present the following system:

$$\begin{cases} \dot{\Theta_1}(\tau) = \Theta_2(\tau) \\ \dot{\Theta_2}(\tau) = \frac{1}{\Pi}\sqrt{|\Theta_1(\tau)|}\arctan\big(\Theta_2(\tau)\big) + \Upsilon(\tau) + \xi\nu(\tau) \end{cases}. \tag{59}$$

The representative point starts at (–4, 0). From the shape of the switching curve and stable sliding motion, one can conclude that the maximum value of Ψ is 1 due to the fact that $\frac{1}{\Pi}\arctan\big(\Theta_2(\tau)\big)$ takes a value from [–1;1]. Absolute value of external disturbances is limited by 3. These disturbances switch 20 times per second between the minimum and maximum admissible value in order to provide the most difficult feedback from the controller. Therefore, Ω takes a value 4. We select parameter $\xi = 1$. We set limits of control signal and systems velocity as follows $\nu_{\mathrm{max}} = 15$ and $\Theta_{2\mathrm{max}} = 4$. Now, let us consider the first case in our paper: when the control signal is constrained. The minimum value of ITAE is

$$E = 0.9848 \tag{60}$$

and parameters related to the switching curve are given as follows:

$$\begin{cases} \gamma = 5.5 \\ \tau_0 = 0.8528s \end{cases}. \tag{61}$$

Representative point reaches the demand state after $\tau_{\mathrm{f}} = 1.2792s$. From the control signal chart shown in Fig. 1, one can observe that it takes its maximum admissible value at the start of a control process. After that, it switches with an amplitude equal to the value of parameter $\Omega = 4$. When the sliding line is in motion, the mean value of the input decreases monotonically. After time τ_0, it stops, and the control signal switches between its minimum admissible value and $-|\nu_{\mathrm{max}} - 2\Omega|$ in order to maintain the stable sliding motion. When it reaches the demand state, it takes values 4 or –4 to remain in it. System position shown in Fig. 2 increases monotonically, and after time τ_{f}, its value is always equal to 0. System's velocity (Fig. 3) starts rising due to the fact that the object must accelerate in order to reach the demand state. After some time, it peaks and starts decreasing. At the time τ_0, we can see a non-differentiable point in our figure. It is the moment when the sliding line stops moving. Again, the demand state is reached in finite time τ_{f}. In the next scenario, when the absolute value of system's velocity is bounded from above, we get the optimal results when the switching curve moves for the whole control process. Minimum value of ITAE is

$$E = 1.5802 \tag{62}$$

and optimal parameter related to the speed of the switching curve is given as follows:

$$\gamma = 3.3750 \tag{63}$$

In this case, due to the fact that the sliding line moves for the whole control process, the control signal decreases monotonically. From Fig. 4, one can observe that in this scenario, it exceeds the value –15 because now we do not require the control signal limitation. After time $\tau_{\mathrm{f}} = 1.5396\ s$, again it switches with amplitude Ω in order to maintain the representative point in a demand state.

DOI 10.2478/ama-2023-0026

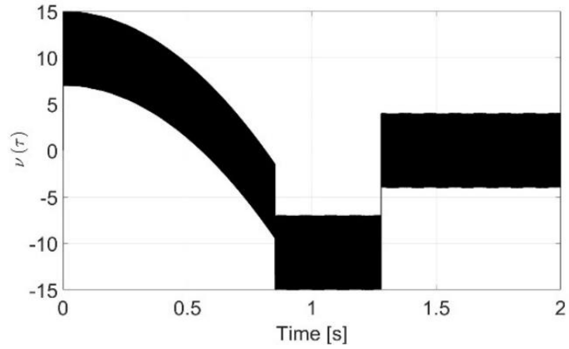


Fig. 1. Control signal

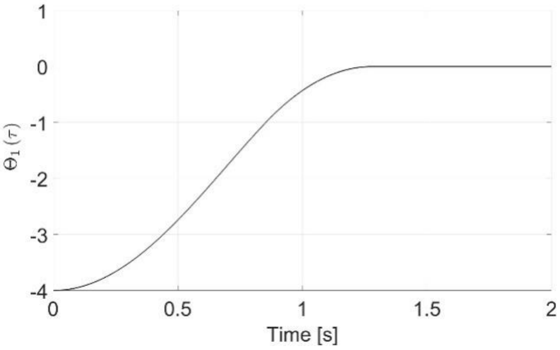


Fig. 2. System's position

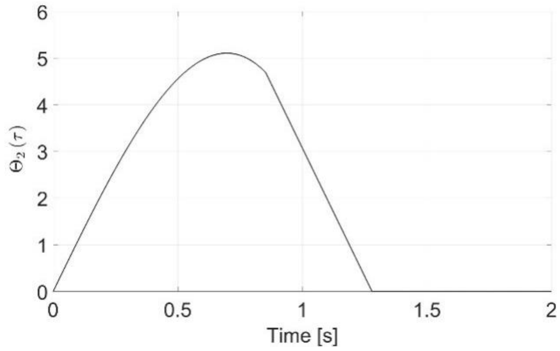


Fig. 3. System's velocity

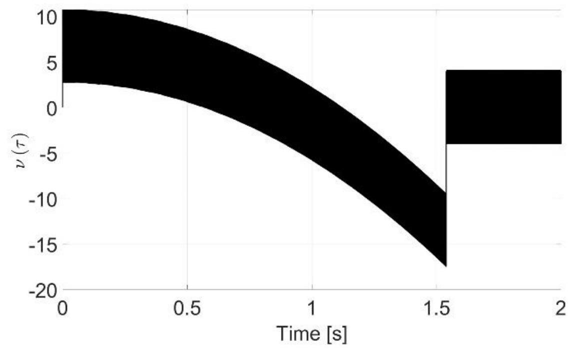


Fig. 4. Control signal

System's position (Figure 5) and velocity behave similarly as in the first case. However, one can observe from Fig. 6 that the limitation of the system's velocity is fulfilled. This is the reason why in this case the regulation time is longer than that in the scenario when the control signal is constrained. The last case covers both control signal and velocity limitations. Minimised ITAE is given as follows:

$$E = 1.5805. \tag{64}$$

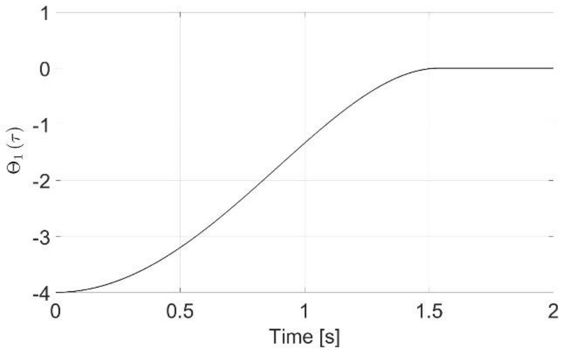


Fig. 5. System's position

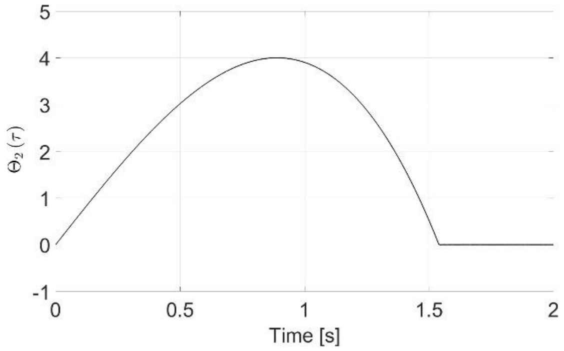


Fig. 6. System's velocity

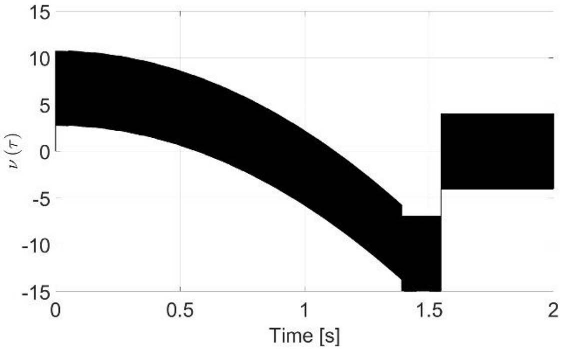


Fig. 7. Control signal

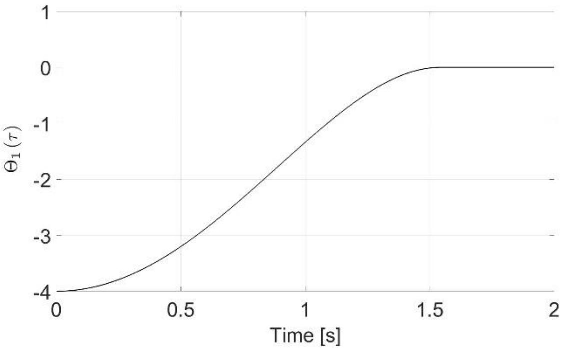


Fig. 8. System's position

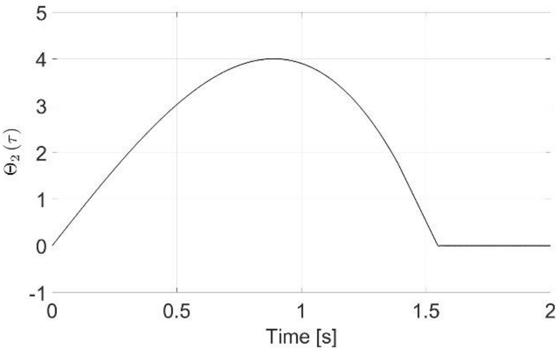


Fig. 9. System's velocity

We can see that it is the highest value of all three scenarios, which is a logical outcome due to the fact that now we have to provide the fulfilment of not one but two limitations. Optimal switching curve parameters are as follows:

$$\begin{cases} \gamma = \ 3.3750 \\ \tau_0 = 0.8528s^{\cdot} \end{cases} \tag{65}$$

The regulation time is $\tau_{\mathrm{f}} = 1.5477$. Again, one can observe that in order to satisfy both constraints, we select the minimum value of γ from both scenarios and the regulation time is the longest one from all three cases. From Figs. 7, 8 and 9, we can see that control signal and system state behave as we expected, and both control signal and velocity limitations are provided. The chattering visible in the control signal in Figs. 1, 4 and 7 could be reduced by changing the sign function of the sliding variable in Eq. (6) to a saturation function. In this paper, we have chosen not to do this to present the main contributions of our approach more clearly. Moreover, unfortunately, using the saturation function would result in a quasi-sliding motion instead of an ideal one. Namely, the state would be constrained to a close vicinity of the sliding line, not necessarily directly to it.

6. CONCLUSIONS

This work comprises the design of a sliding mode controller which can be applied for second-order, nonlinear systems. A switching curve that ensures the elimination of the reaching phase and the robustness for the whole control process is introduced. A finite time convergence of the representative point to the demand state is ensured. Control signal and system's velocity are both constrained separately, and after that, this approach is combined. In order to achieve a satisfying dynamical performance of the system, the ITAE quality index is minimised. It is noticeable that the main difficulty of the approach was considering all the possible scenarios and calculating the optimal parameter values for all of them. However, once this is done, the approach can be used fairly easily, using the final results presented above. Comparing the approach applied in this article with the one used in [14], one can conclude that IAE quality index treats the plant error similarly throughout the whole control process, while ITAE penalises the error more for further time periods. This results in obtaining slightly higher initial values of error for ITAE quality index. However, these values decrease more rapidly.

REFERENCES

1. Ali N, Liu Z, Armghan H, Ahmad I, Hou Y. LCC-S-Based integral terminal sliding mode controller for a hybrid energy storage system using a wireless power system. Energies. 2021;14: 1693.
2. Chang TH, Hurmuzlu Y. Trajectory tracking in robotic systems using variable structure control without a reaching phase, 1992 American Control Conference. IEEE. 1992;1505–1509.
3. Chang TH, Hurmuzlu Y. Sliding control without reaching phase and its application to bipedal locomotion. Journal of Dynamic Systems. Measurement and Control. 1993, 115(3):447–455.
4. Chen Y, Bu W, Qiao Y. Research on the Speed Sliding Mode Observation Method of a Bearingless Induction Motor, Energies. 2021:14.
5. Clemente A, Montiel M, Barreras F, Lozano A, Costa-Castello R, Vanadium Redox Flow Battery State of Charge Estimation Using a Concentration Model and a Sliding Mode Observer. IEEE Access. 2021;9:72368–72376.
6. Drazenović B. The invariance conditions in variable structure systems, Automatica.1969;5(3):287–295.
7. El-Sousy FFM, Alenizi FAF. Optimal Adaptive Super-Twisting Sliding-Mode Control Using Online Actor-Critic Neural Networks for Permanent-Magnet Synchronous Motor Drives, IEEE Access, 2021; 9:82508–82534.
8. Hu J, Zhang H, Liu H, Yu X. A survey on sliding mode control for networked control systems. International Journal of Systems Science. 2021;52(6):1129–1147.
9. Incremona GP, Rubagotti M, Tanelli M, Ferrara A. A general framework for switched and variable-gain higher-order sliding mode control, IEEE Trans. on Aut. Contr. 2020;66(4):1717–1724.
10. Komurcugil H. A new sliding mode control for single-phase UPS inverters based on rotating sliding surface. IEEE International Symposium on Industrial Electronics. 2010;579–584.
11. Komurcugil H. Rotating-sliding-line-based sliding-mode control for single-phase UPS Inverters. IEEE Transactions on Industrial Electronics. 2012;59(10):3719–3726.
12. Li H, Chen X, Zhang H, Cui X. High-Precision Speed Control for Low-Speed Gimbal Systems Using Discrete Sliding Mode Observer and Controller. IEEE Trans. Emerg. Sel. Topics Power Electron. 2022; 10:2871–2880.
13. Pietrala M, Leśniewski P, Bartoszewicz A. Sliding Mode Control with Minimization of the Regulation Time in the Presence of Control Signal and Velocity Constraints. Energies. 2021;14(10):2887.
14. Pietrala M, Leśniewski P, Bartoszewicz A. IAE Minimization in Sliding Mode Control With Input and Velocity Constraints. IEEE Access. 2022;10:28631–28641.
15. Shang W, Jing G, Zhang D, Chen T, Liang Q. Adaptive Fixed Time Nonsingular Terminal Sliding-Mode Control for Quadrotor Formation With Obstacle and Inter-Quadrotor Avoidance. IEEE Access. 2021; 9: 60640–60657,.
16. Skruch P, Długosz M. Design of terminal sliding mode controllers for disturbed non-linear systems described by matrix differential equations of the second and first orders. Applied Sciences. 2019; 9(11):2325–2344.
17. Tang Y. Terminal sliding mode control for rigid robots. Automatica. 1998;34:51–56.
18. Tran AT, Minh BLN, Huynh VV, Tran PT, Amaefule EN, Phan VD, Nguyen TM. Load Frequency Regulator in Interconnected Power System Using Second-Order Sliding Mode Control Combined with State Estimator. Energies. 2021; 14.
19. Ullah N, Mehmood Y, Aslam J, Ali A, Iqbal J. UAVs-UGV, Leader Follower Formation Using Adaptive Non-Singular Terminal Super Twisting Sliding Mode Control. IEEE Access. 2021;9:74385–74405.
20. Utkin V, Drakunov SV. On discrete-time sliding mode control, Proc. IFAC Conf. Nonlinear Control. 1989;484–489.
21. Wang Y, Feng Y, Zhang X, Liang J. A new reaching law for antidisturbance sliding-mode control of PMSM speed regulation system. IEEE Trans. Power Electron. 2020;35(4):4117–4126.

DOI 10.2478/ama-2023-0026

22. Wang T, Tan N, Zhang X, Li G, Su S, Zhou J, Qiu J, Wu Z, Zhai Y, Labati RD, Piuri V, Scotti F. A Time-Varying Sliding Mode Control Method for Distributed-Mass Double Pendulum Bridge Crane With Variable Parameters. IEEE Access. 2021;9:75981–75992.
23. Wang P, Xu Y, Ding R, Liu W, Shu S, Yang X. Multi-Kernel Neural Network Sliding Mode Control for Permanent Magnet Linear Synchronous Motors. IEEE Access. 2021;9:57385–57392.
24. Xu L, Shao X, Zhang W. USDE-Based Continuous Sliding Mode Control for Quadrotor Attitude Regulation: Method and Application. IEEE Access. 2021;9:64153–64164.
25. Yousufzai IK, Waheed F, Khan Q, Bhatti AI, Ullah R, Akmeliawati R. A Linear Parameter Varying Strategy Based Integral Sliding Mode Control Protocol Development and Its Implementation on Ball and Beam Balancer. IEEE Access. 2021;9:74437–74445.
26. Zhihong M, Yu XH, Terminal sliding mode control of MIMO linear systems. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications. 1997;44:1065–1070.

Mateusz Pietrala: iD https://orcid.org/0009-0008-6715-4967

Piotr Leśniewski: iD https://orcid.org/0000-0003-4131-6928

Andrzej Bartoszewicz: iD https://orcid.org/0000-0002-1271-8488

TRAVELLING WAVE SOLUTIONS OF THE NON-LINEAR WAVE EQUATIONS

Jamil A. HAIDER*, Sana GUL*, Jamshaid U. RAHMAN*, Fiazud D. ZAMAN*

*Abdus Salam School of Mathematical Sciences, Government College University,
68-B, New MuslimTown, Lahore 54600, Pakistan

jamilabbashaider@gmail.com, sana_gul_22@sms.edu.pk, jamshaidrahman@gmail.com, f.d.zamann@sms.edu.pk,

received 25 September 2022, revised 28 November 2022, accepted 6 December 2022

Abstract: This article focuses on the exact periodic solutions of nonlinear wave equations using the well-known Jacobi elliptic function expansion method. This method is more general than the hyperbolic tangent function expansion method. The periodic solutions are found using this method which contains both solitary wave and shock wave solutions. In this paper, the new results are computed using the closed-form solution including solitary or shock wave solutions which are obtained using Jacobi elliptic function method. The corresponding solitary or shock wave solutions are compared with the actual results. The results are visualised and the periodic behaviour of the solution is described in detail. The shock waves are found to break with time, whereas, solitary waves are found to be improved continuously with time.

Key words: nonlinear evolution problems, coupled equations, Jacobi elliptic function, periodic solutions

1. INTRODUCTION

In several subfields within physics and engineering, the notion of soliton is an important concept. The solitons are used in modelling of optics, hydrodynamics, nuclear physics, biomechanics, plasma physics and many other fields [1–3]. While in most cases, the solution to nonlinear governing equations is associated with a soliton, this equation describes a wave that maintains its form across time [4]. For the solution of the nonlinear differential equations, several methods are defined such as tangent hyperbolic (tanh) expansion method [5], tanh-sech expansion method [6], exponent expansion method [7], F-expansion method [8], modified simple expansion method [9], exp(-$\emptyset(\alpha)$) expansion method [10], sine-cosine method [11], expansion method [12], coth-expansion function method [13], and some more methods have also been purposed [14–16]. These methods only predict the solution of the solitary and the shock waves but are unable to predict the periodic behaviour of the solutions.

Finding travelling wave solutions of nonlinear partial differential equations is of significant interest, especially in integrable systems [17–26]. In past, the studies presented by different researchers yielded several intriguing forms of solutions, including soliton solutions, cnoidal solutions, compacton solutions and peakon solutions. Nevertheless, as the literature indicates, discovering these answers has not been simple. In recent studies [27–33], the research scholars provided a straightforward method for constructing travelling wave solutions to general nonlinear equations, which may or may not be integrable, beginning with solutions of simple equations (including even linear equations). As proved by nontrivial examples (usually involving equations of the third order), this method is particularly effective for getting travelling wave solutions of nonlinear equations. The generation of travelling wave solutions of a large class of partial differential equations beginning with a trial travelling wave and an invertible map from a trial travelling wave are the primary contributions of this paper. However, we now realise that the underlying reason for the simplicity of our earlier proposal [35] is due to some remarkable properties of travelling wave solutions.

In 1996, a well-known mathematician [36] purposed the shock and solitary wave solutions as well as the periodic solutions by using the Weierstrass elliptic function. In this article, various nonlinear wave equations are solved using the Jacobi elliptic function expansion approach, which is more general than the hyperbolic tangent function expansion method. It is demonstrated that this strategy yields periodic solutions that include certain shock wave and solitary wave solutions. In mathematics, a group of fundamental elliptic functions known as the Jacobi elliptic functions can be found easily. One can find the applications of these functions in the characterisation of the oscillation of a pendulum and in the design of electronic elliptic filters [37]. Jacobi's elliptic operations are the generalisation that refers to those other conics, the ellipse in particular, whereas trigonometry functions are specified concerning a circle. Unlike the Weierstrass elliptic functions, the Jacobi elliptic functions do not need to be explained in terms of complex analysis before they can be used in the real world.

2. METHOD EXPLANATION

The Jacobi elliptic function expansion method is summarised as follows.

Consider a given equation for nonlinear waves

$$G\left(w,\frac{\partial w}{\partial t},\frac{\partial w}{\partial x},\frac{\partial^2 w}{\partial t^2},\frac{\partial^2 \mathrm{w}}{\partial x^2},\dots\right)=0, \tag{1}$$

the travelling wave is the solution to the form

$$w(x,t) = w(\xi)\ , \xi = k(x-ct), \tag{2}$$

where *k* represents the number of waves and *c* represents their speed, correspondingly. A method known as the expansion of the Jacobi elliptic function can be used to represent $w(\xi)$ as a finite series of the Jacobi elliptic function, ξ, which stands for the ansatz.

$$w(\xi) = \sum_{i=0}^{n} a_i\, sn^i\ \xi, \tag{3}$$

is produced based on the highest degree, which is

$$O(w(\xi) = m, \tag{4}$$

which represents the differential equation. From Eq. (3), we have

$$\frac{dw}{d\xi} = \sum_{i=0}^{n} i a_i\, sn^{i-1}\xi(cn\ \xi)dn\ \xi, \tag{5}$$

where all the above functions *sn* ξ, *cn* ξ and $dn\xi$ are the Jacobi elliptic functions of the third kind.

2.1. Relation between the square of the functions

$$sn^2(\xi) + cn^2(\xi) = 1, \tag{6}$$

$$\begin{gathered} dn^2(\xi) + m^2 sn^2(\xi) = 1, \\ cn^2 + m' sn^2 = dn^2, \end{gathered} \tag{7}$$

2.2. Jacobi elliptic functions as solutions of nonlinear ordinary differential equations

The derivatives of the three basic Jacobi elliptic functions are: $(0 \prec m \prec 1)$

$$\begin{aligned} \left(\frac{d}{d\xi}\right)\big(sn(\xi)\big) &= cn(\xi)dn(\xi), \\ \left(\frac{d}{d\xi}\right)\big(cn(\xi)\big) &= -sn(\xi)dn(\xi), \\ \left(\frac{d}{d\xi}\right)\big(dn(\xi)\big) &= -m^2 cn(\xi)sn(\xi), \end{aligned} \tag{8}$$

the ordinary differential equation balances the highest derivative with the highest nonlinear part.

Tab. 1. When $n = 0$ or $n = 1$, the Jacobi elliptic functions are reduced to non-elliptic functions

Function	$\boldsymbol{n = 0}$	$\boldsymbol{n = 1}$
$sn(u,n)$	$sinu$	$tanhu$
$cn(u,n)$	$cosu$	$sechu$
$dn(u,n)$	1	$sechu$
$ns(u,n)$	$cscu$	$cothu$
$nc(u,n)$	$secu$	$coshu$
$nd(u,n)$	1	$coshu$
$sd(u,n)$	$sinu$	$sinhu$
$cd(u,n)$	$cosu$	1
$cs(u,n)$	$cotu$	$cschu$
$ds(u,n)$	$cscu$	$cschu$
$dc(u,n)$	$secu$	1
$sc(u,n)$	$tanu$	$sinhu$

After this process, we choose the order of the ordinary differential equation on its base and choose the series of the function.

$$O\left(\frac{d^q w}{d\xi^q}\right) = m + q\ , q = 1,2,3, \dots \tag{9}$$

$$O\left(w^p\,\frac{d^q w}{d\xi^q}\right) = (p+1)m + q, q = 1,2,3, \dots \tag{10}$$

Therefore, m can be chosen in Eq. (3) to strike a good balance between the highest-order derivative term and the nonlinear component Eq. (1).

Eq. (3) in its following form is shown in Tab. 1.

$$w(\xi) = \sum_{i=0}^{n} a_i tanh^i \xi, \tag{11}$$

Therefore, the Jacobi elliptic function expansion approach is superior than the hyperbolic tangent function expansion method in terms of its universal applicability.

2.3. Applications

A number of the nonlinear models are solved using the Jacobi elliptic function method, and a lot of applications are found in different fields of real word problems.

2.3.1. Model 1

The Korteweg–De Vries (KdV) equation in mathematics is a mathematical model of waves on shallow water surfaces. It is especially significant as the classic example of a perfectly solvable model, i.e., a non-linear partial differential equation whose solutions can be precisely described. The KdV equation can be solved using the inverse scattering transform. The mathematical theory behind the KdV equation is an active area of study. The KdV equation was first presented by Boussinesq and afterwards found by Diederik Korteweg and Gustav de Vries (1895) [38]. Zabusky and Kruskal (1965) [34] discovered statistically that the solutions of the KdV equation appeared to break down at large times into a collection of ‘solitons’: well-separated solitary waves. Moreover, the shape of the solitons appears to be essentially unaffected by their passage through one another (though this could cause a change in their position). By demonstrating that the KdV equation represented the continuum limit of the FPUT system, they established a relationship to earlier numerical experiments by Fermi, Pasta, Ulam and Tsingou. In 1967, Gardner, Greene, Kruskal and Miura developed an analytic solution utilising the inverse scattering transform [39–41].

$$\left(\frac{\partial q}{\partial t}\right) + 12\, q\left(\frac{\partial q}{\partial x}\right) + 2\beta\left(\frac{\partial^3 q}{\partial x^3}\right) = 0 \tag{11}$$

Step 1: Using transformation for the above Eq. (11

$$q(x,t) = q(\xi), \xi = k(x-ct), \tag{12}$$

by using Eq. (12) transformation, the above partial differential equations are converted into the ordinary differential equation, and in the above transformation *k* and *c* represent the wave number and the wave speed.

$$\left(\frac{\partial q}{\partial x}\right) = k\left(\frac{\partial q}{\partial \xi}\right) \tag{13}$$

$$\left(\frac{\partial^3 q}{\partial x^3}\right) = k^3\left(\frac{\partial^3 q}{\partial \xi^3}\right) \tag{14}$$

Putting Eqs (13) and (14) in Eq. (11), the above equation becomes,

$$-2c\left(\frac{\partial q}{\partial \xi}\right)+12\mu q\left(\frac{\partial q}{\partial \xi}\right)+2\beta k^2\left(\frac{\partial^3 q}{\partial \xi^3}\right)=0, \tag{15}$$

which is the required ordinary differential equation that can be obtained by travelling wave solution.

Step 2: Now balancing the above ordinary differential equation balancing the highest derivative with the nonlinear part after this process we balanced our required ordinary differential equation balancing

$$O\left(q\left(\frac{\partial q}{\partial \xi}\right)\right)\ =\ 2m+1, \tag{16}$$

$$O\left(\left(\frac{\partial^3 q}{\partial \xi^3}\right)\right)=m+3, \tag{17}$$

Comparing Eqs (16) and (17)

$$m\ =\ 2 \tag{18}$$

The above ordinary differential equation is balanced at $\mathrm{m}=2$

Step 3: In this method, the finite series method of the Jacobi elliptic function is employed for the nonlinear equation and the values of constants are found. Here the trigonometric cosine and sine functions are used and are represented with *cn* and *sn*.

The solution to the previous equation could take the form of a travelling wave

$$q(\xi)=a_0+a_1 sn(\xi)+a_2 sn^2(\xi)+O\big(sn^3(\xi)\big), \tag{19}$$

we truncated up to two terms because the above is balanced at *m* = 2.

$$\left(\frac{\partial q}{\partial \xi}\right)=\ \big(a_1+2a_2 sn(\xi)\big)cn(\xi)dn(\xi), \tag{20}$$

$$q\left(\frac{\partial q}{\partial \xi}\right)=\ [a_0a_1+(a_1^2+2a_0a_2)sn(\xi)+ \\ 3a_1a_2sn^2(\xi)+2a_2^2sn^3(\xi)]cn(\xi)dn(\xi), \tag{21}$$

$$\left(\frac{\partial^2 q}{\partial \xi^2}\right)= \\ \left.\begin{array}{c}2a_2-(1+m^2)a_1sn(\xi)-4a_2(1+m^2)sn^2(\xi)\\ +2m^2a_1sn^3(\xi)+6m^2a_2sn^4(\xi)\end{array}\right\}, \tag{22}$$

$$\left(\frac{\partial^3 q}{\partial \xi^3}\right)= \\ \left.\begin{array}{c}[-(1+m^2)a_1-8(1+m^2)a_2sn(\xi)]cn(\xi)dn(\xi)\\ +[6m^2a_1sn^2(\xi)+24m^2a_2sn^3(\xi)]cn(\xi)dn(\xi)\end{array}\right\}, \tag{23}$$

Substituting Eqs (20–23) in Eq. (15)

$$\left.\begin{array}{c}-2c\big[\big(a_1+2a_2sn(\xi)\big)cn(\xi)dn(\xi)\big]\\ +12\mu\left[\begin{bmatrix}a_0a_1+(a_1^2+2a_0a_2)sn(\xi)\\ +3a_1a_2sn^2(\xi)+2a_2^2sn^3(\xi)\end{bmatrix}cn(\xi)dn(\xi)\right]\\ +2\beta k^2\left[\begin{bmatrix}-(1+m^2)a_1-8(1+m^2)a_2sn(\xi)\\ +6m^2a_1sn^2(\xi)+24m^2a_2sn^3(\xi)\end{bmatrix}cn(\xi)dn(\xi)\right]\end{array}\right\}=0, \tag{24}$$

Comparing coefficients on both side

$$[-ca_1+a_0a_1-k^2\beta(1+m^2)a_1]=\ 0, \tag{25}$$

$$[-2a^2c+a_1^2+2a_0a_2-8k^2\beta(1+m^2)a_2]=\ 0, \tag{26}$$

$$3a_1a_2+6m^2\beta k^2=\ 0, \tag{27}$$

$$2a_2^2+24m^2\beta k^2a_2=\ 0, \tag{28}$$

By solving the above-mentioned equations using Mathematica Bulit in Software for finding the values of the constants $\mathrm{a}_0, \mathrm{a}_1$ and a_2.

$$a_0=\ c+4\beta k^2+4\beta k^2m^2, \tag{29}$$

$$a_1=\ 0, \tag{30}$$

$$a_2=\ -12k^2m^2\beta, \tag{31}$$

Putting the above values of the constant in the above-mentioned equation

$$q(\xi)=c+4\beta k^2+4\beta k^2m^2- \\ 12k^2m^2\beta sn^2(\xi)+\mathrm{O}\,(sn^2(\xi)), \tag{32}$$

This is the periodic solution precise to the model Eq. (1). In common usage, this solution is known as the cnoidal wave solution to the model equation that came before it. When $\mathrm{m}\ =\ 1$, the expression Eq. (22) is simplified as

$$q(\xi)=c-4\beta k^2+12k^2\beta sech^2(\xi), \tag{33}$$

Which is the model equation's solitary wave solution particularly when c = 1.

$$q(\xi)=3c\big(sech^2(x-ct)\big)\sqrt{\frac{c}{4\beta}}(x-ct), \tag{34}$$

2.3.2. Model 2

This model represents the coupled system of partial differential equations [37].

$$q\left(\frac{\partial q}{\partial t}\right)+12q\left(\frac{\partial q}{\partial x}\right)+\left(\frac{\partial z}{\partial x}\right)+\alpha\left(\frac{\partial^3 q}{\partial t\partial x^2}\right)=\ 0, \tag{35}$$

$$\left(\frac{\partial q}{\partial t}\right)+\left(\frac{\partial (qz)}{\partial x}\right)+6\beta\left(\frac{\partial^3 q}{\partial x^3}\right)=\ 0, \tag{36}$$

These are known as couple equations, so we solve these equations by using a travelling wave solution.

Step 1: Here we suppose the possible travelling wave solution for the above-coupled equations

$$q(x,t)\ =\ q(\xi),\xi=k(x-ct), \\ z(x,t)\ =\ z(\xi),\xi=k(x-ct), \tag{37}$$

By using Eq. (37), we transform the above coupled partial differential equation into an ordinary differential equation. In the above transformation, k and c represent the wave number and the wave speed, respectively.

$$(q)=\ -ck\left(\frac{\partial q}{\partial \xi}\right), \\ \left(\frac{\partial q}{\partial x}\right)=\ k\left(\frac{\partial q}{\partial \xi}\right),\left(\frac{\partial z}{\partial x}\right)=\ k\left(\frac{\partial z}{\partial \xi}\right), \\ \left(\frac{\partial^2 q}{\partial x^2}\right)=\ k^2\left(\frac{\partial^2 q}{\partial \xi^2}\right), \\ \left(\frac{\partial^3 q}{\partial x^3}\right)=\ k^3\left(\frac{\partial^3 q}{\partial \xi^3}\right), \tag{38}$$

Putting Eq. (38) in Eq. (35), we have the following equation:

$$-cq\left(\frac{\partial q}{\partial \xi}\right)+12q\left(\frac{\partial q}{\partial \xi}\right)+\left(\frac{\partial z}{\partial \xi}\right)-ck^2\alpha\left(\frac{\partial^3 q}{\partial \xi^3}\right)=0, \tag{39}$$

which is the required ordinary differential equation that can be obtained using travelling wave solution.

Step 2: Now balancing the above ordinary differential equation balancing the highest derivative with the nonlinear part after this process we balanced our required ordinary differential equation balancing

$$O\left(q\left(\frac{\partial q}{\partial \xi}\right)\right) = 2m+1, \tag{40}$$

$$O\left(\left(\frac{\partial^3 q}{\partial \xi^3}\right)\right) = m+3, \tag{41}$$

Comparing Eqs (40) and (41)

$$m = 2 \tag{41}$$

The above ordinary differential equation is balanced at $\mathrm{m}=2$.

Step 3: In this method, we apply a finite series method of the Jacobi elliptic function on nonlinear equation and find the values of constants. Here we use cosine and sine functions which are represented by *cn* and *sn*.

The solution to the previous equation could take the form of a travelling wave.

$$\begin{aligned} q(\xi) &= a_0 + a_1 sn(\xi) + a_2 sn^2(\xi) + O\big(sn^3(\xi)\big), \\ z(\xi) &= b_0 + b_1 sn(\xi) + b_2 sn^2(\xi) + O\big(sn^3(\xi)\big), \end{aligned} \tag{42}$$

We truncated up to two terms because the above is balanced at *m* = 2.

$$\left(\frac{\partial q}{\partial \xi}\right) = (a_1 + 2a_2 sn\xi)cn(\xi)dn(\xi), \tag{43}$$

$$q\left(\frac{\partial q}{\partial \xi}\right) = [a_0a_1 + (a_1^2 + 2a_0a_2)sn(\xi) + 3a_1a_2sn^2(\xi) + 2a_2^2sn^3\xi]cn(\xi)dn(\xi), \tag{44}$$

$$\left(\frac{\partial z}{\partial \xi}\right) = (b_1 + 2b_2 sn\xi)cn(\xi)dn(\xi), \tag{45}$$

$$\left(\frac{\partial^3 q}{\partial \xi^3}\right) = \left.\begin{matrix}[-(1+m^2)a_1 - 8(1+m^2)a_2sn(\xi)]cn(\xi)dn(\xi) \\ +[6m^2a_1sn^2(\xi) + 24m^2a_2sn^3(\xi)]cn(\xi)dn(\xi)\end{matrix}\right\}, \tag{46}$$

By using the defining values of the above equations, we have,

$$\left.\begin{matrix} -cq[(a_1 + 2a_2sn\xi)cn(\xi)dn(\xi)] + [a_0a_1 + (a_1^2 + 2a_0a_2)sn(\xi)]cn(\xi)dn(\xi) \\ +12[3a_1a_2sn^2(\xi) + 2a_2^2sn^3\xi]cn(\xi)dn(\xi) + [(b_1 + 2b_2sn\xi)cn(\xi)dn(\xi)] \\ -ck^2\alpha\begin{bmatrix} [-(1+m^2)a_1 - 8(1+m^2)a_2sn(\xi)]cn(\xi)dn(\xi) \\ -ck^2\alpha[+6m^2a_1sn^2(\xi) + 24m^2a_2sn^3(\xi)]cn(\xi)dn(\xi)\end{bmatrix}\end{matrix}\right\}, = 0 \tag{47}$$

Comparing coefficients on both sides

$$[-ca_1 + a_0a_1 + b_1 + ck^2\alpha(1+m^2)a_1] = 0, \tag{48}$$

$$[-2a_2c + a_1^2 + 2a_0a_2 + 2b_2 + 8ck^2\alpha(1+m^2)a^2] = 0, \tag{49}$$

$$3a_1a_2 + 6m^2 = 0, \tag{50}$$

$$2a_2^2 - 24m^2a_2 = 0, \tag{51}$$

Step 2: Again, we repeated Step 2 for the coupled partial differential equation.

Now balancing the above ordinary differential equation balancing the highest derivative with the nonlinear part after this process we balanced our required ordinary differential equation balancing

$$O\left(z\left(\frac{\partial q}{\partial \xi}\right)\right) = O\left(q\left(\frac{\partial z}{\partial \xi}\right)\right) = 2m+1, \tag{52}$$

$$O\left(\left(\frac{\partial^3 q}{\partial \xi^3}\right)\right) = m+3, \tag{53}$$

Comparing Eqs (52) and (53)

$$m = 2 \tag{54}$$

The above ordinary differential equation is balanced at $m = 2$

Step 3: In this method, we apply a finite series method of the Jacobi elliptic Function on a nonlinear equation and find the values of constants. Here we use cosine and sine functions which are represented by *cn* and *sn*.

The solution to the previous equation could take the form of a travelling wave.

$$\left(\frac{\partial z}{\partial \xi}\right) = (b_1 + 2b_2 sn\xi)cn(\xi)dn(\xi), \tag{55}$$

$$q\left(\frac{\partial z}{\partial \xi}\right) = \left.\begin{matrix}[a_0b_1 + (a_1b_1 + 2a_0b_2)sn(\xi)]cn(\xi)dn(\xi) \\ +[(2a_1b_2 + a_2b_1)sn^2(\xi) + 2a_2b_2sn^3\xi]cn(\xi)dn(\xi)\end{matrix}\right\}, \tag{56}$$

$$z\left(\frac{\partial q}{\partial \xi}\right) = \left.\begin{matrix}[b_0a_1 + (a_1b_1 + 2b_0a_2)sn(\xi)]cn(\xi)dn(\xi) \\ +[(2b_1a_2 + a_2b_1)sn^2(\xi) + 2b_2a_2sn^3\xi]cn(\xi)dn(\xi)\end{matrix}\right\}, \tag{57}$$

$$\left(\frac{\partial^3 q}{\partial \xi^3}\right) = \left.\begin{matrix}[-(1+m^2)a_1 - 8(1+m^2)a_2sn(\xi)]cn(\xi)dn(\xi) \\ +[6m^2a_1sn^2(\xi) + 24m^2a_2sn^3(\xi)]cn(\xi)dn(\xi)\end{matrix}\right\}, \tag{58}$$

By substituting the values of Eqs (55)–(58) in Eq. (36), we have the following equation:

$$\left.\begin{matrix} -c[(b_1 + 2b_2sn\xi)cn(\xi)dn(\xi)] + [a_0b_1 + (a_1b_1 + 2a_0b_2)sn(\xi)]cn(\xi)dn(\xi) \\ +[(2a_1b_2 + a_2b_1)sn^2(\xi) + 2a_2b_2sn^3\xi]cn(\xi)dn(\xi) \\ +[b_0a_1 + (a_1b_1 + 2b_0a_2)sn(\xi)]cn(\xi)dn(\xi) \\ +[(2b_1a_2 + a_2b_1)sn^2(\xi) + 2b^2a_2sn^3\xi]cn(\xi)dn(\xi) \\ +\beta k^2[-(1+m^2)a_1 - 8(1+m^2)a_2sn(\xi)]cn(\xi)dn(\xi) \\ +\beta k^2[6m^2a_1sn^2(\xi) + 24m^2a_2sn^3(\xi)]cn(\xi)dn(\xi)\end{matrix}\right\}, = 0 \tag{59}$$

By comparing the coefficients on both sides, we have,

$$[-cb_1 + a_0b_1 + a_1b_0 - k^2\beta(1+m^2)a_1] = 0, \tag{60}$$

$$[-2b_2c + 2a_0b_2 + 2a_1b_1 + 2b_0a_2 - 8k^2\beta(1+m^2)a_2] = 0, \tag{61}$$

$$3a_1b_2 + 3a_2b_1 + 6\beta k^2m^2a_1 = 0, \tag{62}$$

$$4a_2b_2 + 24\beta k^2m^2a_2 = 0, \tag{63}$$

By solving the above-mentioned equations by using Mathematica Bulit in Software for finding the values of the constants $\mathrm{a}_0, \mathrm{a}_1, \mathrm{a}_2, \mathrm{b}_0, \mathrm{b}_1$ and b_2.

$$a_0 = c + \left(\frac{\beta}{2\alpha c}\right) - 4ck^2\alpha(1+m^2), \tag{64}$$

$$a_1 = 0, \tag{65}$$

$$a_2 = 12k^2m^2\alpha c, \tag{66}$$

$$b_0 = \left(\frac{\beta^2}{4\alpha^2c^2}\right) + 2\beta(1+m^2)k^2, \tag{67}$$

$$b_1 = 0, \tag{68}$$

$$b_2 = -6c\beta^2m^2, \tag{69}$$

Hence the series solution of the sn for the coupled Eqs (35) and (36) written as

$$q(\xi) = c + \left(\frac{\beta}{2\alpha c}\right) - 4ck^2\alpha(1+m^2) + [12k^2m^2\alpha c]sn^2(\xi) + O\big(sn^3(\xi)\big), \tag{70}$$

$$z(\xi) = \left(\frac{\beta^2}{4\alpha^2c^2}\right) + 2\beta(1+m^2)k^2 + [-6c\beta^2m^2]sn^2(\xi) + O\big(sn^3(\xi)\big), \tag{71}$$

It is the solution for cnoidal waves and the precise periodic solution of Eqs (70) and (71), while the corresponding solitary wave solution for them is

$$q(\xi) = c + \left(\frac{\beta}{2\alpha c}\right) + 4ck^2\alpha - [12k^2\alpha c]sech^2(\xi), \tag{72}$$

$$z(\xi) = -\left(\frac{\beta^2}{4\alpha^2c^2}\right) - 2\beta k^2 - [6c\beta^2]sech^2(\xi), \tag{73}$$

2.3.3. Model 3

Exact solutions of nonlinear evolution equations (NLEEs) are very important to figure out how complex physical phenomena work on the inside. In this work, the new generalised Jacobi elliptic function expansion method is used to look at the exact travelling wave solutions of the Boussinesq equation [42]. With this method, one can get a lot of travelling wave solutions with any parameters, and the wave solutions are written in terms of elliptic functions. It is shown that the new generalised Jacobi elliptic function expansion method is a powerful and clear way to solve nonlinear partial differential equations in mathematical physics and engineering [37].

$$\left(\frac{\partial^2 q}{\partial t^2}\right) - 12c_0^2\left(\frac{\partial^2 q}{\partial x^2}\right) - \alpha q\left(\frac{\partial^4 q}{\partial x^4}\right) - 6\beta\left(\frac{\partial^2 q^2}{\partial x^2}\right) = 0, \tag{75}$$

The solution to Eq. (75) is

$$c^2\big(a_2 - (1+m^2)a_1sn(\xi) - 4a_2(1+m^2)sn^2(\xi) + 2m^2a_1sn^3(\xi) + 6m^2a_2sn^4(\xi)\big) \tag{76}$$

By comparing the coefficients on both sides of the Eq. (76), we find the values of the constants given below:

$$a_0 = \left(\frac{c^2-c_0^2}{2\beta}\right) + \left(\frac{2\alpha k^2}{\beta}\right) + \left(\frac{2m^2\alpha k^2}{\beta}\right), \tag{77}$$

$$a_1 = 0, \tag{78}$$

$$a_2 = -\left(\frac{6m^2\alpha k^2}{\beta}\right), \tag{79}$$

The solution of Model 3 is in the form of the periodic solitary wave by using Eqs (77–79) in Eq. (19) we have,

$$q(\xi) = \left(\frac{c^2-c_0^2}{2\beta}\right) + \left(\frac{2\alpha k^2}{\beta}\right) + \left(\frac{2m^2\alpha k^2}{\beta}\right) - \left(\frac{6m^2\alpha k^2}{\beta}\right)sn^2(\xi) + O\big(sn^3(\xi)\big), \tag{80}$$

The solitary wave solution that corresponds to this one is

$$q(\xi) = \left(\frac{c^2-c_0^2}{2\beta}\right) - \left(\frac{2\alpha k^2}{\beta}\right) + \left(\frac{6\alpha k^2}{\beta}\right)sech^2(\xi), \tag{81}$$

3. CONCLUDING REMARKS

The Jacobi elliptic function expansion approach is presented and applied to nonlinear wave equations in this study. This method is more general than the hyperbolic tangent function expansion method, as demonstrated. In addition, the shock wave and solitary wave solutions are included in the periodic wave solutions derived from the Jacobi elliptic function expansion approach. In the applications, it is demonstrated that the Jacobi elliptic function expansion approach applies to both single and coupled equations. In fact, this method can be used to solve more nonlinear wave equations so long as the odd-order and even-order derivative terms do not overlap in the nonlinear wave equations. It is found that the shock waves break with time, whereas, solitary waves are improved continuously with time.

REFERENCES

1. Asghar S, Haider JA, Muhammad N. The modified KdV equation for a nonlinear evolution problem with perturbation technique. International Journal of Modern Physics B. 2022 Sep 30;36(24):2250160.
2. Fang J, Nadeem M, Habib M, Akgül A. Numerical investigation of nonlinear shock wave equations with fractional order in propagating disturbance. Symmetry. 2022 Jun 8;14(6):1179.
3. Shah NA, El-Zahar ER, Akgül A, Khan A, Kafle J. Analysis of fractional-order regularized long-wave models via a novel transform. Journal of Function Spaces. 2022 Jun 6;2022.
4. Rabie WB, Seadawy AR, Ahmed HM. Highly dispersive Optical solitons to the generalized third-order nonlinear Schrödinger dynamical equation with applications. Optik. 2021 Sep 1;241:167109.
5. Rabie WB, Ahmed HM. Cubic-quartic optical solitons and other solutions for twin-core couplers with polynomial law of nonlinearity using the extended F-expansion method. Optik. 2022 Mar 1;253:168575.
6. Malfliet W. Solitary wave solutions of nonlinear wave equations. American journal of physics. 1992 Jul;60(7):650-4.
7. He JH, Wu XH. Exp-function method for nonlinear wave equations. Chaos, Solitons & Fractals. 2006 Nov 1;30(3):700-8.

8. Zhang JL, Wang ML, Wang YM, Fang ZD. The improved F-expansion method and its applications. Physics Letters A. 2006 Jan 30;350(1-2):103-9.
9. Zayed EM, Arnous AH. The modified wg-expansion method and its applications for solving the modified generalized Vakhnenko equation. Italian Journal of Pure and Applied Mathematics. 2014;32:477-92.
10. Yang AM, Yang XJ, Li ZB. Local fractional series expansion method for solving wave and diffusion equations on Cantor sets. InAbstract and Applied Analysis 2013 Jan 1 (Vol. 2013). Hindawi.
11. Wazwaz AM. A sine-cosine method for handlingnonlinear wave equations. Mathematical and Computer modelling. 2004 Sep 1;40(5-6):499-508.
12. Islam MT, Akbar MA, Azad AK. A rational (G/G)-expansion method and its application to modified KdV-Burgers equation and the (2+ 1)-dimensional Boussineq equation. Nonlinear Stud. 2015 Sep 1;6(4):1-1.
13. Parkes EJ. Observations on the tanh–coth expansion method for finding solutions to nonlinear evolution equations. Applied Mathematics and Computation. 2010 Oct 15;217(4):1749-54.
14. Haider, J.A. and Muhammad, N., 2022. Computation of thermal energy in a rectangular cavity with a heated top wall. *International Journal of Modern Physics B*, *36*(29), p.2250212.
15. Haider JA, Ahmad S. Dynamics of the Rabinowitsch fluid in a reduced form of elliptic duct using finite volume method. International Journal of Modern Physics B. 2022 Dec 10;36(30):2250217.
16. Nadeem S, Haider JA, Akhtar S, Ali S. Numerical simulations of convective heat transfer of a viscous fluid inside a rectangular cavity with heated rotating obstacles. International Journal of Modern Physics B. 2022 Nov 10;36(28):2250200.
17. Hashemi MS, Akgül A. Solitary wave solutions of time–space nonlinear fractional Schrödinger's equation: Two analytical approaches. Journal of Computational and Applied Mathematics. 2018 Sep 1;339:147-60.
18. Hashemi MS, Inc M, Kilic B, Akgül A. On solitons and invariant solutions of the Magneto-electro-elastic circular rod. Waves in Random and Complex Media. 2016 Jul 2;26(3):259-71.
19. Haider JA, Muhammad N. Mathematical analysis of flow passing through a rectangular nozzle. International Journal of Modern Physics B. 2022 Oct 20;36(26):2250176.
20. Seadawy AR, Ahmed HM, Rabie WB, Biswas A. An alternate pathway to solitons in magneto-optic waveguides with triple-power law nonlinearity. Optik. 2021 Apr 1;231:166480.
21. Ahmed HM, Rabie WB, Arnous AH, Wazwaz AM. Optical solitons in birefringent fibers of Kaup-Newell's equation with extended simplest equation method. Physica Scripta. 2020 Oct 16;95(11):115214.
22. Bilal M, Seadawy AR, Younis M, Rizvi ST, Zahed H. Dispersive of propagation wave solutions to unidirectional shallow water wave Dullin–Gottwald–Holm system and modulation instability analysis. Mathematical Methods in the Applied Sciences. 2021 Mar 30;44(5):4094-104.
23. Seadawy AR, Ali A, Albarakati WA. Analytical wave solutions of the (2+ 1)-dimensional first integro-differential Kadomtsev-Petviashivili hierarchy equation by using modified mathematical methods. Results in Physics. 2019 Dec 1;15:102775.
24. Ali I, Seadawy AR, Rizvi ST, Younis M, Ali K. Conserved quantities along with Painleve analysis and Optical solitons for the nonlinear dynamics of Heisenberg ferromagnetic spin chains model. International Journal of Modern Physics B. 2020 Dec 10;34(30):2050283.
25. Khan MA, Akbar MA, binti Abd Hamid NN. Traveling wave solutions for space-time fractional Cahn Hilliard equation and space-time fractional symmetric regularized long-wave equation. Alexandria Engineering Journal. 2021 Feb 1;60(1): 1317-24.
26. Rizvi ST, Seadawy AR, Ashraf F, Younis M, Iqbal H, Baleanu D. Lump and interaction solutions of a geophysical Korteweg–de Vries equation. Results in Physics. 2020 Dec 1;19:103661.
27. Xu C, Farman M, Hasan A, Akgül A, Zakarya M, Albalawi W, Park C. Lyapunov stability and wave analysis of Covid-19 omicron variant of real data with fractional operator. Alexandria Engineering Journal. 2022 Dec 1;61(12):11787-802.
28. Ahmed HM, Rabie WB. Structure of optical solitons in magneto–optic waveguides with dual-power law nonlinearity using modified extended direct algebraic method. Optical and Quantum Electronics. 2021 Aug;53(8):438.
29. Seadawy AR, Lu D, Iqbal M. Application of mathematical methods on the system of dynamical equations for the ion sound and Langmuir waves. Pramana. 2019 Jul;93:1-2.
30. Lu D, Seadawy AR, Arshad M. Bright–dark solitary wave and elliptic function solutions of unstable nonlinear Schrödinger equation and their applications. Optical and Quantum Electronics. 2018 Jan;50:1-0.
31. Ahmad H, Seadawy AR, Khan TA. Numerical solution of Korteweg–de Vries-Burgers equation by the modified variational iteration algorithm-II arising in shallow water waves. Physica Scripta. 2020 Feb 13;95(4):045210.
32. Seadawy AR, Arshad M, Lu D. The weakly nonlinear wave propagation theory for the Kelvin-Helmholtz instability in magnetohydrodynamics flows. Chaos, Solitons & Fractals. 2020 Oct 1;139:110141.
33. Khan MA, Ali Akbar M, Ali NH, Abbas MU. The New Auxiliary Method in the Solution of the Generalized Burgers-Huxley Equation. Journal of Prime Research in Mathematics. 2020;16(2):16-26.
34. Zabusky NJ, Kruskal MD. Interaction of" solitons" in a collisionless plasma and the recurrence of initial states. Physical review letters. 1965 Aug 9;15(6):240.
35. Seadawy AR, Ali S, Rizvi ST. On modulation instability analysis and rogue waves in the presence of external potential: The (n+ 1)-dimensional nonlinear Schrödinger equation. Chaos, Solitons & Fractals. 2022 Aug 1;161:112374.
36. Chen Y, Yan Z. The Weierstrass elliptic function expansion method and its applications in nonlinear wave equations. Chaos, Solitons & Fractals. 2006 Aug 1;29(4):948-64.
37. Haider JA, Asghar S, Nadeem S. Travelling wave solutions of the third-order KdV equation using Jacobi elliptic function method. International Journal of Modern Physics B. 2022 Oct 26:2350117.
38. Korteweg DJ, De Vries G. XLI. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1895 May 1;39(240):422-43.

39. Grébert B. KdV & KAM ergebnisse der mathematik und ihrer grenzgebiete 3. The Mathematical Intelligencer. 2004 Sep;26(3):76-7.
40. Korteweg DJ, De Vries G. XLI. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1895 May 1;39(240):422-43.
41. Wang M, Li X, Zhang J. The (G' G)-expansion method and travelling wave solutions of nonlinear evolution equations in mathematical physics. Physics Letters A. 2008 Jan 21;372(4):417-23.
42. Alam MN, Akbar MA, Roshid HO. Traveling wave solutions of the Boussinesq equation via the new approach of generalized (G'/G)-expansion method. SpringerPlus. 2014 Dec;3:1-9.

Jamil Abbas Haider: https://orcid.org/0000-0002-7008-8576

Sana Gul: https://orcid.org/0000-0002-5075-2599

Jamshaid Ul Rahman: https://orcid.org/0000-0001-8642-0660

Fiazud D Zaman: https://orcid.org/0000-0002-6498-0664

COMPUTATIONAL ANALYSIS OF SORET AND DUFOUR EFFECTS ON NANOFLUID FLOW THROUGH A STENOSED ARTERY IN THE PRESENCE OF TEMPERATURE-DEPENDENT VISCOSITY

Nidhish K. MISHRA*

*Department of Basic Science, College of Science and Theoretical Studies,
Saudi Electronic University, Riyadh 11673, Saudi Arabia

n.kumar@seu.edu.sa

received 9 December 2022, revised 18 January 2023, accepted 19 January 2023

Abstract: In this study, the Soret and Dufour effects in a composite stenosed artery were combined with an analysis of the effect of varying viscosity on copper nanofluids in a porous medium. Blood viscosity, which changes with temperature, is taken into account using the Reynolds viscosity model. The finite difference approach is used to quantitatively solve the governing equations. For use in medical applications, the effects of the physical parameters on velocity, temperature and concentration along the radial axis have been investigated and physically interpreted. The results are graphically displayed and physically defined in order to facilitate comprehension of the various phenomena that occur in the artery when nanofluid is present. It is observed that the Soret effect increases the rate of heat transfer but decreases the rate of mass transfer. The new study enhances knowledge of non-surgical treatment options for stenosis and other abnormalities, hence reducing post-operative complications. Additionally, current research may have biomedical applications such as magnetic resonance angiography (MRA), which provide a picture of an artery and enable identification of any anomalies, and thus may be useful

Keywords: Soret and Dufour effect, variable viscosity, stenosed artery, nanofluid

1. INTRODUCTION

According to the World Health Organization (WHO), about 30% of all deaths globally are caused by cardiovascular diseases (CVD), which are the leading cause of death (1). One of the main causes of health hazards is the reduced blood flow throughout our bodies as a result of artery blockages. In the circulatory system of the human body, arteries transport oxygenated blood and nutrients from the heart to each cell in the body. Red blood cells, white blood cells, platelets and plasma make up blood, a biomagnetic fluid (2). In aqueous plasma, red blood cells, white blood cells and platelets are suspended (3). Aqueous lipoproteins, nutrients and clotting factors make up plasma (4). Blood takes waste products away from the cells and assists in supplying nutrients and oxygen to the cells. The red blood cells' haemoglobin aids in the transportation of oxygen (5). Blood coagulation is aided by the platelets suspended in the plasma (6).

A frequent cause of coronary artery disease is the buildup of fatty substances inside the arterial wall lumen. Through a combination of reduced blood vessel radius and increased blood flow resistance, this process essentially limits the amount of oxygenated blood that the heart can pump to the rest of the body. Due to its critical applications in CVD such as angina, heart attacks, atherosclerosis and aneurysms, which are among the leading causes of death in the world, the haemodynamical study of stenosed arteries is an important area of research. The causes of stenosis and how it affects blood flow have been the subject of several theoretical and practical research studies. Blood travels through arteries and veins, two types of blood vessels. Based on the diameter of the arteries, blood can be classified as either a Newtonian or non-Newtonian fluid (7). Stenosis, or the accumulation of fatty materials inside the lumen, impairs blood flow in the arteries (8). The buildup of lipids in the artery wall or pathological alterations in the tissue structure cause arteriosclerosis, which affects the blood vessels (9). Heart attacks are brought on by stenosis because of increases in blood flow resistance and pressure (3). For this reason, understanding blood flow through a stenosed artery is crucial to understanding cardiovascular disorders. Blood viscosity is a crucial factor to examine because it significantly affects blood flow. Plasma viscosity and packed cell volume are the main determinants of blood viscosity (10). Viscosity is also influenced by red blood cells' mechanical characteristics. Due to their high deformability, red blood cells help the flow of blood during bulk flow and microcirculation (11). Gandhi and Sharma (12) and Poonam and Sharma (13) have examined heat and mass transfer in magnetobiofluid flow under various physical conditions.

Since there is a plethora of applications of heat and mass transfer, the effects of heat and mass transfer occurrence in porous media are receiving more attention. In this instance, the flow phenomena is more complicated than it is in a pure thermal/solutal convection process. Due to its viscosity, fluid experiences a retarding force during flow. A certain amount of energy is lost during this process. Viscous dissipation is the name given to this phenomenon. Since metallic and non-metallic nanoparticles are widely used in nano-haemodynamics, blood purification systems, nanopharmacology and the treatment of haemodynamic disorders, it is becoming increasingly popular to study the rheological aspects of blood flow through arteries affected by cardiovascular disorders while suspending metallic or non-metallic nanoparticles. The effects of pulsing hydromagnetic flow of Au-blood non-Newtonian nanofluid in a channel in the presence of Joule heating, viscosity dissipation and thermal radiation are currently being studied by Thamizharasan and Reddy (14). Tripathi et al. (15–17) explored the interaction of magnetohydrodynamics (MHD) with a number of other important variables, including changing viscosity, Joule heating, radiation and chemical reaction. Hayat et al. (18) and Hafeez et al. (19) discussed computational analysis for velocity slip and diffusion species in the presence of magnetic field.

When heat and mass transfer occur at the same time in a moving fluid, the interactions between the fluxes and the driving potentials are more complicated. It is known that the Soret effect, which is caused by a temperature gradient, and the Dufour effect, which is caused by the concentration effect, are two different types of heat transmission. It was discovered that the Dufour effect is of a size that cannot be disregarded. Eckert and Drake (20) and Sharma et al. (21,22) analyse the thermal and mass diffusion effects on unsteady MHD free and forced convection with radiation and chemical reaction effects. Effects of Soret and Dufour on Jeffrey fluid peristaltic movement in a curved artery/channel in MHD is discussed by Hayat et al. (23). Recently, Siddique et al. (24) explore the effects of Soret and Dufour on the Casson fluid's MHD flow across a stretched surface.

In recent decades, nanofluids have been recognised as an important advancement in biomedical engineering. Theoretical and practical investigations into the potential applications of nanoparticles in blood flow problems have had a significant impact on current bioscience literature. Nanoparticles have a variety of uses, such as surgical tools for treating hyperthermia, medicine carriers, MRI, tracking agents and gene therapy. Microelectronics, fuel cells, pharmaceutical procedures, hybrid-powered engines, household refrigerators, chillers, nuclear reactor coolants, grinding and space technologies are just a few of the many heat transfer uses for nanofluids (25). Nanoparticle suspension in a base fluid is known as a nanofluid. Metals, oxides, carbides, nitrites and nanotubes are among the nanoparticles utilised in the nanofluid (26). Water, ethylene glycol, oil, biofluids and polymer solutions are the most often utilised base fluids (27). Arteriovenous stenosis is treated with nanofluids. Thermophysical characteristics such as thermal conductivity, viscosity of fluid, thermal diffusivity and convective heat transfer are improved in nanofluids (28). Gandhi et al. (29) have recently described the drug delivery applications of magnetohybrid nanoparticles (Au-Al2O3/blood) through different shape-occluded arteries by considering the viscous dissipation effect, Joule heating and variable viscosity of the fluid. As a result, monitoring blood flow while a nanofluid is present is crucial to the healing process. Some more researchers (30–34) developed mathematical models to illustrate the effect of different nanoparticles on fluid flow through different geometries.

The flow system used in all of the aforementioned research is for a fluid with constant viscosity. The viscosity of the substance might, however, differ significantly. It is vital to consider a viscosity fluctuation in order to correctly forecast flow behaviour. Hematocrit and temperature have an impact on blood viscosity, cerebral blood flow adaptation to oxygen metabolism in the brain, or both, which can affect how quickly blood flows through the brain. Blood viscosity can alter as a result of hematocrit changes. When blood temperature drops, blood viscosity rises. As temperature is lowered from 37°C to 22°C, blood viscosity rises by 50%–300% (35,36). When deliberately inducing profound hypothermia for cardiac or thoracic aortic procedures requiring temporary circulatory arrest, blood temperatures as low as 8–12°C are frequently measured (37–39). The impact of variable viscosity (temperature dependent) on MHD chemically reacting blood flow with heat and mass transfer was examined by Sharma et al. (40). Further, a mathematical model was recently created by Sharma et al. (41) to show how temperature-dependent viscosity affects blood flow through the stretching surface.

Based on the literature review and to our knowledge, no attempt is made to examine the Soret and Dufour effects on nanofluid flow through a stenosed artery in the presence of temperature-dependent viscosity. The following are the objectives and novelty of the current study: As blood viscosity is temperature-dependenttso, so the Reynolds viscosity model is considered which states that blood viscosity varies exponentially with temperature. Soret and Dufour effects on nanofluid flow through a stenosed artery are considered. No-slip boundary condition is taken into account in this case. The velocity, concentration and temperature profiles are presented using a constant pressure gradient and a steady flow assumption. Additionally covered are the Brinkmann number, viscosity parameter and Soret and Dufour effects. These findings are explained in a way that highlights the impact of different blood flow characteristics in a sick state.

2. MATHEMATICAL FORMULATION

We consider an axisymmetric flow of blood through a composite stenosed artery in a circular tubule of finite length L, as shown in Fig. 1. The geometry of the composite stenosis in an arterial wall is described by Joshi and Srivastava (42,43) as

$$\frac{R(z)}{R_0} = \left\{ \begin{array}{ll} 1 - \frac{2\delta}{R_0 L_0}(z-d) & d < z \leq d + \frac{L_0}{2} \\ 1 - \frac{\delta}{2R_0}\left(1 + \cos\frac{2\pi}{L_0}\left(z - d - \frac{L_0}{2} \right) \right) & d + \frac{L_0}{2} < z \leq d + L_0 \\ 1 & otherwise \end{array} \right\} \tag{2.1.}$$

where *R(z)* is the radius of the artery in the obstructed region, and R_0 is the radius of normal artery. Δ, L_0 and *d*, are the height, length and location of the stenosis, respectively. β is the angle of inclination of the artery from the vertical axis.

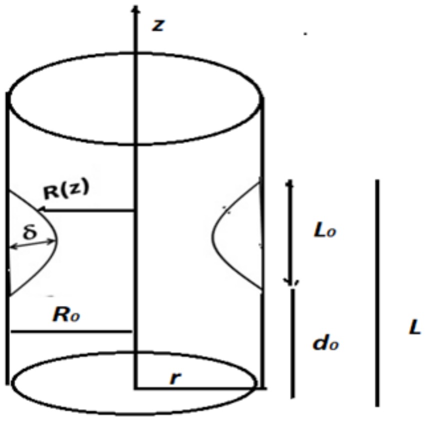


Fig. 1. Geometry of the problem

Using the above assumptions, the governing equations of the physical model can be written as the following:

Continuity equation:

$$\frac{1}{r}\frac{\partial(rv)}{\partial r}+\frac{\partial u}{\partial z}=0 \tag{2.2}$$

Momentum equation (r-direction):

$$\rho_f\left(v\frac{\partial v}{\partial r}+u\frac{\partial v}{\partial z}\right)=-\frac{\partial p}{\partial r}-\left[\frac{1}{r}\frac{\partial(r\tau_{rr})}{\partial r}+\frac{\partial\tau_{zr}}{\partial z}-\frac{\tau_{\theta\theta}}{r}\right] \tag{2.3}$$

Momentum equation (z-direction):

$$\rho_f\left(v\frac{\partial u}{\partial r}+u\frac{\partial u}{\partial z}\right)=-\frac{\partial p}{\partial z}-\left[\frac{1}{r}\frac{\partial(r\tau_{rz})}{\partial r}+\frac{\partial\tau_{zz}}{\partial z}\right]$$
$$-\mu(T)\frac{u}{k_1}+\rho_f g\alpha_T(T-T_0)\ +\rho_f g\alpha_c(C-C_0) \tag{2.4}$$

Energy equation:

$$\left(\rho C_p\right)_f\left(v\frac{\partial T}{\partial r}+u\frac{\partial T}{\partial z}\right)=k_T\left[\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}+\frac{\partial^2 T}{\partial z^2}\right]+$$
$$\left\{\frac{D_m K_T}{T_m}\left[\frac{\partial^2 C}{\partial r^2}+\frac{1}{r}\frac{\partial C}{\partial r}+\frac{\partial^2 C}{\partial z^2}\right]\right\}+\ \mu(T)\left(\frac{\partial u}{\partial r}\right)^2 \tag{2.5}$$

Concentration equation:

$$\left(v\frac{\partial C}{\partial r}+u\frac{\partial C}{\partial z}\right)=D_m\left[\frac{\partial^2 C}{\partial r^2}+\frac{1}{r}\frac{\partial C}{\partial r}+\frac{\partial^2 C}{\partial z^2}\right]$$
$$+\ \frac{D_m K_T}{T_m}\left[\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}+\frac{\partial^2 T}{\partial z^2}\right] \tag{2.6}$$

The legends for the symbols used in the above equations have been defined in the nomenclature section.

The corresponding boundary conditions are the symmetry at the centreline and no-slip at the walls, as indicated below:

– At r = 0,

$$\frac{\partial u}{\partial r}=0,\qquad \frac{\partial T}{\partial r}=0,\qquad \frac{\partial C}{\partial r}=0 \tag{2.7}$$

– At r = $R(z)/R_0$

$$u=0,\qquad T=T_0,\qquad C=C_0 \tag{2.8}$$

We proceed to introduce the following non-dimensional parameters:

$$\bar{r}=\frac{r}{R_0},\quad \bar{p}=\frac{R_0^2}{UL_0\mu_f}p,\quad \bar{z}=\frac{z}{L_0},\quad \mu(\theta)=\frac{\mu(T)}{\mu_f},$$

$$D_f=\ \frac{D_m K_T C_0}{\mu_f T_m C_{pf} T_0},\quad \bar{R}=\frac{R}{R_0},\ K_1=\frac{k_1}{R_0}$$

$$\bar{u}=\frac{u}{U},\ \ E_c=\ \frac{U^2}{C_{pf}T_0},Br=E_c Pr,\quad \sigma=\frac{C-C_0}{C_0},$$

$$S_c=\frac{\mu_f}{p_f D_m},\ \ Gr=\frac{g\alpha_T R_0^2 T_0\rho_f}{U\mu_f},\ P_r=\ \frac{\mu_f c_{pf}}{k},$$

$$\bar{v}=\frac{L_0}{\delta U}v,\ \ \bar{d}=\frac{d}{L_0},\quad S_r=\frac{D_m K_T\rho T_0}{\mu T_m C_0},$$

$$Cr=\ \frac{g\alpha_C R_0^2 C_0\rho_f}{U\mu_f},\quad \bar{\delta}=\frac{\delta}{R_0},\theta=\frac{T-T_0}{T_0},\quad \alpha=\frac{k}{(\rho C_p)_f}$$

Making use of these non-dimensional variables and applying the additional conditions $\epsilon=\frac{R_0}{L_0}\cong O(1)$ for the case of mild stenosis $\left(\frac{\delta}{R_0}\ll 1\right)$ (44) to Eqs (2.1)–(2.6), the governing equations in non-dimensional form after dropping the dashes can be written as follows:

$$\frac{\partial p}{\partial r}=0 \tag{2.9}$$

$$\frac{\partial p}{\partial z}=\frac{1}{r}\frac{\partial}{\partial r}\left(r\mu(\theta)\frac{\partial u}{\partial r}\right)-\left(\frac{\mu(\theta)}{K_1}\right)u+Gr\theta+Cr\sigma \tag{2.10}$$

$$\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\theta}{\partial r}\right)\right)+DfPr\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\sigma}{\partial r}\right)\right]+Br\mu(\theta)\left(\frac{\partial u}{\partial r}\right)^2=0 \tag{2.11}$$

$$\frac{1}{S_c}\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\sigma}{\partial r}\right)\right)+S_r\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\theta}{\partial r}\right)\right)=0 \tag{2.12}$$

Reynold's model of viscosity is considered in this study, and the same is defined as follows:

$$\mu(\theta)=e^{-\omega\theta} \tag{2.13}$$

Taking the Maclaurins expansion of the exponential terms in Eq. (2.13) and considering up to two terms, we obtain $\mu(\theta)=1-\omega\theta$.

The boundary conditions of velocity, temperature and concentration in dimensionless form are the following:

– At r = 0

$$\frac{\partial u}{\partial r}=0,\qquad \frac{\partial\theta}{\partial r}=0,\qquad \frac{\partial\sigma}{\partial r}=0 \tag{2.14}$$

– At r = R(z)

$$u=0,\qquad \theta=0\qquad \sigma=0 \tag{2.15}$$

The geometry of the arterial wall in non-dimensional form will be:

$$R(z)=\left\{\begin{matrix} 1-2\delta(z-d) & d<z\underline{\leq} d+\frac{1}{2} \\ 1-\frac{\delta}{2}\left(\begin{matrix}1+\\ \cos 2\pi\left(\begin{matrix}z-d\\ -\frac{1}{2}\end{matrix}\right)\end{matrix}\right) & d+\frac{1}{2}<z\underline{\leq} d+1 \\ 1 & otherwise\end{matrix}\right\}. \tag{2.16}$$

3. NUMERICAL SOLUTION OF THE PROBLEM

The explicit finite difference approach has been used to derive numerical solutions for the non-linear dimensionless partial differential equations with corresponding boundary conditions as mentioned above. The first order forward difference scheme is the foundation for the discretisation of first order derivative terms, whereas the central difference scheme serves as the foundation for second order terms. The blood flow zone is segmented into a grid or mesh of lines in order to obtain the difference equations. At the junction of these mesh lines, known as nodes, these difference schemes' solutions are found. A tri-diagonal system of equations is made up of the finite difference equations at every internal nodal point on a certain *n*-level. Using the Thomas algorithm (45), these equations are resolved. These are the finite difference equations:

$$\left(\frac{\partial p}{\partial z}\right)_{i,j}=\frac{(\mu_{i,j})}{r_{i,j}}\left(\frac{u_{i+1,j}-u_{i,j}}{\Delta r}\right)+\left(\mu_{i,j}\right)\frac{u_{i+1,j}-2u_{i,j}+u_{i-1,j}}{(\Delta r)^2}-$$
$$\left(\frac{(\mu_{i,j})}{K}\right)u_{i,j}+\left(\frac{u_{i+1,j}-u_{i,j}}{\Delta r}\right)\left(\frac{\mu_{i+1,j}-\mu_{i,j}}{\Delta r}\right)Gr\frac{\theta_{i+1,j}-\theta_{i,j}}{\Delta r}+$$
$$Cr\frac{\sigma_{i+1,j}-\sigma_{i,j}}{\Delta r} \tag{3.1}$$

$$\left(\frac{1}{r_{i,j}}\frac{\theta_{i+1,j}-\theta_{i,j}}{\Delta r}+\frac{\theta_{i+1,j}-2\theta_{i,j}+\theta_{i-1,j}}{(\Delta r)^2}\right)+Br\left(\mu_{i,j}\right)\left(\left(\frac{u_{i+1,j}-u_{i,j}}{(\Delta r)^2}\right)^2\right)+$$
$$D_f Pr\left(\frac{1}{r_{i,j}}\frac{\sigma_{i+1,j}-\sigma_{i,j}}{\Delta r}+\frac{\sigma_{i+1,j}-2\sigma_{i,j}+\sigma_{i-1,j}}{(\Delta r)^2}\right)=0 \tag{3.2}$$

$$\frac{1}{S_c}\left(\frac{1}{r_{i,j}}\frac{\sigma_{i+1,j}-\sigma_{i,j}}{\Delta r}+\frac{\sigma_{i+1,j}-2\sigma_{i,j}+\sigma_{i-1,j}}{(\Delta r)^2}\right)+S_r\left(\frac{1}{r_{i,j}}\frac{\theta_{i+1,j}-\theta_{i,j}}{\Delta r}+\frac{\theta_{i+1,j}-2\theta_{i,j}+\theta_{i-1,j}}{(\Delta r)^2}\right)=0 \qquad (3.3)$$

The appropriate mesh size for the above calculation is $\Delta r = 0.0625$. The procedure was carried out iteratively till the error was $<10^{-5}$. In order to ensure the accuracy of our results, the computation is carried out for slightly changed values of Δr. After each cycle of iteration in which the convergence criteria are assessed to confirm their validity, the values of velocity, temperature and concentration are taken note of, and on studying these values, we observe negligible change.

4. RESULTS AND DISCUSSION

The influence of the Soret and Dufour effects with permeable arterial wall and temperature-dependent viscosity, on MHD blood flow through a stenosed artery with heat and mass transfer, has been examined in this paper. Gr = 0.5, Cr = 0.3, Br = 0.3, Sc = 0.5, Sr = 0.5, Pr = 0.7, =0.02, Df*Sr = 0.12 and d = 0.25, are the default parameter values used in the calculations (46–51). Numerous numerical estimates have been made for various Sr, Df, Sc, Pr, Cr, Gr and Br values. By keeping the mean temperature *Tm* constant, the values of the Soret and Dufour numbers are selected in a way that ensures their product remains constant. Figs. 2–7 show the numerical outcomes for the velocity, temperature, concentration and shear stress profiles for various parameters. In Fig. 2, the viscosity reduces as the velocity increases. Fig. 3 depicts that with increasing the values of Sc, velocity increases. This is consistent with the physical behaviour that shows that as copper's volume percentage rises, thermal conductivity rises as well, leading to an increase in thermal boundary-layer thickness. Fig. 4 depicts the velocity curve for various values of Gr. It has been found that velocity grows as Gr increases. This is because the buoyant force is increasing as a result of the geographical variation in fluid density brought on by widening temperature differences. Additionally, as shown in Fig. 5, velocity falls as the adjusted Grashof number rises. As illustrated in Fig. 6, velocity rises as the Brinkman number rises due to an increase in heat conduction. Fig. 7 shows that blood velocity increases as the parameter K1 is raised.

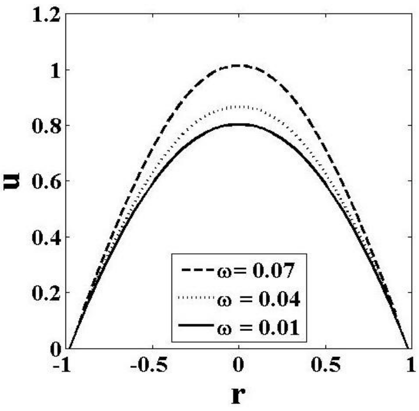


Fig. 2. Velocity profile against radial axis for varying ω

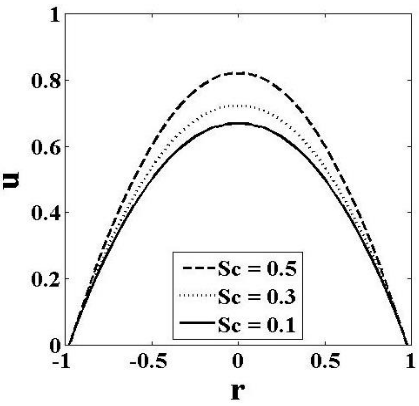


Fig. 3. Velocity profile against radial axis for varying Sc

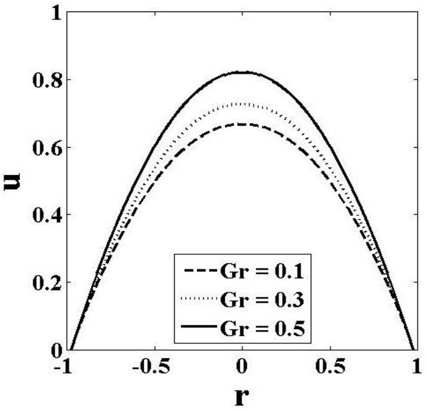


Fig. 4. Velocity profile against radial axis for varying Gr

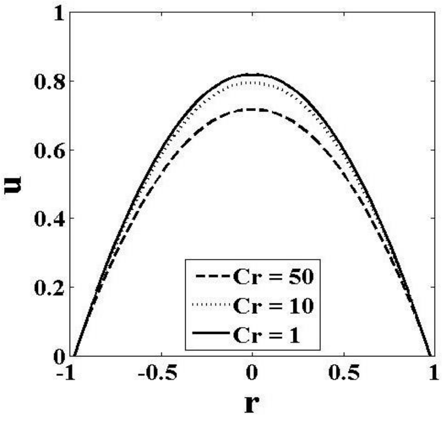


Fig. 5. Velocity profile against radial axis for varying Cr

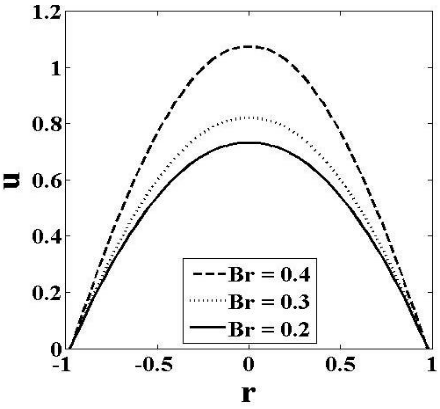


Fig. 6. Velocity profile against radial axis for varying Br

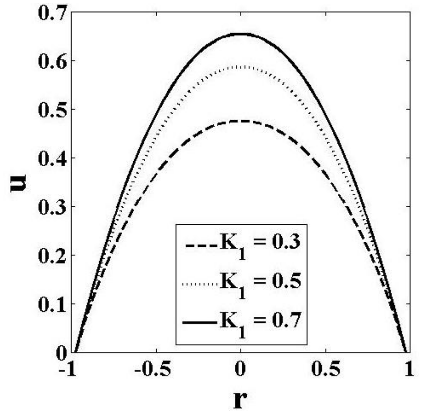


Fig. 7. Velocity profile against radial axis for varying $\mathrm{K1}$

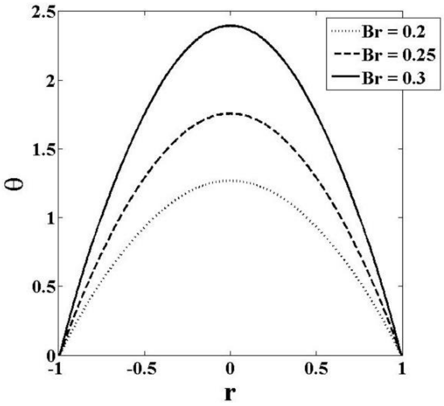


Fig. 8. Temperature profile against radial axis for varying Br

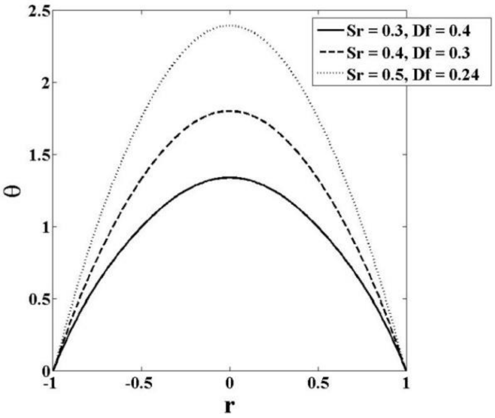


Fig. 9. Temperature profile against radial axis for varying $\mathrm{Sr} - \mathrm{Df}$

Temperature profiles for different values of Sr, Df, ω and Br are shown in Figs. 8–10. The influence of the Brinkman number on temperature distribution is depicted in Fig. 8. The Brinkman number measures the relationship/ratio prevailing between the heat produced by viscous dissipation and that by molecular transport.. As can be seen from Fig. 8, temperature rises as Br increases, suggesting that the viscous dissipation effect predominates over thermal diffusion in terms of temperature variation. As seen in Fig. 9, temperature rises with increasing Soret number and falls with increasing Dufour number. Owing to the temperature gradient, temperature rises as the Soret number of thermophoretic diffusion rises. Additionally, it has been found that temperature drops as the Dufour number rises because a decrease in concentration gradient results in a reduction in energy flux. As seen in Fig. 10, it is further found that temperature rises with concentration profiles for different values of ω, Sc, Sr, Df and Br are shown in Figs. 11–14.

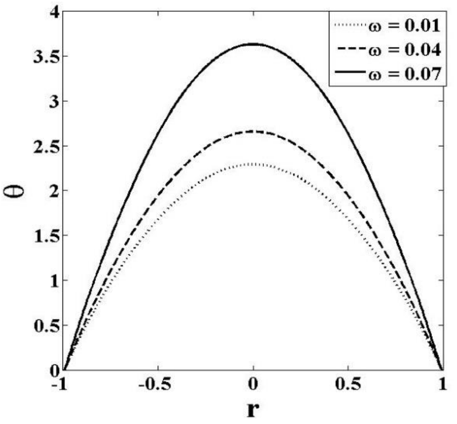


Fig. 10. Temperature profile against radial axis for varying ω

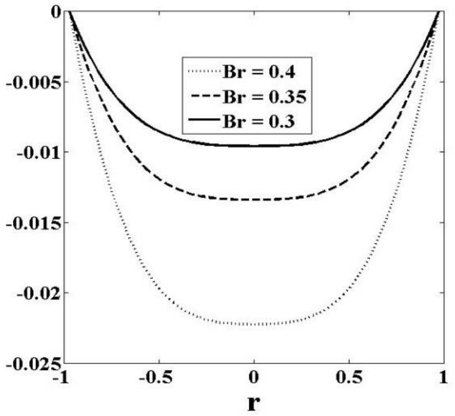


Fig. 11. Concentration profile against radial axis for varying Br

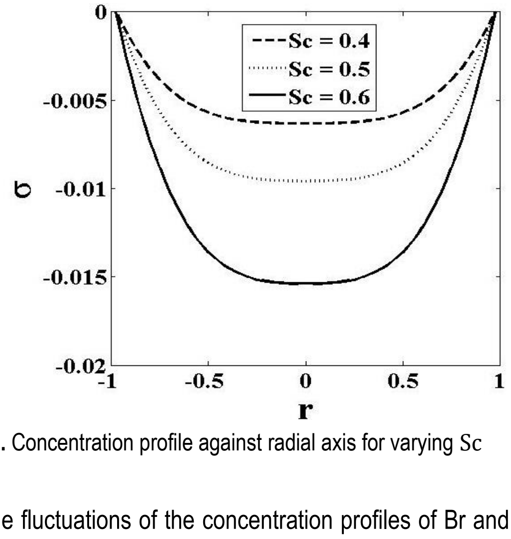


Fig. 12. Concentration profile against radial axis for varying Sc

The fluctuations of the concentration profiles of Br and Sc are depicted in Figs. 11 and 12. Additionally, it has been noted that the concentration profile behaves differently from the temperature profile. It is observed that concentration falls as the Brinkman number rises. This happens as a result of the rise in energy flux brought on by viscous dissipation. According to Fig. 12, the rise in momentum diffusivity causes concentration to fall as Sc increases. As a result, the concentration buoyancy effects lessen, which lowers the fluid velocity. The thickness of the momentum and concentration boundary layers simultaneously decreases with the drops in the velocity, temperature and concentration profiles. Fig. 13 depicts the influence of Soret and Dufour numbers on the concentration field. We observe that the concentration profile descends in concatenation with the increase in the Soret number and the decrease in the Dufour number, which is attributable the fact that the gradient of concentration has gotten steeper in pursuance of the change. The influence of the viscosity parameter is shown in Fig. 14; as the viscosity parameter increases, the concentration also increases.

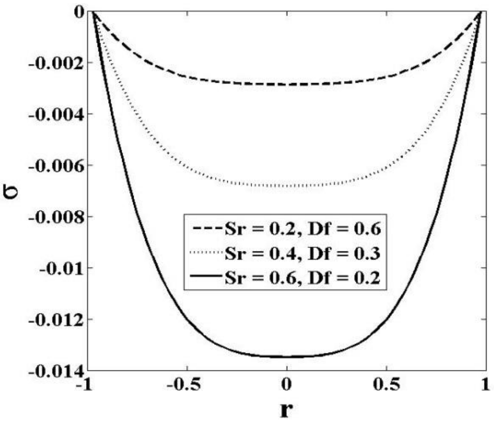


Fig. 13. Concentration profile against radial axis for varying $\mathrm{Sr} - \mathrm{Df}$

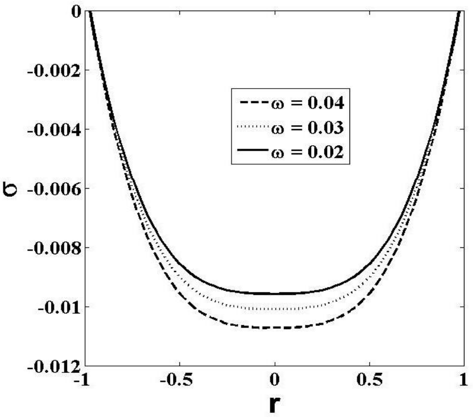


Fig. 14. Concentration profile against radial axis for varying ω

5. CONCLUSIONS

To analyse the problem of unsteady boundary layer, as well as the heat and mass transfer flow of a viscous incompressible electrically-conducting fluid (blood) through a porous medium over a composite stenosed artery subject to the Soret and Dufour effects, the finite difference method is used to solve the governing highly non-linear partial differential equations, along with the boundary conditions. The dependence of blood viscosity on temperature is taken into consideration. The key points of the above analysis are as follows:

- It is seen that velocity field rises with increasing Br, Gr, Sc and ω, whereas it decreases with increasing Cr.
- It is also noted that a reverse phenomenon is demonstrated in the concentration profile in comparison with the temperature profile.
- It is observed that increase in Br and ω increases the temperature whereas increase in Df and Pr decreases the temperature.
- The Soret number has a positive impact on the rate of heat transfer but a negative impact on the rate of mass transfer.

In addition to its utility in examining more complex difficulties involving the Soret and Dufour effects, it is envisaged that the current work will be helpful in the study of arterial blood flow issues, which can serve as the foundation for numerous scientific and industrial applications.

REFERENCES

1. Chan BT, Lim E, Chee KH, Osman NAA. Review on CFD simulation in heart with dilated cardiomyopathy and myocardial infarction. Comput Biol Med. 2013;43(4):377–385. https://doi.org/10.1016/j.compbiomed.2013.01.013. Epub 2013
2. Akbar NS, Nadeem S. Blood flow analysis in tapered stenosed arteries with pseudoplastic characteristics. Int J Biomath. 2014;7(6):1450065. https://doi.org/10.1142/S179352451450065X
3. Sankar DS. Two-phase non-linear model for blood flow in asymmetric and axisymmetric stenosed arteries. Int J Non Linear Mech. 2011;46(1):296–305. https://doi.org/10.1016/j.ijnonlinmec.2010.09.011
4. Thurston GB. Erythrocyte rigidity as a factor in blood rheology: viscoelastic dilatancy. J Rheol (N Y N Y). 1979;23(6):703–719. https://doi.org/10.1122/1.549506
5. Wolberg AS, Campbell RA. Thrombin generation, fibrin clot formation and hemostasis. Transfus Apher Sci. 2008;38(1):15–23. https://doi.org/10.1016/j.transci.2007.12.005
6. Roustaei M, Nikmaneshi MR, Firoozabadi B. Simulation of Low Density Lipoprotein (LDL) permeation into multilayer coronary arterial wall: Interactive effects of wall shear stress and fluid-structure interaction in hypertension. J Biomech. 2018;67(4):114–122.
7. Huckabe CE, Hahn AW. A generalized approach to the modeling of arterial blood flow. Bull Math Biophys. 1968;30(4):645–662. https://doi.org/10.1007/BF02476681
8. Ellahi R, Rahman SU, Nadeem S, Akbar NS. Blood flow of nanofluid through an artery with composite stenosis and permeable walls. Appl Nanosci. 2014;4(8):919–926. https://doi.org/10.1007/s13204-013-0253-6
9. Liepsch D, Singh M, Lee M. Experimental analysis of the influence of stenotic geometry on steady flow. Biorheology. 1992;29(4):419–431. https://doi.org/10.3233/bir-1992-29405
10. Lowe GD, Drummond MM, Lorimer AR, Hutton I, Forbes CD, Prentice CR, et al. Relation between extent of coronary artery disease and blood viscosity. Br Med J. 1980;280(6215):673–674. https://doi.org/ 10. 1136/bmj.280.6215.673
11. Baskurt OK, Meiselman HJ. Blood Rheology and Hemodynamics. Semin Thromb Hemost [Internet]. 2003 Nov 21 [cited 2021 Sep 11];29(05):435–450. Available from: https://doi.org/10.1055/s-2003-44551
12. Gandhi R, Sharma BK. Unsteady MHD Hybrid Nanoparticle (Au-Al 2 O 3/Blood) Mediated Blood Flow Through a Vertical Irregular Stenosed Artery: Drug Delivery Applications. In: Nonlinear Dynamics and Applications: Proceedings of the ICNDA 2022. Springer; 2022;12(2):325–337. https://doi.org/10.1007/978-3-030-99792-2_28
13. Sharma BK, Kumawat C, Vafai K. Computational biomedical simulations of hybrid nanoparticles (Au-Al2O3/blood-mediated) transport in a stenosed and aneurysmal curved artery with heat and mass transfer: Hematocrit dependent viscosity approach. Chem Phys Lett. 2022;800:139666. https://doi.org/10.1007/978-3-030-99792-2_34

14. Thamizharasan T, Reddy AS. Pulsating hydromagnetic flow of au-blood Jeffrey nanofluid in a channel with joule heating and viscous dissipation. Nanoscience and Technology: Nanosci Technol An Int J. 2022;13(2):1-13. https://doi.org/10.1615/NanoSciTechnolIntJ.2022039247
15. Tripathi B, Sharma BK, Sharma M. Modeling and analysis of MHD two-phase blood flow through a stenosed artery having temperature-dependent viscosity. Eur Phys J Plus. 2019;134(1):466. https://doi.org/10.1140/epjp/i2019-12813-9
16. Tripathi B, Sharma BK. Influence of heat and mass transfer on two-phase blood flow with joule heating and variable viscosity in the presence of variable magnetic field. Int J Comput Methods. 2020;17(03):1850139. https://doi.org/10.1142/S0219876218501396
17. Tripathi B, Sharma BK. Two-phase analysis of blood flow through a stenosed artery with the effects of chemical reaction and radiation. Ric di Mat. 2021;3(2):1-7.
18. Hayat T, Hussain Z, Alsaedi A, Hobiny A. Computational analysis for velocity slip and diffusion species with carbon nanotubes. Results Phys. 2017;7:3049–58. https://doi.org/10.1016/j.rinp.2017.07.070
19. Hafeez MB, Krawczuk M, Shahzad H. An overview of heat transfer enhancement based upon nanoparticles influenced by induced magnetic field with slip condition via finite element strategy. acta Mech Autom. 2022;16(3):200–206. https://doi.org/10.2478/ama-2022-0024
20. Eckert ERG, Drake Jr RM. Analysis of heat and mass transfer. MC Graw Hill Publishing;1974. https://doi.org/10.1016/j.rinp.2017.07.070
21. Sharma BK, Yadav K, Mishra NK, Chaudhary RC. Soret and Dufour effects on unsteady MHD mixed convection flow past a radiative vertical porous plate embedded in a porous medium with chemical reaction. 2012; 3(7):717-723. https://doi.org/ 10.4236/am.2012.37105
22. Sharma BK, Gupta S, Krishna VV, Bhargavi RJ. Soret and Dufour effects on an unsteady MHD mixed convective flow past an infinite vertical plate with Ohmic dissipation and heat source. Afrika Mat. 2014;25(3):799–821.
23. Xiao X, Wu Z-C, Chou K-C. A multi-label classifier for predicting the subcellular localization of gram-negative bacterial proteins with both single and multiple sites. PLoS One. 2011;6(6):e20592.
24. Siddique I, Nadeem M, Awrejcewicz J, Pawłowski W. Soret and Dufour effects on unsteady MHD second-grade nanofluid flow across an exponentially stretching surface. Sci Rep. 2022;12(1):11811. https://doi.org/10.1038/s41598-022-16173-8
25. Nowar K. Peristaltic flow of a nanofluid under the effect of Hall current and porous medium. Math Probl Eng. 2014;2014:1-15. https://doi.org/10.1155/2014/389581
26. Nadeem S, Ijaz S, Akbar NS. Nanoparticle analysis for blood flow of Prandtl fluid model with stenosis. Int Nano Lett. 2013;3(1):1–13. https://doi.org/10.1186/2228-5326-3-35
27. Su X, Zheng L. Hall effect on MHD flow and heat transfer of nanofluids over a stretching wedge in the presence of velocity slip and Joule heating. Cent Eur J Phys. 2013;11(12):1694–703. https://doi.org/10.2478/s11534-013-0331-0
28. Ellahi R, Rahman SU, Nadeem S. Blood flow of Jeffrey fluid in a catherized tapered artery with the suspension of nanoparticles. Phys Lett A. 2014;378(40):2973–80. https://doi.org/10.1016/j.physleta.2014.08.002
29. Ghandi R, Sharma BK, Kumawat C, Beg OA. Modeling and analysis of magnetic hybrid nanoparticle (Au-Al2O3/blood) based drug delivery through a bell-shaped occluded artery with Joule heating, viscous dissipation and variable viscosity effects. Proc Inst Mech Eng Part E J Process Mech Eng. 2022; 236(5):2024-43. https://doi.org/10.1177/09544089221080273
30. Hayat T, Hussain Z, Alsaedi A, Muhammad T. An optimal solution for magnetohydrodynamic nanofluid flow over a stretching surface with constant heat flux and zero nanoparticles flux. Neural Comput Appl. 2018;29:1555–62. https://doi.org/10.1007/s00521-016-2685-x
31. Bhandari A. Mathematical Modelling of Water-Based FeO Nanofluid Due to Rotating Disc and Comparison with Similarity Solution. acta Mech Autom. 2021;15(3):113–121. https://doi.org/10.2478/ama-2021-0016
32. Hussain Z, Hayat T, Alsaedi A, Anwar MS. Mixed convective flow of CNTs nanofluid subject to varying viscosity and reactions. Sci Rep. 2021;11(1):22838. https://doi.org/10.1038/s41598-021-02228-9
33. Hussain Z, Alshomrani AS, Muhammad T, Anwar MS. Entropy analysis in mixed convective flow of hybrid nanofluid subject to melting heat and chemical reactions. Case Stud inTherm Eng. 2022;34:101972. https://doi.org/10.1016/j.csite.2022.101972
34. Miri R, Abbassi MA, Ferhi M, Djebali R. Second Law Analysis of MHD Forced Convective Nanoliquid Flow Through a Two-Dimensional Channel. acta Mech Autom. 2022;16(4):417–431. https://doi.org/10.2478/ama-2022-0050
35. Cho HW, Hyun JM. Numerical solutions of pulsating flow and heat transfer characteristics in a pipe. Int J Heat Fluid Flow. 1990;11(4):321–330. https://doi.org/10.1016/0142-727X(90)90056-H
36. Sharma BK, Gandhi R. Combined effects of Joule heating and non-uniform heat source/sink on unsteady MHD mixed convective flow over a vertical stretching surface embedded in a Darcy-Forchheimer porous medium. Propuls Power Res. 2022;11(2):276–292. https://doi.org/10.1016/j.jppr.2022.06.001
37. Craciunescu OI, Clegg ST. Pulsatile blood flow effects on temperature distribution and heat transfer in rigid vessels. J Biomech Eng. 2001;123(5):500–505. https://doi.org/ 10.1115/1.1392318
38. Naqvi SMRS, Farooq U, Aiyashi MA, Waqas H. Comprehensive analysis of thermally radiative transport of Sisko fluid over a porous stretchable curved surface with gold nanoparticles. Int J Mod Phys B. 2022;36(03):2250028. https://doi.org/10.1142/S021797922250028X
39. Hussain Z. Heat transfer through temperature dependent viscosity hybrid nanofluid subject to homogeneous-heterogeneous reactions and melting condition: A comparative study. Phys Scr. 2020;96(1):15210. https://doi.org/10.1088/1402-4896/abc5ef
40. Sharma BK, Kumawat C, Makinde OD. Hemodynamical analysis of MHD two phase blood flow through a curved permeable artery having variable viscosity with heat and mass transfer. Biomech Model Mechanobiol. 2022;21(3):797–825. https://doi.org/10.1007/s10237-022-01561-w
41. Sharma BK, Kumawat C. Impact of temperature dependent viscosity and thermal conductivity on MHD blood flow through a stretching surface with ohmic effect and chemical reaction. Nonlinear Eng. 2021;10(1):255–271. https://doi.org/10.1515/nleng-2021-0020
42. Chakravarty S, Mandal PK. Mathematical modelling of blood flow through an overlapping arterial stenosis. Math Comput Model. 1994;19(1):59–70. https://doi.org/10.1016/0895-7177(94)90116-3
43. Chakravarty S, Mandal P. A nonlinear two-dimensional model of blood flow in an overlapping arterial stenosis subjected to body acceleration. Math Comput Model. 1996;24(1):43–58. https://doi.org/10.1016/0895-7177(96)00079-9
44. Nadeem S, Ijaz S. Theoretical analysis of metallic nanoparticles on blood flow through stenosed artery with permeable walls. Phys Lett A. 2015;379(6):542–554. https://doi.org/10.1016/j.physleta.2014.12.013
45. Datta BN. Numerical linear algebra and applications: Siam. 2010; 116. https://www.mdpi.com/journal/mathematics/special issues/ numelinear algebra
46. Sharma BK, Gandhi R, Mishra NK, Al-Mdallal QM. Entropy generation minimization of higher-order endothermic/exothermic chemical reaction with activation energy on MHD mixed convective flow over a stretching surface. Sci Rep. 2022;12(1):17688. https://doi.org/10.1038/s41598-022-22521-5
47. Sharma BK, Poonam, Chamkha AJ. Effects of heat transfer, body acceleration and hybrid nanoparticles (Au–Al2O3) on MHD blood flow through a curved artery with stenosis and aneurysm using hematocrit-dependent viscosity. Waves in Random and Complex Media. 2022;2(3):1–31. https://doi.org/10.1080/17455030.2022.2125597

48. Ali U, Irfan M, Rehman KU, Alqahtani AS, Malik MY, Shatanawi W. On the Cattaneo–Christov heat flux theory for mixed convection flow due to the rotating disk with slip effects. Waves in Random and Complex Media. 2022;4(3):1–15.
49. Sharma BK, Khanduri U, Mishra NK, Mekheimer KS. Combined effect of thermophoresis and Brownian motion on MHD mixed convective flow over an inclined stretching surface with radiation and chemical reaction. Int J Mod Phys B. 2022;2350095. https://doi.org/10.1142/S0217979223500959
50. Gandhi R, Sharma BK, Makinde OD. Entropy analysis for MHD blood flow of hybrid nanoparticles (Au–Al2O3/blood) of different shapes through an irregular stenosed permeable walled artery under periodic body acceleration: Hemodynamical applications. ZAMM-Journal Appl Math Mech für Angew Math und Mech. 2022;e202100532. https://doi.org/10.1002/zamm.202100532
51. Sharma BK, Kumar A, Gandhi R, Bhatti MM. Exponential space and thermal-dependent heat source effects on electro-magneto-hydrodynamic Jeffrey fluid flow over a vertical stretching surface. Int J Mod Phys B. 2022;36(30):2250220. https://doi.org/10.1142/S0217979222502204

Nomenclature: *u* represents the velocity component in the z-direction, *v* the velocity component in the r-direction, *R(z)* the radius of the artery in the obstructed region, R_0 the normal artery's radius, ρ_f the effective density of the nanofluid, α_T the volume expansion coefficient with temperature, α_c the volumetric expansion coefficient with concentration, C_{pf} the specific heat of the fluid at constant pressure, *σ* the electrical conductivity, T_m the mean fluid temperature, D_m the coefficient of mass diffusivity, the K_T the thermal-diffusion ratio, k_T the thermal conductivity of the fluid, *μ* the viscosity, Pr Prandtl number, Ec Eckert number, Sr Soret number, Sc Schmidt number, *ω* viscosity parameter, Gr Grashof number, Cr the local concentration number, Df Dufour number and Br Brinkman number.

Nidhish Kumar MISHRA: https://orcid.org/0000-0003-4502-261X

DOI 10.2478/ama-2023-0029

SYNTHESIS OF PNEUMATIC SYSTEMS IN THE CONTROL OF THE TRANSPORT LINE OF ROLLING ELEMENTS

Adam SZCZEŚNIAK, Zbigniew SZCZEŚNIAK*, Leszek CEDRO****

*Faculty of Electrical Engineering, Automation and Computer Science, Kielce University of Technology,
Al. Tysiąclecia Państwa Polskiego 7, 25-314 Kielce, Poland
**Faculty of Mechatronics and Mechanical Engineering, Kielce University of Technology,
Al. Tysiąclecia Państwa Polskiego 7,25-314 Kielce, Poland

z.szczesniak@tu.kielce.pl, adam_szczesniak@o2.pl, lcedro@tu.kielce.pl

received 6 December 2022, revised 18 January 2023, accepted 23 January 2023

Abstract: This paper presents the synthesis of a pneumatic control system for a selected configuration of the transport path for the delivery of rolling elements to spiral storage in inter-operational transport. The sequential control system sets the state of the manifolds to ensure a flow of workpieces to serve the subsequent storage. The essential module of the control system is the memory block. It is developed based on a storage filling sequence graph. The filling level of the storages can be monitored in one or two points using sensors. The rolling element displacement control sensors work together with appropriately designed systems to execute the delay of the rising and falling edge input signal. By using a two-level control of the filling level of the storages, it is possible to control the emptying status of the storages as a function of the technological time of removal of the items from the storage between the two control points. Control systems were synthesised and verified using Festo's FluidSim computer programme.

Key words: pneumatic systems, delay execution systems, synthesis, verification of sequential systems

1. INTRODUCTION

In high-volume production of rolling elements of all kinds, automation of the transport route between subsequent production stages is simply necessary. In this type of production, special machine tools with a high degree of automation are used to tool the components. Activities such as supplying components to workstations, applying products for machining, etc., result in increased worker utilisation, and these can be performed much more quickly by a unit of equipment configured in a mass handling system. The automation of the inter-operational transport route increases the productivity of machine tools almost to the maximum and reduces the costs associated with the transport of workpieces. The transport system must also be capable of being tuned, for example, to produce components of a different type, and thus it requires the control equipment to use universal modules that are independent of the transport path configuration.

There are many tasks involved in designing, among which one of the most significant is the analysis and synthesis of the schematic diagram of the device [1,2].

In sequential circuits [3], the current state of outputs depends not only on the current state of inputs but also on the sequence of previous input states, while in combinational circuits [4] the current state of the outputs depends only on the current state of the inputs.

In sequential asynchronous systems, the clock signal does not occur. The input signals directly affect the internal state of the system at all times. Thus, each input change causes an immediate (taking into account the signal propagation time through the system) reaction of the system [5].

The lack of a clock signal makes the synthesis of asynchronous circuits, in general, more difficult than the synthesis of synchronous circuits [6].

The synthesis of sequential circuits using the conventional method of transition and output tables is simple if the number of inputs and the number of internal states are not large. However, for systems with more than three inputs and eight internal states, the burden of using the transition and output tables increases significantly, the synthesis algorithms become more complicated (for example, the excitation functions then depend on seven arguments) and the chance of obtaining an optimal solution decreases [7].

In order to meet the expectations of system designers regarding the minimisation of the mathematical apparatus in the analysis of systems, an algorithmic approach to the synthesis of sequential systems was presented [8]. The programming language that has gained the greatest popularity among PLC programmers is the Ladder Diagram language [9]. The reason for this is that it is easy to understand due to its similarity to contact-relay diagrams [10]. Ladder logic also allows the user to perform more complex operations such as arithmetic and time operations. The control scheme in this language is in the form of symbols placed in circuits resembling the ladder of a relay scheme. This language allows the user to build control systems based on logical dependencies resulting from Boolean algebra [7].

The material presented in the article is a continuation of the research issues discussed in an earlier paper in the literature from the authors of the present study [8], in which the synthesis of the sequential electropneumatic system with the use of logical elements was presented, and discussed in other studies in the literature from the same authors [11, 12], in which the problems of the synthesis of sequential electropneumatic systems were presented.

The control of the transport routes is carried out through distributors consisting of actuators working together with two-state valves [13,14].

In the case of continuous actuator position control systems, proportional valves are used [15,16], while digital position transducers [17] are used to measure actuator position, enabling precise actuator position control.

The literature [18,19] provides selected examples of control system design using pneumatic components and devices. In control systems, the positioning accuracy of actuator elements is of vital importance [20,21]. An analysis of accuracy in signal processing is presented in the literature [22,23].

The general configuration of the transport line for delivering rolling elements to the storages is shown in Fig. 1.

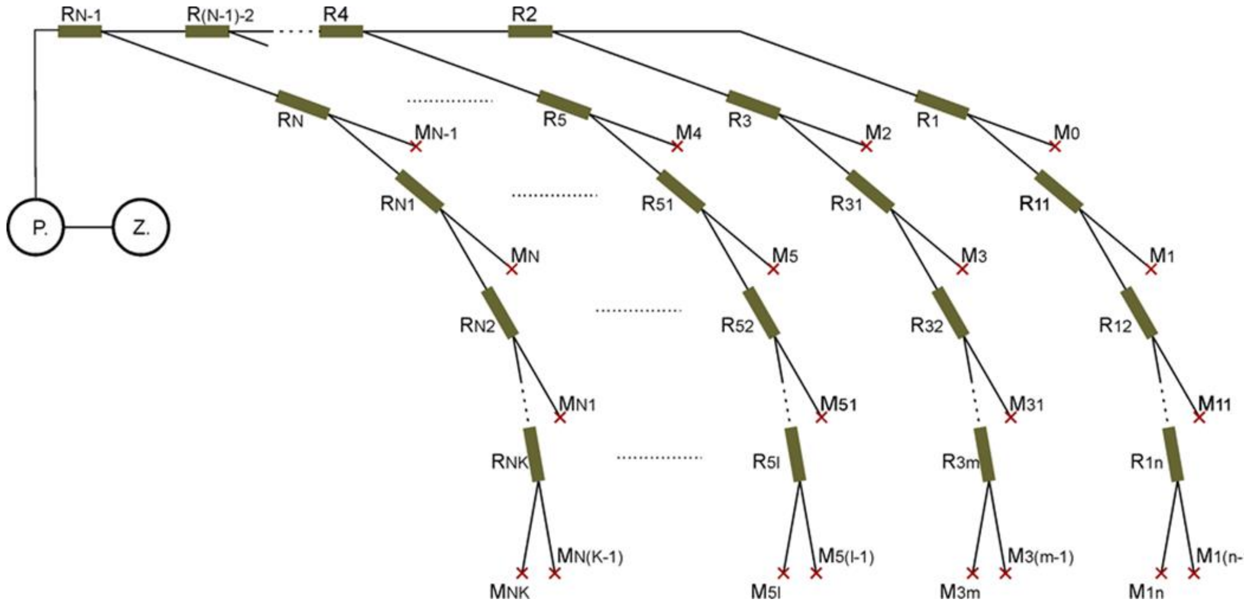


Fig. 1. Generalised configuration of the transport line section: Z – bin, P – vertical lift, R – distributors of transport routes, M – spiral storages

The transport line can be configured in any way by means of a suitable connection of the R distributors. The rolling elements are lifted from the Z bin to a certain height by means of the P vertical lift and directed via the R distributors to the corresponding M spiral storage using the force of gravity and the properties of the rolling elements.

In the general case, the number of M storages is greater than the number of R distributors by one, i.e. for R = X, M = X + 1.

This paper presents a method for the synthesis of an automatic control system for the distribution of rolling elements, using the example of a transport line section with M = 8 storages and R = 7 distributors. The configuration shown in Fig. 2 is an outtake of a generalised transport line section configuration.

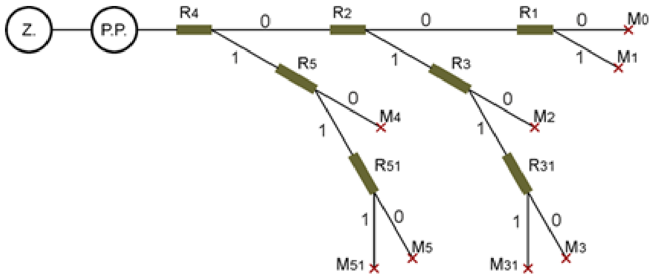


Fig. 2. Configuration of the analysed transport line M = 8, R = 7

The choice of transport route depends on the signals coming from the storage fill level sensors. Tab. 1 shows the status of setting up the R distributors to provide a flow of workpieces to the selected storage.

Tab. 1. The state of the R distributors to ensure the flow of workpieces to relevant M storages

	R1	R2	R3	R31	R4	R5	R51
M0	0	0	-	-	0	-	-
M1	1	0	-	-	0	-	-
M2	-	1	0	-	0	-	-
M3	-	1	1	0	0	-	-
M31	-	1	1	1	0	-	-
M4	-	-	-	-	1	0	-
M5	-	-	-	-	1	1	0
M51	-	-	-	-	1	1	1

For example, in order to provide flow to the M31 storage, distributors R2, R3 and R31 must be in state 1, i.e. on, and distributor R4 in state 0 (off), while the other distributors may be in any state.

The filling level of the storages can be monitored in one or two points using sensors. Sensors for the control of moving rolling elements must work in conjunction with appropriate delay execution systems.

This article presents the pneumatic measuring and control systems developed for use in the inter-operational transport of rolling elements. In pneumatic systems, divide valves, check and throttle valves and pneumatic capacities, among others, are used as input signal time delay systems [24]. Through various combinations of these elements, a delay in the appearance or disappearance of the output signal in relation to the input signal is obtained. By routing the input signal to the check and throttle valve, the output of which is connected to the pneumatic capacity and the valve control input, a delay in switching the valve on or off is achieved, depending on the direction in which the check and throttle valve is switched on. The time delays are selected by the parameters of pneumatic capacity and damping of the check and throttle valve. By connecting opposite-acting check and throttle valves in series, the output of which is connected to a pneumatic capacity and a valve control input, a delay in valve switching on and off is achieved, reflected by the delay in the appearance and disappearance of the output signal relative to the input signal. Transducer (time-dependent) systems built based on a series connection of check and throttle valves do not have the ability to control the duration of input signals with a set time in the case of input signal sequences with a duration less than the set time, which is a major drawback. The above drawback was eliminated by appropriately designed circuits for executing the delay of the rising and falling edge input signal, working in conjunction with sensors to control moving components. Pneumatic jet sensors are used as control sensors [25]. They work on the principle of sensing the reflected air stream from an aperture (in this case, a rolling element) or interruption of the air stream by a moving workpiece. The filling level of the storages can be monitored in one or two points using sensors. By using a two-level control of the filling level of the storages, it is possible to control the emptying status of the storages as a function of the technological time of removal of the items from the storage between the two control points.

The developed system is dedicated to applications with high requirements for reliable operation in adverse environments (dust, high temperature, magnetic field, etc.).

2. SINGLE-POINT SPIRAL STORAGE CONTROL SYSTEM FOR ROLLING ELEMENTS

Fig. 3 shows the system developed to work with a pneumatic sensor for single-point measurement of the fill level of a rolling element storage.

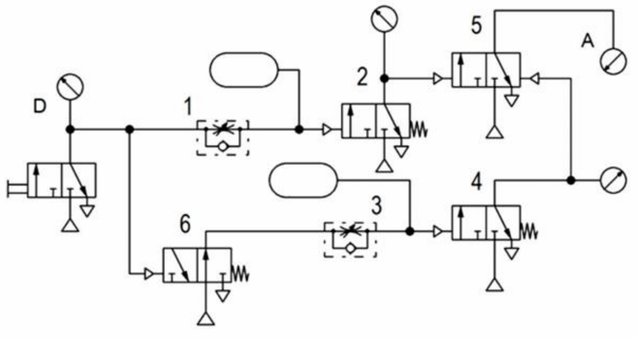


Fig. 3. Single-point storage state control system

The system was set up using diverter valves, throttle valves and pneumatic capacities. The delaying system diagram shows all these elements, regardless of whether they form a structural whole or are independent elements connected by wires. Through various combinations of these elements, a delay in the appearance or disappearance of the output signal in relation to the input signal was obtained. This is important, since moving rolling elements in the control sensor zone cause the generation of a brief pulse of change in the signal state of this sensor.

The system does not react to short pulses, shorter than the set delay time of the rising edge of the input signal (signal head), and shorter than the set delay time of the falling edge of the input signal (signal rear). The head delay of the D signal of the control sensor was obtained at the output of valve 2 of the system, while the rear delay of the signal D of the control sensor was obtained at the output of valve 4 of the system. The signal at the output of valve 2 causes output valve 5 to be switched on, generating an A signal at high level (logical 1) – that is, it reflects the state of the presence of rolling elements in the storage. The high state of the signal at valve output 4 switches off output valve 5, changing signal A to a low level (logical 0), reflecting the absence of rolling elements in the zone of the storage control sensor.

The cyclogram of the system operation is shown in Fig. 4, which explains the operation of the system in Fig. 3.

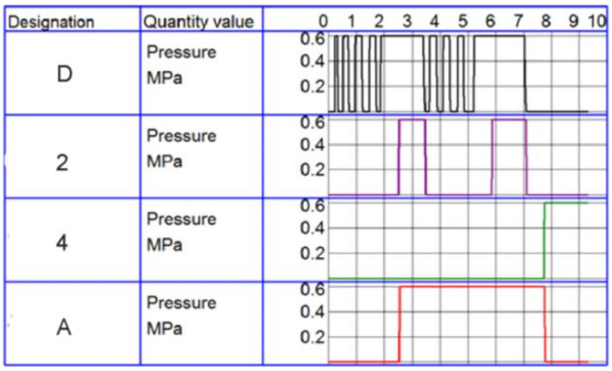


Fig. 4. Cyclogram of the operation of the single-point storage state control system

It should be emphasised that the rise and fall delay time of the input signal is selected in the process depending on the speed of movement of the components in the rolling element control sensor zone.

3. TWO-POINT SPIRAL STORAGE CONTROL SYSTEM FOR ROLLING ELEMENTS

Fig. 5 shows the developed system for controlling the fill level of the storage by means of two measuring sensors, i.e. sensor D for the lower position of the rings in the storage and sensor G for the upper position of the rings in the spiral storage. A rising and falling edge input delay system works with each sensor. These systems are identical to the systems discussed in item 2, and shown in Figs. 3 and 4.

The essence of the developed system consists in linking two sensors D and G cooperating with the head and rear delay systems of the input signal, with a corresponding memory system.

The arrangement of the combination of the two valves 7 and 8 and the sum element 8 constitute a memory system C, in which the output signal C reflects the state of the spiral storage. When the storage is unfilled, the output signal of the C memory system is in a low state (logic state 0). When the storage is filled to the lower D level, the output signal of the A delay system is in a high state (logic state 1), overdriving valve 8 by feeding the low level C output signal of the memory system to the 9 sum valve.

When the storage is filled to the upper G level, the output signal of the B delay system is in a high state, and it is fed to the 9 sum element of the memory system.

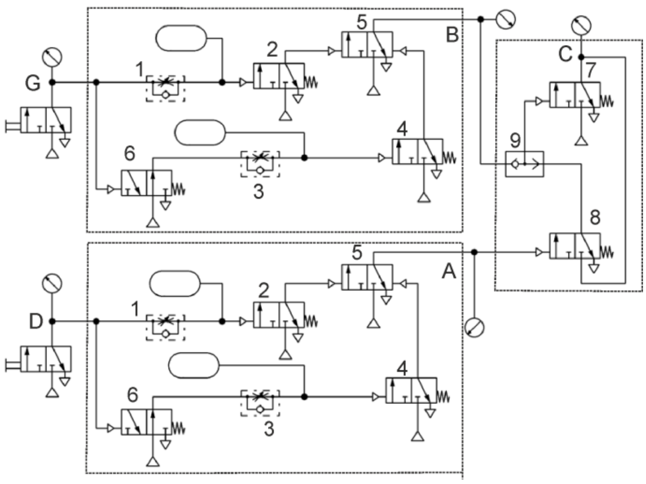


Fig. 5. Two-point spiral storage control system for rolling elements

This switches valve 7 and applies high pressure to output C of the memory system. This state is maintained by valve 8 controlled by signal A and 9 sum system, providing feedback to the memory system. When the storage is emptied in the process, the first step is to change the state of the upper G level sensor of the storage (the B-output signal of valve 5 changes from high to low). The state of the C signal of the memory chip is still high. Further emptying of the storage below the level of the lower storage D level sensor changes the signal of the delay system A to low and thus changes the state of valve 8 controlled by this signal.

There is then a change in the state of the memory system, thereby changing the C signal to a low state, which is signalled as a storage requirement for elements.

The exact operation of the designed system is illustrated by the cyclogram shown in Fig. 6.

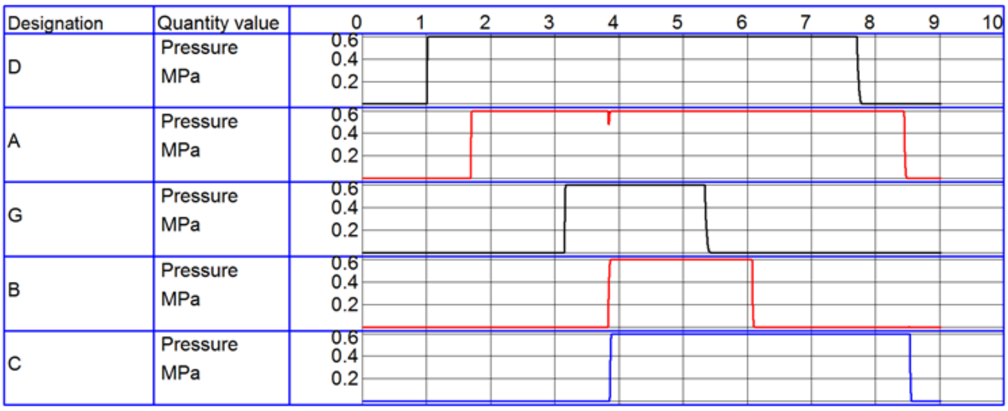


Fig. 6. Cyclogram of the operation of the double-point rolling element spiral storage state control system

It should be emphasised that the time between the change in state of signal B from high to low and the change in state of signal A from high to low is the storage delay time known as the technological process adaptation time.

4. MINIMISED TWO-POINT SPIRAL STORAGE CONTROL SYSTEM FOR ROLLING ELEMENTS

Fig. 7 shows the developed system for controlling the fill level of the storage by means of two measuring sensors, i.e. sensor D for the lower position of the rings in the storage and sensor G for the upper position of the rings in the spiral storage, which work with a single system for delaying the input signal of the rising and falling edge and the system implementing the logic function (¬D+A)¬G. The delay system for the rising and falling edge input signal is identical to the system discussed in item 2, which is shown in Fig. 3. The system implementing the logic function (¬D+A)¬G is set up using appropriately connected elements 7, 8, 9 and 10. The output of this system is connected to the input of the rising and falling edge delay system.

The exact operation of the designed minimised two-point spiral storage control system for rolling elements is illustrated by the cyclogram shown in Fig. 8.

It should be emphasised that in the systems in Figs. 3, 5 and 7, time delay valves, type VZ-3-PK-3 by Festo [24], were used. For the analysis of the developed systems, this valve was adopted because it meets the requirements of setting the time delay range (0.25–5.0 s).

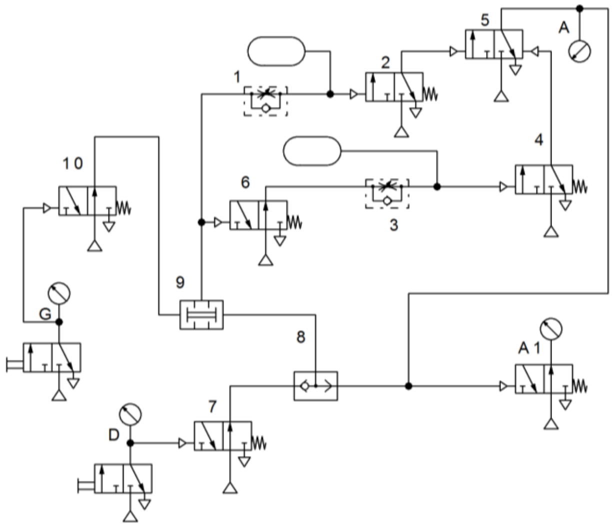


Fig. 7. Minimised two-point spiral storage control system for rolling elements

DOI 10.2478/ama-2023-0029

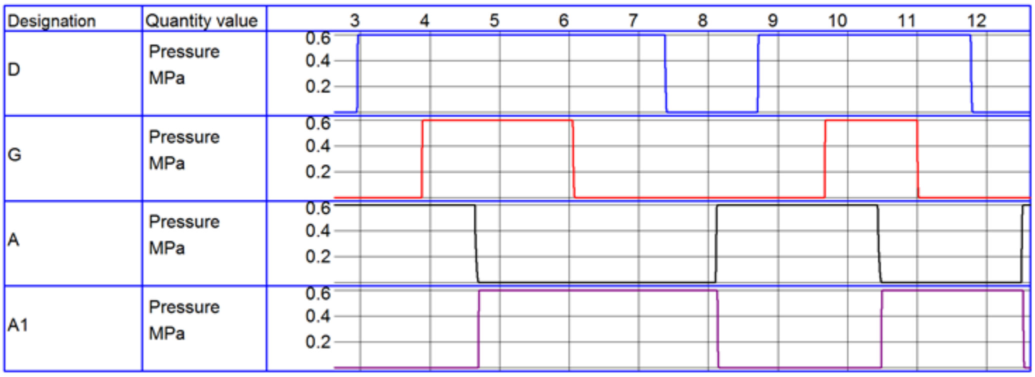


Fig. 8. Cyclogram of the operation of the two-point spiral storage control system for rolling elements

5. SYNTHESIS OF A CONTROL SYSTEM FOR THE SUPPLY OF ROLLING ELEMENTS TO A SPIRAL STORAGE

The essential module of the control system is the memory block. It is developed based on a storage filling sequence graph. In the graph, the order in which the storages are filled is assumed from M0 to M51, as shown in Fig. 9.

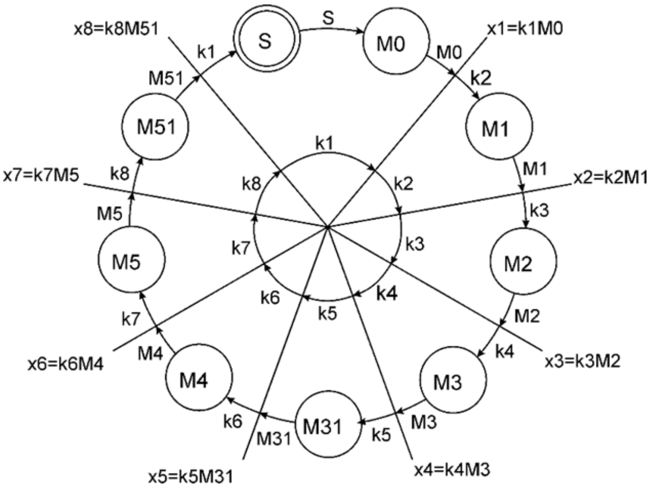


Fig. 9. Graph of storage filling sequence

The graph was divided radially by assigning a k memory state to the individual M storages. On the division line, the signals xn specifying the transition to the next state kn were defined (filling a storage in kn state allows the transition to filling the next unfilled storage).

The graph indicates the memory state that the sequential storage filling system is in when filling the detail storage. This is illustrated in Tab. 2. The memory is in a high state for the filled storage. The selection of the next storage is possible in the next memory high state (logical state 1), i.e. in the wandering memory high state.

Based on the graph (Fig. 9) and the status of the distributors, a sequential control system for filling the storages was developed to ensure the flow rate of the workpieces to the corresponding M storages (Tab. 1), as shown in Fig. 10.

The system consists of three main elements:

– M0–M51 single- or two-point storage filling control systems in accordance with Figs. 3, 5 or 7;
– memory system k1–k8; and
– control system for the flow divider elements R1–R51 for the storages M0–M51.

Tab. 2. Sequence of storage filling depending on memory status

	k1	k2	k3	k4	k5	k6	k7	k8
M0	1	0	0	0	0	0	0	0
M1	0	1	0	0	0	0	0	0
M2	0	0	1	0	0	0	0	0
M3	0	0	0	1	0	0	0	0
M31	0	0	0	0	1	0	0	0
M4	0	0	0	0	0	1	0	0
M5	0	0	0	0	0	0	1	0
M51	0	0	0	0	0	0	0	1

Pneumatic memory has been implemented on 4/2 valves (four-connection two-position valves) controlled by x1–x8 signals, with each subsequent memory state erasing the previous one. The status of the storages is reflected by the signals M0–M51 (the signals represent the status of the control sensors together with the delay systems). The state of the flow distributors is controlled by the sum elements as shown in Tab. 1. Distributors not involved in the flow setting for filling the currently selected storage are in the setting from the previous flow setting – it is uneconomical to overdrive them to the zero position. In addition to selecting successive memory states, the storage filling control system is responsible for switching the vertical feeder on or off and stopping the cycle in state k1, according to the relation S = k1 [(M0) + (M1) + (M2) + (M3) + (M31) + (M4) + (M5) + (M51)].

The system shown in Fig. 11 is undergoing simulation in the filling state of storage M5, with distributors R4 and R5 on (state 1 setting) and distributor R51 off (logic state 0), and in memory state k7.

Fig. 12 shows an example of a simulation cyclogram of a sequential control system for filling spiral storages for rolling elements. In the initial state, storages M1, M3, M31 and M4 are filled.

In the simulation presented, the following phases of the system should be distinguished:

– in the first instance, the rolling elements are delivered to the M0 storage, distributors R1, R2 and R4 are in the off state (logic state 0) and the vertical feeder is switched on;

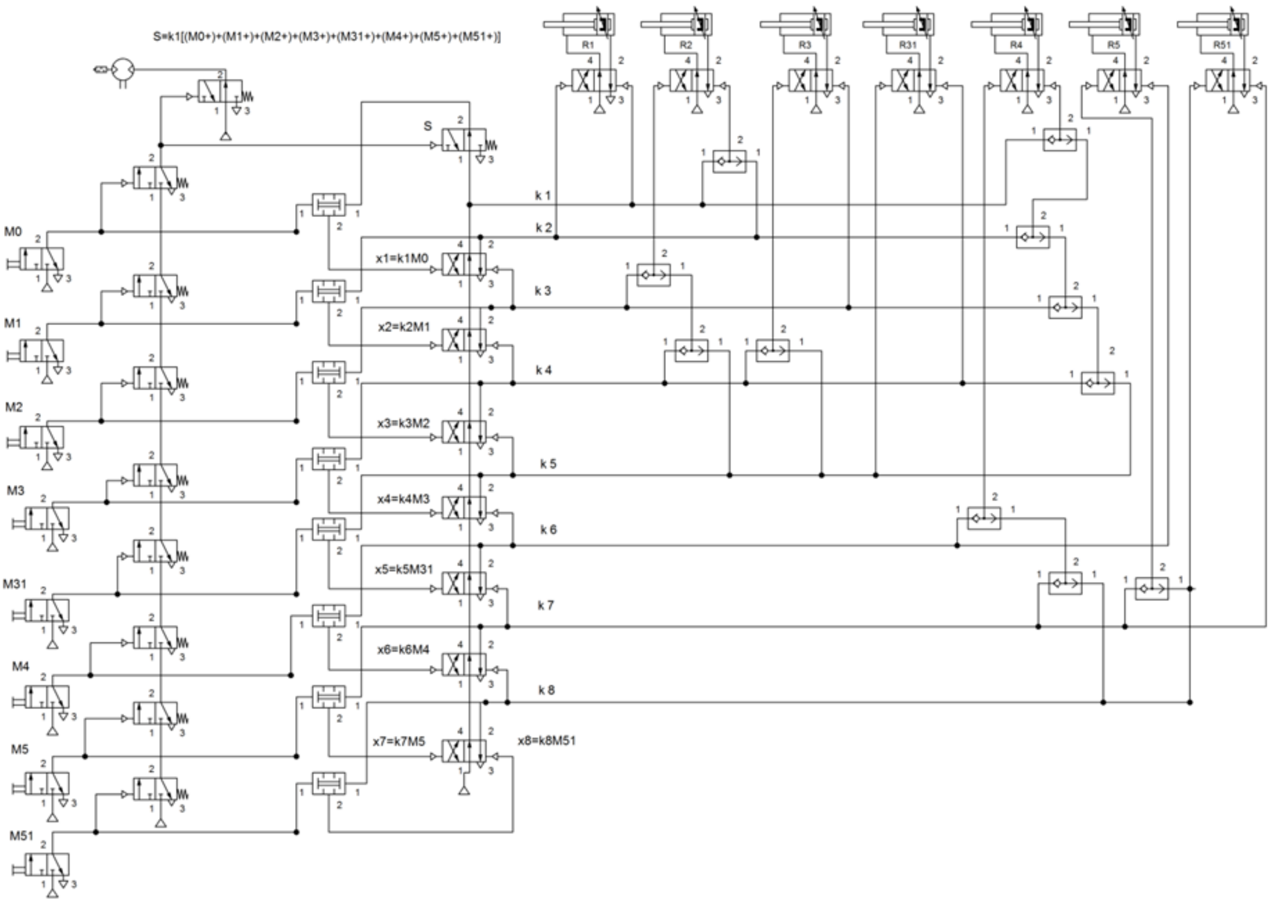


Fig. 10. Sequential control system for filling spiral storages for rolling elements

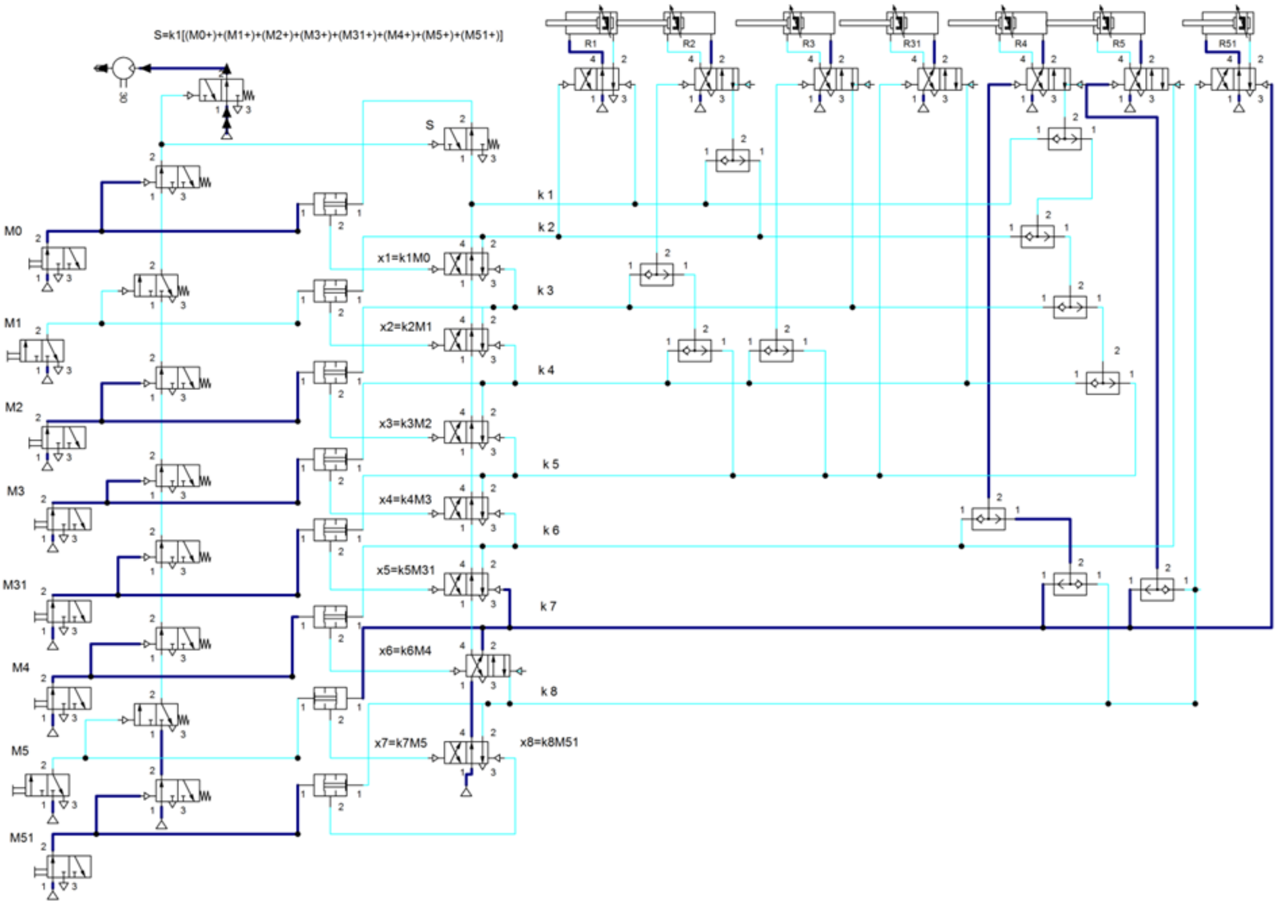


Fig. 11. Simulation of the sequential control system for filling spiral storages for rolling elements

DOI 10.2478/ama-2023-0029

- once storage M0 is filled (logic state 1), storage M2 is filled, with distributor R2 switched on (setting to state 1) and distributors R3 and R4 switched off (logical state 0);
- once storage M2 is filled (logic state 1), storage M5 is filled, with distributors R4 and R5 switched on (setting to state 1) and distributor R51 switched off (logic state 0);
- when storage M5 is filled (logic state 1), storage M51 is filled, with distributors R4, R5 and R51 switched on (setting to state 1);
- when the M51 storage is filled, the rolling element feeder is switched off;
- the M31 storage is then declared for filling, distributors R2, R3 and R31 are in the on state (logic state 1), manifold R4 is off (logic state 0) and the rolling element feeder is switched on;
- once storage M31 is filled (logic state 1), storage M5 is filled, with distributors R4 and R5 switched on (setting to state 1) and distributor R51 switched off (logic state 0); and
- once storage M5 (logic state 1) is filled, the empty storage is filled again in sequential order from M0 to M51, etc. – the cycle repeats.

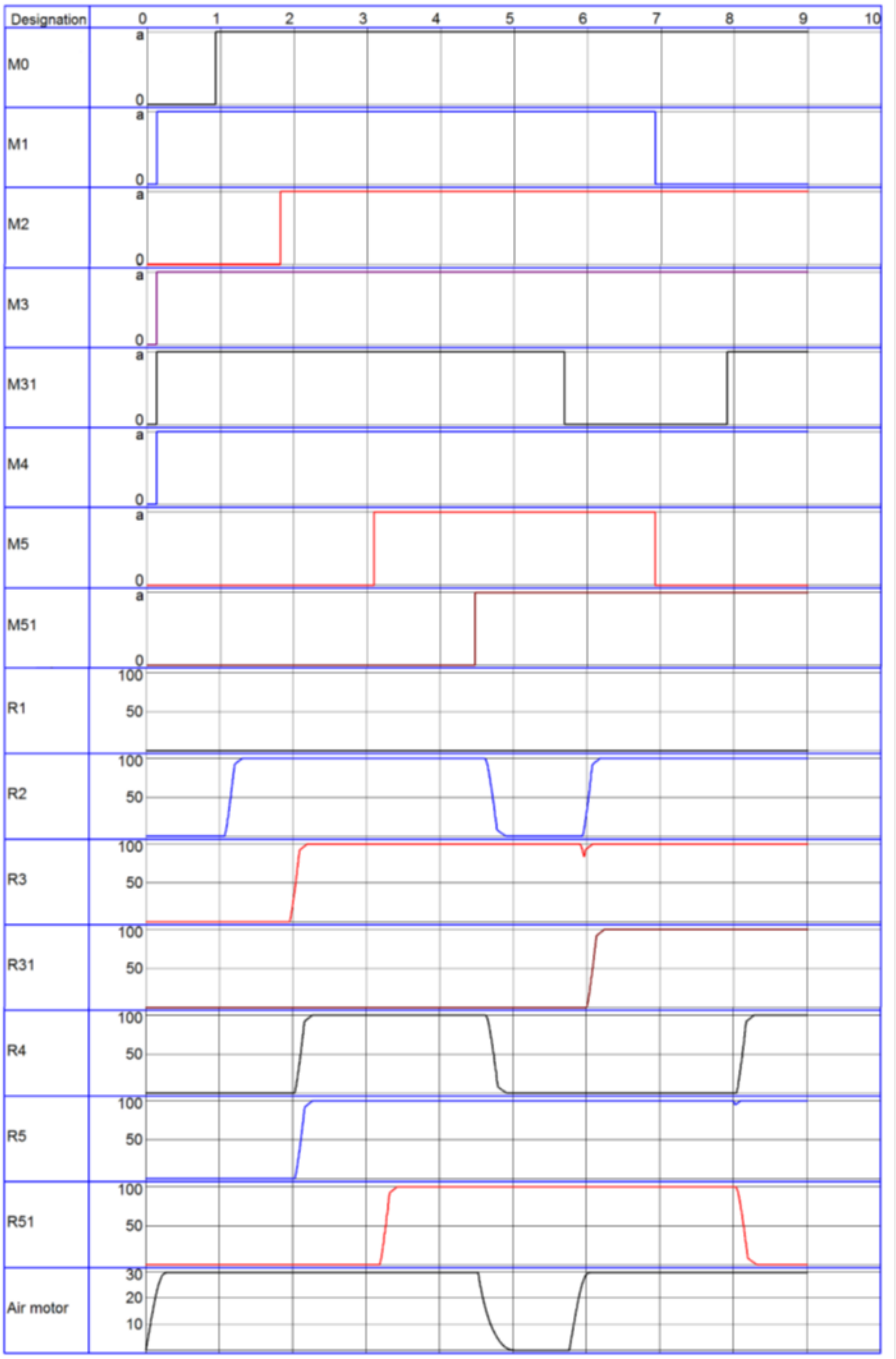


Fig. 12. Cyclogram of the sequential control system for filling spiral storages for rolling elements

It should be emphasised that in order to initially characterise the mass service system, three basic parameters must be identified: the intensity of the notification stream, the intensity of the service process and the rule of the queue [26]. The average intensity of the notification stream (λ) is defined as:

$\lambda = 1/t_\lambda$

where t_λ represents average time interval between successive notifications flowing into the system in the examined period.

The average service stream intensity (μ) is defined as:

μ = 1/t_μ

where t_μ represents the average time of handling a single notification in the analysed period.

If the stream of requests has a Poisson distribution, the intensity of service is described by an exponential distribution and first in first out (FIFO) discipline is maintained in the queue, then the functioning of such a system [27] can be expressed by the system utilisation rate (ρ), also called the Erlang constant:

$\rho = \lambda/s\mu$

where s represents the number of service desks designed to be serviced.

If ρ >1 at $t \rightarrow \infty$, the queue grows to infinity (system is unstable), while when ρ <1 the queue problem does not exist (system is stable). When $\rho = 1$ (indicating that the system is on the verge of stability) and $\rho >1$ (indicating that the system is unstable), then the system's operation would not be not very practical under the above assumptions. The system stability condition is: $0 \leq \rho < 1$. The target reduction of the queue to the zero level ($Tk \rightarrow 0$) at a certain intensity of the call stream ($\lambda = const$) can be achieved in two ways, either by reducing the time ($\mu \rightarrow min$), or by increasing the number of service desks ($S \rightarrow max$). In the case of single-station service, the appropriate service capacity should be selected (in the system considered, the capacity of the vertical lift of the rolling elements), which in turn can reduce the time of magazine service.

6. CONCLUSIONS

Testing of the designed pneumatic control system allows us to conclude that:

– by synthesising the system, according to the procedure outlined, it is possible to quickly obtain a control system for any configuration of the transport route for the delivery of components to storages in inter-operational transport;

– by using a two-level control of the filling level of the storages, it is possible to control the emptying status of the storages as a function of the technological time of removal of the items from the storage between the two control points;

– by using a system for the execution of the rise and fall delay of the input signal, short-term states of the presence and absence of an element in the sensor's area of operation are eliminated, which makes it possible to determine an unambiguous state of the presence (absence) of an element in the sensor's area of operation; and

– the time delays of the sensor cooperation system are set according to the speed of movement of the elements in the sensor control zone.

– The designed solution of pneumatic systems is dedicated to the control of process lines in intermediate transport in the production of rolling bearing components, especially inner and outer bearing rings.

– The designed pneumatic system for the execution of the head and rear delays of the input signal with adaptive technological time has been applied for a patent to the Patent Office of the Republic of Poland.

Further research will aim to develop methods and principles for creating control systems for transport routes for the delivery of components to storages in inter-operational transport using electronic systems and PLC programming ladder diagrams.

REFERENCES

1. Mychuda Z, Mychuda L, Antoniv U, Szcześniak A. Logarithmic ADC with accumulation of charge and impulse feedback – construction, principle of operation and dynamic properties. International Journal of Electronics and Telecommunications. 2021 Dec 1;67(4):699–704. doi: 10.24425/ijet.2021.137865
2. Szcześniak A, Mychuda Z. Analiza prądów upływu logarytmicznego przetwornika analogowo-cyfrowego z sukcesywną aproksymacją. Przegląd Elektrotechniczny. 2012;R. 88, nr 5a:247–50.
3. El-Maleh A. A Note on Moore Model for Sequential Circuits. 2016. https://www.researchgate.net/publication/305268049_A_Note_on_Moore_Model_for_Sequential_Circuits
4. Horowitz P, Hill W. The art of electronics. Cambridge university press Cambridge; 2002.
5. Chhillar K, Dahiya S. Design of Sequential Circuits with Timing Analysis and Considerations. Int J Eng Sci Comput. 2017;7: 808–11809.
6. Widmer NS, Moss GL, Tocci RJ. Digital systems: principles and applications. Twelfth edition. Boston: Pearson; 2017.
7. Gorzałczany M.B. Układy Cyfrowe—Metody Syntezy. Tom II: Układy Sekwencyjne, Układy Mikroprogramowane. Kielce, Poland: Wydawnictwo Politechniki Świętokrzyskiej; 2003.
8. Szcześniak Z, Szcześniak A. Projektowanie układów sterowania dla automatyzacji procesów technologicznych. Kielce, Poland: Wydawnictwo Politechniki Świętokrzyskiej; 2015.
9. Borden TR, Cox RA, Cox RA. Technician's guide to programmable controllers. 6th ed. Clifton Park, NY: Delmar, Cengage Learning; 2013.
10. Fernandez P, del Carpio C, Rocca E, Vinces L. An Automatic Control System Using the S7-1200 Programmable Logic Controller for the Ethanol Rectification Process. In: 2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON). Lima: IEEE; 2018 p. 1–4. Available from: https://ieeexplore.ieee.org/document/8526382/
11. Szcześniak A, Szcześniak Z. Algorithmic Method for the Design of Sequential Circuits with the Use of Logic Elements. Applied Sciences. 2021 Nov 23;11(23):11100. doi: 10.3390/app112311100
12. Szcześniak A, Szcześniak Z. Fast Designing Ladder Diagram of Programmable Logic Controller for a technological process. International Journal of Electronics and Telecommunications. 2022 Nov 30;68(4):709–14. doi: 10.24425/ijet. 022.141289
13. Phan VD, Vo CP, Dao HV, Ahn KK. Actuator Fault-Tolerant Control for an Electro-Hydraulic Actuator Using Time Delay Estimation and Feedback Linearization. IEEE Access. 2021;9:107111–23. doi: 10.1109/ACCESS.2021.3101038
14. Herbuś K, Ociepka P. Verification of operation of the actuator control system using the integration the B&R Automation Studio software with a virtual model of the actuator system. IOP Conf Ser: Mater Sci Eng. 2017 Aug;227:012056. doi:10.1088/1757-899X/227/1/012056
15. Acuña-Bravo W, Canuto E, Agostani M, Bonadei M. Proportional electro-hydraulic valves: An Embedded Model Control solution. Control Engineering Practice. 2017 May;62:22–35. doi: 10.1016/j.conengprac.2017.01.013

16. Wu D, Wang X, Ma Y, Wang J, Tang M, Liu Y. Research on the dynamic characteristics of water hydraulic servo valves considering the influence of steady flow force. Flow Measurement and Instrumentation. 2021 Aug 1;80:101966.
doi: 10.1016/j.flowmeasinst.2021.101966
17. Szcześniak A, Szcześniak Z. Mikroprocesorowe przetwarzanie sygnałów optoelektronicznego przetwornika położenia. Przegląd Elektrotechniczny. 2009;R. 85, nr 4:153–8.
18. Vo CP, Ahn KK. High-precision Position Control of Soft Actuator Systems - The 3rd International Workshop on Active Materials and Soft Mechatronics (AMSM2018). In KAIST, Daejeon, South Korea
19. Zhang Y, Yue H, Li K, Cai M. Analysis of Power Matching on Energy Savings of a Pneumatic Rotary Actuator Servo-Control System. Chin J Mech Eng. 2020 Dec;33(1):30. doi: 10.1186/s10033-020-00445-3
20. Szcześniak A. Analiza przetwarzania sygnałów logarytmicznego przetwornika analogowo - cyfrowego z sukcesywną aproksymacją. Kielce, Poland: Wydawnictwo Politechniki Świętokrzyskiej; 2019
21. Mychuda Z, Zhuravel I, Mychuda L, Szcześniak A, Szcześniak Z, Yelisieieva H. Mathematical Modelling of the Influence of Parasitic Capacitances of the Components of the Logarithmic Analogue-to-Digital Converter (LADC) with a Successive Approximation on Switched Capacitors for Increasing Accuracy of Conversion. Electronics. 2022 May 6;11(9):1485. doi:10.3390/electronics11091485.
22. Mychuda Z, Mychuda L, Antoniv U, Szcześniak A. Logarithmic ADC with Accumulation of Charge and Impulse Feedback : Analysis and Modeling. International Journal of Electronics and Telecommunications. 2021;Vol. 67, No. 4:705–10. doi: 10.24425/ijet.2021.137866
23. Mychuda Z, Zhuravel I, Mychuda L, Szcześniak A, Szcześniak Z. Modelling a New Multifunctional High Accuracy Analogue-to-Digital Converter with an Increased Number of Inputs. Electronics. 2022 May 25;11(11):1677. doi:10.3390/electronics11111677
24. Time delay valve VZ-3-PK-3 data sheet. Available from: https://www.festo.com/tw/en/a/download-document/datasheet/5755
25. Air gap sensors catalogue. Available from:
https://www.festo.com/pl/pl/c/produkty/automatyka-przemyslowa/czujniki/czujniki-szczelinowe-powietrzne-id_pim139/
26. Kisielewski P, Sobota Ł. Zastosowanie teorii masowej obsługi do modelowania systemów transportowych. Autobusy: technika, eksploatacja, systemy transportowe. 2016;17(6):600–4.
27. Ficoń K. Zastosowanie teorii masowej obsługi do analizy systemu zabezpieczenia logistycznego sytuacji kryzysowych. SLW. 2017 Dec 29;47(2):59–79.

Adam Szcześniak: https://orcid.org/0000-0003-2411-9279

Zbigniew Szcześniak: https://orcid.org/0000-0002-7896-3291

Leszek Cedro: https://orcid.org/0000-0002-2419-4044

SORET AND DUFOUR EFFECTS ON CHEMICALLY REACTING AND VISCOUS DISSIPATING NANOFLUID FLOWING PAST A MOVING POROUS PLATE IN THE PRESENCE OF A HEAT SOURCE/SINK

Aastha AASTHA*, **Khem CHAND***

*Department of Mathematics and Statistics, Himachal Pradesh University, Summer Hill, Shimla-171005, India

aastha2101@gmail.com, khemthakur99@gmail.com

received 18 November 2022, revised 25 January 2022, accepted 26 January 2022

Abstract: This study performed a numerical investigation of the Soret and Dufour effects on unsteady free convective chemically reacting nanofluid flowing past a vertically moving porous plate in the presence of viscous dissipation and a heat source/sink. The equations directing the flow are non-dimensionalised, modified to ordinary differential equations and emerging equations are resolved computationally by using the bvp4c function in MATLAB software. The results obtained from this analysis indicate that the resulting velocity of the nanofluid increases with increasing Grashof number, mass Grashof number and porosity parameter. An increase in the Dufour number increases the fluid temperature, whereas the concentration profile declines with the increase in the Schmidt number. It is also observed that the skin friction coefficient, Nusselt number and Sherwood number increase with increasing magnetic field parameter, Eckert number and Schmidt number, respectively. The present study reveals the impact of Soret and Dufour effects on heat and mass transfer rates in chemically reacting and viscous dissipating nanofluids.

Keywords: Soret and Dufour effects, viscous dissipation, heat source/sink, chemically reacting nanofluid, free convection

1. INTRODUCTION

The heat transfer characteristics of fluids are used in all heat transfer applications. Well-known conventional heat transfer fluids are water, ethylene glycol and propylene glycol. Numerous investigations have been carried out to improve the poor heat transfer properties of these fluids. A novel method adopted to enhance their properties is formulation of nanofluids. Nanofluids are manufactured by mixing base fluid having low thermal conductivity with solid nanoparticles having high thermal conductivity, and consequently, the new fluid, i.e., nanofluid, is formed, which has a higher heat transfer characteristic than the base fluid. In addition, nanofluids contain nanometer-sized particles, which are prepared by colloidal suspension of nanoparticles in heat transfer fluids such as water, oil, diesel, ethylene and glycol. Nanoparticles used in nanofluids are generally made of metals, oxides, carbides or carbon nanotubes. Nanofluids have enhanced thermophysical properties like thermal conductivity, specific heat, viscosity and convective heat transfer in comparison to base fluids. These thermophysical properties of nanofluids are still under exploration and needs to be further elaborated. The thermophysical properties of nanofluids and the elements that can reform these properties were investigated by Gupta et al. [1]. The found that the main factors influencing these properties are shape, size, material, concentration and temperature of nanoparticles, together with the base fluid. Nanofluids have a wide range of advantages such as high specific surface area and thus more heat transfer surface between the particles and fluids, high dispersion stability with predominant Brownian motion of particles, reduced pumping power as compared to pure liquid to achieve equivalent heat transfer intensification, and reduced particle clogging as compared to conventional slurries, thus promoting miniaturisation. The characteristic features of high thermal conductivities and heat transfer coefficients in comparison to those of conventional fluids make nanofluids suitable for the subsequent generation of flow and heat transfer fluids. Nanofluids have attracted the attention of many physicists, chemists and engineers around the world. Turkyilmazoglu and Pop [2] analysed the heat and mass transfer characteristics of some nanofluids flowing past a vertical infinite flat plate and the radiation effect for two distinct types of thermal boundary conditions. Dalir and Nourazar [3] found out the boundary layer flow of various nanofluids flowing past a moving semi-infinite plate using the homotopy perturbation method. Ghalambaz et al. [4] found the numerical solution for the natural convective flow of nanofluids over a convectively heated vertical plate in a saturated Darcy porous medium. Sulochana et al. [5] addressed the unsteady magnetohydrodynamic boundary layer flow of a nanofluid flowing past a permeable stretching surface sheet immersed in a porous medium. Mishra et al. [6] carried out a numerical study of an oscillatory unsteady Magnetohydrodynamic flow, heat and mass transfer in a vertical rotating channel with an inclined uniform magnetic field and hall effect. Ashwinkumar et al. [7] determined the momentum, heat and mass transfer characteristics of a magnetic nanofluid flowing past a vertical plate embedded in a porous medium filled with ferrous nanoparticles. Samrat et al. [8] found the heat and mass transfer characteristics of an unsteady flow of Casson nanofluid flowing past an elongated surface with a thermal radiation effect. Shaw et al. [9] formulated the transfer of mass and heat of nanofluid flowing over three different geometries of a non-Darcy permeable vertical cone/wedge/vertical plate under viscous dissipation and thermal radiation. The Soret, Dufour, hall current and rotation effects on MHD natural convective heat and mass transfer flow past an accelerated vertical plate through a porous medium were calculated by Kumar et al. [10]. Khan et al. [11] scrutinised the magnetohydrodynamic flow of nanomaterials over a stretchable surface with melting heat effect, dissipation, heat flux, Joule heating effect, thermophoresis, Brownian motion and entropy generation. Rasheed et al. [12] discussed the impact of nanofluid flowing over an elongated moving surface with a uniform hydromagnetic field and non-linear heat reservoir. Kumawat et al. [13] analysed the entropy generation of MHD blood flowing through a stenosed permeable curved artery with a heat source and chemical reaction. Tlili et al. [14] studied the effect of fibre laser welding parameters on temperature distribution, weld bead dimensions, melt flow velocity and microstructure by using finite volume and experimental methods. Hejazi et al. [15] explored the role of velocity slip effects for mixed convection flow of nanofluid due to an inclined surface. Anantha Kumar et al. [16] studied the control of electromagnetic induction over the flow and heat transmission in shear-thickening hybrid nano- and ferrofluids for cooling/heating applications. Veera Krishna et al. [17] studied the effect of thermal conductivity on temperature on free convective movement of an incompressible viscous fluid through a heated uniform and vertical wavy surface. Khanduri et al. [18] focussed on the effect of Hall and ion slips on MHD blood flowing through a catheterised multi-stenosis artery with thrombosis. Sharma et al. [19] analysed the higher order endothermic/exothermic chemical reactions with activation energy by considering thermophoresis and Brownian motion effects on MHD mixed convective flow across a vertical stretching surface. The entropy optimisation of MHD flow past a continuously stretching surface was carried out by Khanduri et al. [20].

After reviewing the aforementioned research studied, we analysed the free convective flow with heat and mass transfer of a heat-generating and chemically reacting nanofluid passing through a vertical moving porous plate in a conducting field taking into account the viscous dissipation effect along with Soret and Dufour effects. To the best of our knowledge, no study has focussed on the chemical reaction, heat source/sink and viscous dissipation effects on MHD flow of a nanofluid passing through a vertically infinite moving porous plate subjected to Soret and Dufour effects. Therefore, to bridge this research gap and due to its realistic significance in the number of effective applications in geophysics and energy-related problems, we carried out this study. The study of heat and mass transfer rates of nanofluids as a combination of base fluid and low concentration of nano-sized particles of metals is noteworthy due to its wide spectrum of applications in engineering devices in power and chemical engineering, military, surveillance cameras, microchips and medicine for drug delivery, specifically for cancer cells. Continuing on this line of research, our aim is to solve the governing equations corresponding to the physical model under analysis with the help of the bvp4c numerical scheme and to identify effects of different flow parameters in the equations. For this, coupled non-linear partial differential equations are transformed to ordinary differential equations by the Laplace transform technique and then solved numerically by using the bvp4c function in MATLAB. The numerical results attained are explored using contour plots, and main research outcomes are mentioned at the end of this article.

2. MATHEMATICAL FORMULATION

In this research, a magneto-nano and heat-generating fluid flowing towards an infinite porous plate, which is oriented in vertical direction and moving with an impulsive motion as shown in Fig. 1, is considered.

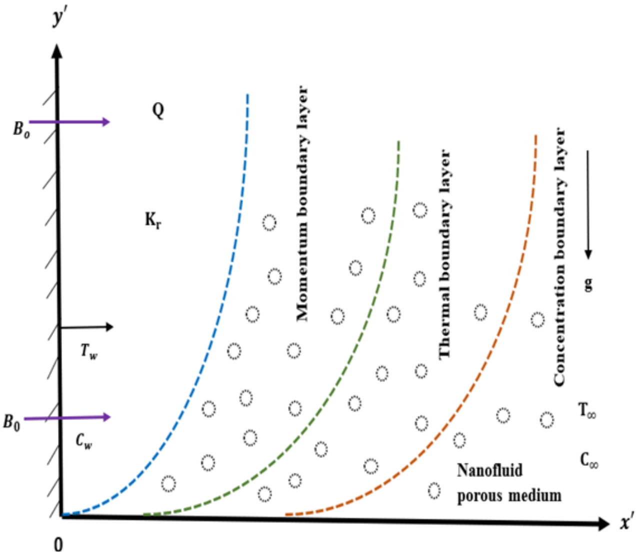


Fig 1. Physical configuration of the problem

We take into account an unsteady free convective, heat and mass transfer flow by incorporating Soret and Dufour effects along with viscous dissipation. The y' axis is vertically upwards along with the plate, while the x' axis is in a direction perpendicular to it. The magnetic field B_0, which is uniformly transverse, is applied parallel to the x' axis. The temperature of the plate fluctuated from T'_∞ to T'_w, and the concentration of the plate varied from C'_∞ to C'_w when the movement of the plate starts in its plane with velocity λu_0 at time $t' > 0$. Water with nanoparticles of copper (Cu) is taken as the base fluid, and both are in thermal equilibrium. In equations governing the flow of problem, density is considered to be linearly dependent on temperature. The magnetic field is supposed to be $\vec{B} \equiv (0, B_o, 0)$ because in comparison to the applied magnetic field, the induced magnetic field generated by the flow of fluid is considered negligible. Furthermore, we suppose the electric field as $\vec{E} = (0,0,0)$ to consider the effect of polarisation of the fluid negligible (Cramer et al. [21]). The basic equations of momentum, energy and concentration limited only to spherical nanoparticles governing the problem in the presence of magnetic field, thermal diffusion, heat source or sink, viscous dissipation, chemical reaction and Soret and Dufour effects are given as follows [22]:

$$\rho_{nf}\ \frac{\partial u'}{\partial t'} = \ \mu_{nf}\ \frac{\partial^2 u'}{\partial x'^2} + \ g(\rho\beta)_{nf}(T' - T'_\infty)$$
$$+g(\rho\beta')_{nf}(C' - C'_\infty) - \sigma_{nf} B_o^2 u' - \frac{\mu_{nf}}{K'_p}\ u' \tag{1}$$

$$\left(\rho C_p\right)_{nf} \frac{\partial T'}{\partial t'} = k_{nf} \frac{\partial^2 T'}{\partial x'^2} + \mu_{nf}\left(\frac{\partial u}{\partial x'}\right)^2 + \ D\rho_{nf} C_p \frac{\partial^2 C'}{\partial x'^2} +$$
$$q_0(T' - T'_\infty) \tag{2}$$

$$\frac{\partial C'}{\partial t'} = D\,\frac{\partial^2 C'}{\partial x'^2} + D_1 \frac{\partial^2 T'}{\partial x'^2} + \ K_1(C' - C'_\infty) \tag{3}$$

The thermal conductivity of nanofluid is given as follows:

$$k_{nf} = k_f \left[\frac{k_s+2k_f-2\phi(k_f-k_s)}{k_s+2k_f+\phi(k_f-k_s)}\right] \tag{4}$$

Also,

$$\mu_{nf} = \frac{\mu_f}{(1-\phi)^{2.5}}, \rho_{nf} = (1-\phi)\rho_f + \phi\rho_s,$$
$$(\rho C_p)_{nf} = (1-\phi)(\rho C_p)_f + \phi(\rho C_p)_s$$
$$(\rho\beta)_{nf} = (1-\phi)(\rho\beta)_f + \phi(\rho\beta)_s$$
$$(\rho\beta')_{nf} = (1-\phi)\,(\rho\beta')_f + \phi(\rho\beta')_s$$
$$\sigma_{nf} = \sigma_f\left[1+\frac{3(\sigma-1)}{(\sigma+2)-(\sigma-1)\phi}\right], \sigma = {}^{\sigma_s}/_{\sigma_f} \tag{5}$$

where *nf*, *f* and *s* in Eqs. (1)–(5) indicate thermophysical properties of the nanofluid, base fluid and the nanoparticles, respectively.

The corresponding initial and boundary conditions are given as follows:

$$u' = 0, T' = T'_\infty, C' = C'_\infty \; for\ all\ x' \geq 0\ and\ t' = 0$$
$$u' = \lambda u_o, T' = T'_w, C' = C'_w\ at\ x' = 0\ and\ t' > 0$$
$$u' \to 0, T' \to T'_\infty, C' \to C'_\infty\ as\ x' \to \infty\ and\ t' > 0 \tag{6}$$

where λ = 0 signifies the direction of the moving plate when the plate is at rest, while λ = ± 1 indicates the direction of the plate moving in both forward and backward directions.

Using the non-dimensional variables in Eqs. (1)–(3),

$$x = \frac{u_0 x'}{v_f}, t = \frac{u_0^2\,t'}{v_f}, u = \frac{u'}{u_0}, \theta = \frac{T'-T'_\infty}{T'_w-T'_\infty}, C = \frac{C'-C'_\infty}{C'_w-C'_\infty},$$
$$K'_P = \frac{K\,v_{nf}^2}{u_0^2}$$

The governing equations in the non-dimensional form are as given follows:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + Gr\,A_2\theta + Gc\,A_5 C - M^2 A_3 u - \frac{1}{K}u \tag{7}$$
$$\frac{\partial\theta}{\partial t} = \frac{1}{Pr}\frac{\partial^2\theta}{\partial x^2} + Ec\left(\frac{\partial u}{\partial x}\right)^2 + Du\,\frac{\partial^2 C}{\partial x^2} + \frac{v}{Pr}\,Q\,\theta \tag{8}$$
$$\frac{\partial C}{\partial t} = \frac{1}{S_c}\frac{\partial^2 C}{\partial x^2} + S_o\frac{\partial^2\theta}{\partial x^2} + K_r C \tag{9}$$

Where:

Gr(Grashof number) $= \frac{g\beta_f\mu_f(T'_w-T'_\infty)}{u_0^3}$,

Gc(mass Grashof number) $= \frac{g\beta_f^*\mu_f(C'_w-C'_\infty)}{u_0^3}$,

M^2(magnetic parameter)$= \frac{\sigma_f B_0^2 v_f}{\rho_f u_o^2}$,

K(porosity parameter) $= \frac{K'_p u_0^2}{v_f^2}$,

Pr(Prandtl number) $= \frac{\mu_f\,C_p}{k_f}$,

Ec(Eckert number) $= \frac{u_0^2}{C_p(T'_w-T'_\infty)}$,

Du(Dufour number) $= \frac{D}{v_f}\left(\frac{C'_w-C'_\infty}{T'_w-T'_\infty}\right)$,

Q(heat source/sink parameter) $= \frac{q_0 v_f^2}{k_f u_0^2}$,

S_c(Schmidt number) $= \frac{v_f}{D}$,

S_o(Soret number) $= \frac{D_1}{v_f}\left(\frac{T'_w-T'_\infty}{C'_w-C'_\infty}\right)$,

Kr(Chemical reaction parameter) $= \frac{K_1 v_f}{u_0^2}$

A_2, A_3 and A_5 are constants defined as follows:

$$A_2 = \left[\frac{(1-\phi)}{\rho_{nf}} + \frac{\phi(\rho\beta)_s}{(\rho\beta)_f\,\rho_{nf}}\right]$$
$$A_3 = \frac{\left[1+\frac{3(\sigma-1)}{(\sigma+2)-(\sigma-1)\phi}\right]}{\left[\left(1-\phi+\phi\,\frac{\rho_s}{\rho_f}\right)\right]}$$
$$A_5 = \left[\frac{(1-\phi)}{\rho_{nf}} + \frac{\phi(\rho\beta^*)_s}{(\rho\beta^*)_f\,\rho_{nf}}\right]$$

On application of the Laplace Transform technique, Eqs. (7)–(9) are transformed to the following forms of ordinary differential equations:

$$\frac{d^2\bar{u}}{dx^2} - \left(s + M^2 A_3 + \frac{1}{K}\right)\bar{u} + Gr\,A_2\bar{\theta} + Gc\,A_5\bar{C} = 0 \tag{10}$$
$$\frac{1}{Pr}\frac{d^2\bar{\theta}}{dx^2} + Du\,\frac{d^2\bar{C}}{dx^2} + Ec\left(\frac{d\bar{u}}{dx}\right)^2 + \frac{v}{Pr}\,Q\bar{\theta} - s\bar{\theta} = 0 \tag{11}$$
$$\frac{1}{S_c}\frac{d^2\bar{C}}{dx^2} + S_o\frac{d^2\bar{\theta}}{dx^2} - (s-Kr)\bar{C} = 0 \tag{12}$$

as in Eqs. (10)–(12) is a complex parameter also known as Laplace operator, which is the same for all the equations.

The corresponding initial and boundary conditions in the non-dimensional form are given as follows:

$t = 0{:}\ u = 0, \theta = 0, C = 0$ for all $x \geq 0$
$t > 0{:}\ u = \lambda, \theta = 1, C = 1$ at $x = 0$
$t > 0{:}\ u \to 0, \theta \to 0, C \to 0$ as $x \to \infty$ (13)

2.1. Numerical solution

Due to the difficulty in finding exact solutions of non-linear partial differential Eqs. (7)–(9), the equations after transformation to ordinary differential equations (10–12) are numerically solved using the bvp4c function by MATLAB software. bvp4c is a method for obtaining numerical solution to the boundary value problems. The basic solution method is subjected to polynomial collocation with four Lobatto points for bvp4c. It uses the three-stage Lobatto IIIa formula. It is a collocation formula, and the collocation polynomial provides a solution that is accurate up to fourth order on a given interval. Before using bvp4c for solving a problem, we rewrite the given second or higher order ODEs as a system of first-order ODEs, which in the present problem in the following manner:

$$\bar{\theta} = y(1) \qquad \bar{C} = y(3) \qquad \bar{u} = y(5)$$
$$\overline{\theta'} = y(2) \qquad \overline{C'} = y(4) \qquad \overline{u'} = y(6)$$

The approach is using to obtain parameterised initial conditions at the initial point of the interval. A standard ODE solver is used for arriving at the solution across the domain. After this, root finding is adopted to find the appropriate parameter values to apply the boundary conditions at the end point of the interval, which are then used for finding the solution across the domain. Hence, it is clear that the solver for initial value problems is turned into a solver for boundary value problems. This is the overall working algorithm for this numerical technique. The code used for this analysis is given in the appendix.

3. RESULT AND DISCUSSIONS

The follow-up of present investigation is currently premeditated with the help of significant sketched graphical features. Several important features of heat-generating and chemically reacting nanofluid with viscous dissipation incorporating Soret and Dufour effects in a conducting field by varying strengths of Dufour number (Du), Soret number (S_o), Schmidt number (S_c), magnetic parameter (M^2), porosity parameter (K), Grashof number (G_r), mass Grashof number (G_c), chemical reaction parameter (Kr), Prandtl number (Pr), Eckert number (Ec) and heat source/sink parameter (Q) have been explored through contours, as shown in Figs. 2–12. The values of the parameters are chosen arbitrarily. The dashed line shows variation of the motion of the fluid for the plate moving in backward direction, i.e., $\lambda = -1$, while the solid line shows variation of the plate moving in forward direction, i.e., $\lambda = 1$.

3.1. Effect of variation of different parameters on velocity field

Figs. 2–5 depict the effect of various governing parameters on velocity of the fluid.

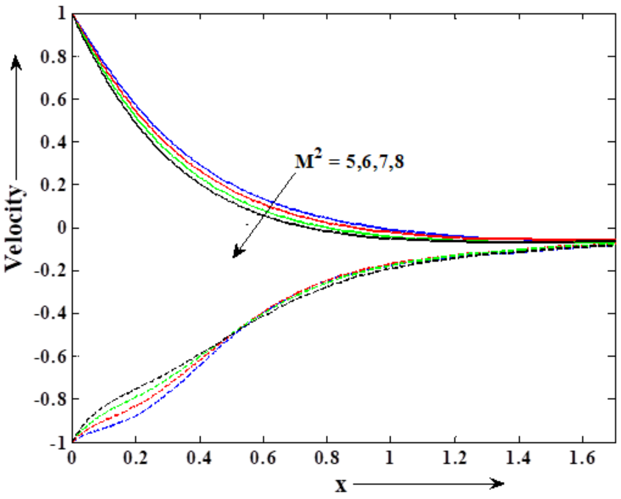


Fig. 2. Effect of M^2 on velocity when $\mathrm{Du} = 5, \mathrm{S_o} = 7, \mathrm{S_c} = 2.5, \mathrm{Ec} = 0.08, \mathrm{K} = 3, \mathrm{Gr} = 4, \mathrm{Gc} = 6, \mathrm{Pr} = 1.9, \mathrm{R} = 6, \mathrm{Q} = 4, \mathrm{Kr} = 3$

The influence of the magnetic parameter (M^2) on velocity distribution is portrayed in Fig. 2. An electric charge is more violently pushed in contact with stronger magnetic field and from the regions of high magnetic field gradient. With an increase in the magnetic field strength, the velocity of nanofluid is found to decrease for motion of the plate in forward direction, while for motion of the plate in backward direction, velocity of the fluid first increases, and after a certain point of time, it starts reversing this trend. As there is a rise in resistive type of force called the Lorentz force, the influence of a transverse magnetic field on the conducting fluid has a tendency to slow down the motion of fluid in the boundary layer region. It produces more resistance to the phenomenon of transportation. The graphical representation of porosity parameter (K) on the velocity profile is depicted in Fig. 3. It is observed that with growing K, velocity in the boundary layer increases with the motion of the plate in forward direction, and for backward motion, velocity first decreases and afterwards slightly increases. This is due to the fact that increasing the value of K assists the fluid considerably to move fast due to reduction in drag force. The porous material is the cause for restriction of flow of the fluid. Fig. 4 displays the effect of thermal Grashof number (Gr) on the velocity profiles. As the Grashof number increases, the velocity of the fluid increases for both forward and backward motions of the plate. For different values of the mass Grashof number (Gc), the velocity profile is plotted in Fig. 5. As the mass Grashof number increases, the velocity of the fluid increases for both forward and backward motions of the plate. This is due to the fact that an increase in the values of the thermal Grashof number and mass Grashof number has a tendency to increase thermal and mass buoyancy effects, which gives rise to an increase in the induced flow.

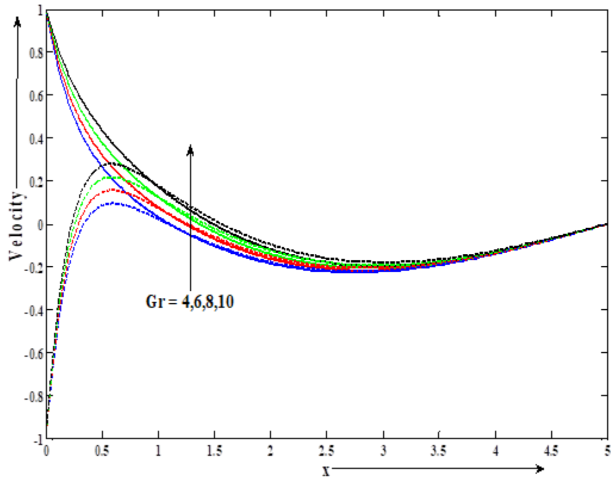


Fig. 4. Effect of Gr on velocity when $\mathrm{Du} = 5, \mathrm{S_o} = 7, \mathrm{S_c} = 2.5, \mathrm{Ec} = 0.08, \mathrm{M}^2 = 4, \mathrm{K} = 0.7, \mathrm{Gc} = 6, \mathrm{Pr} = 1.9, \mathrm{R} = 6, \mathrm{Q} = 4, \mathrm{Kr} = 3$

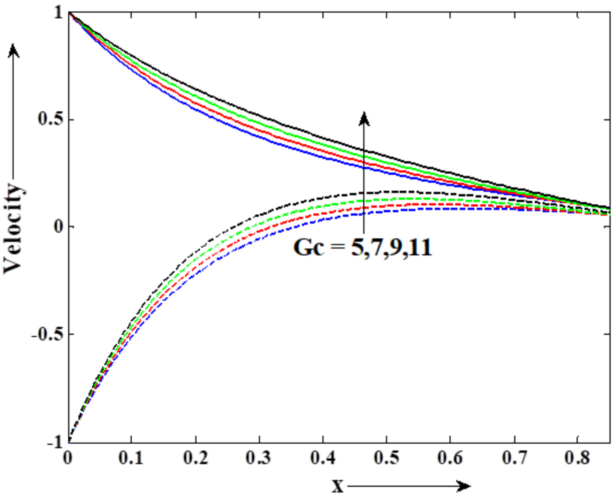


Fig. 5. Effect of Gc on velocity when $\mathrm{Du} = 5, \mathrm{S_o} = 7, \mathrm{S_c} = 2.5, \mathrm{Ec} = 0.08, \mathrm{M}^2 = 4, \mathrm{Gr} = 4, \mathrm{K} = 0.7, \mathrm{Pr} = 1.9, \mathrm{R} = 6, \mathrm{Q} = 4, \mathrm{Kr} = 3$

3.2. Effect of variation of different parameters on skin friction

The values of skin friction (τ) for various parameters like thermal Grashof number (Gr), mass Grashof number (Gc), magnetic parameter (M^2) and porosity parameter (K) are given in Tabs. 1–4 ($R = 4$ taken as constant). It can be interpreted from Tabs. 1 and 2 that values of skin friction decrease with increasing values of Grashof number (Gr) and mass Grashof number (Gc). Table 3 shows the variation of skin friction with respect to the porosity parameter (K). Clearly, skin friction decreases with an increase in the values of the porosity parameter (K). The impact of magnetic parameter (M^2) on skin friction is evaluated and given in Table 4. Evidently, there is an increase in skin friction with increasing magnetic parameter. It also highlights a decrease in viscous drag forces.

Tab. 1. Variation with Grashof number

Gr	Gc	Du	S_o	S_c	Ec	M^2	K	Pr	Q	Kr	τ
5	6	2	4	0.44	1	9	0.5	0.79	4	0.5	2.9622
6	6	2	4	0.44	1	9	0.5	0.79	4	0.5	2.8326
7	6	2	4	0.44	1	9	0.5	0.79	4	0.5	2.7028
8	6	2	4	0.44	1	9	0.5	0.79	4	0.5	2.5729

Tab. 2. Variation with mass Grashof number

Gc	Gr	Du	S_o	S_c	Ec	M^2	K	Pr	Q	Kr	τ
6	5	2	4	0.44	1	9	0.5	0.79	4	0.5	2.9622
7	5	2	4	0.44	1	9	0.5	0.79	4	0.5	2.7850
8	5	2	4	0.44	1	9	0.5	0.79	4	0.5	2.6113
9	5	2	4	0.44	1	9	0.5	0.79	4	0.5	2.4406

Tab. 3. Variation with porosity parameter

K	Du	S_o	S_c	Ec	M^2	Gr	Gc	Pr	Q	Kr	τ
1.1	2	4	0.44	1	9	5	6	0.79	4	0.5	2.8198
1.3	2	4	0.44	1	9	5	6	0.79	4	0.5	2.8012
1.5	2	4	0.44	1	9	5	6	0.79	4	0.5	2.7875
1.7	2	4	0.44	1	9	5	6	0.79	4	0.5	2.7771

Tab. 4. Variation with magnetic parameter

M^2	Du	S_o	S_c	Ec	K	Gr	Gc	Pr	Q	Kr	τ
2	2	4	0.44	0.02	0.5	4	6	0.79	3	0.5	-0.4956
3	2	4	0.44	0.02	0.5	4	6	0.79	3	0.5	-0.0329
4	2	4	0.44	0.02	0.5	4	6	0.79	3	0.5	0.4994
5	2	4	0.44	0.02	0.5	4	6	0.79	3	0.5	1.0630

3.3. Effect of variation of different parameters on temperature field

Figs. 6–9 illustrate the effect of various governing parameters on temperature of the fluid. The low Eckert number ($Ec \ll 1$) allows making the judgement if the effects of self-heating due to dissipation can be overlooked or not. From Fig. 6, it can be concluded that the temperature of fluid decreases with increasing values of Eckert number (Ec), which shows a decrease in the thermal boundary layer thickness, thus consequently increasing the rate of heat transfer. Fig. 7 reveals that the temperature distribution profile of nanofluid decreases with increasing values of Prandtl number (Pr) predominantly due to a decrease in viscosity of the nanofluid. For a low Prandtl number (< 1), the thermal boundary layer is thicker than the velocity boundary layer. Fig. 8 depicts the effect of the heat source parameter (Q) on heat transfer processes. The heat source/sink parameter $Q > 0$ decreases the thermal conductivity, which results in the decrease in the temperature of the fluid.

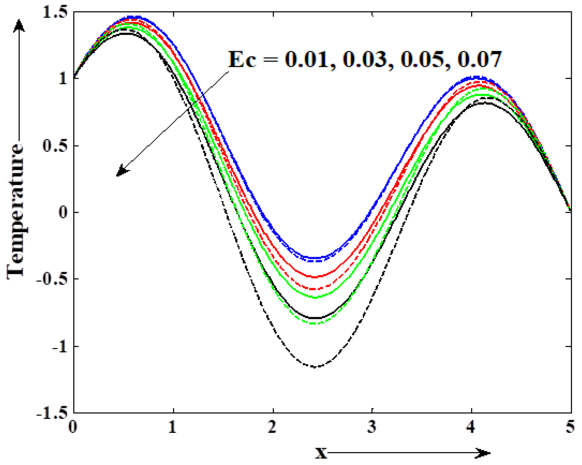


Fig. 6. Effect of Ec on temperature when Du = 5, $S_o = 2, M^2 = 9, K = 5, Gr = 4, Gc = 6, Pr = 0.76, R = 0.4, Q = 0.2, Kr = 5, S_c = 0.44$

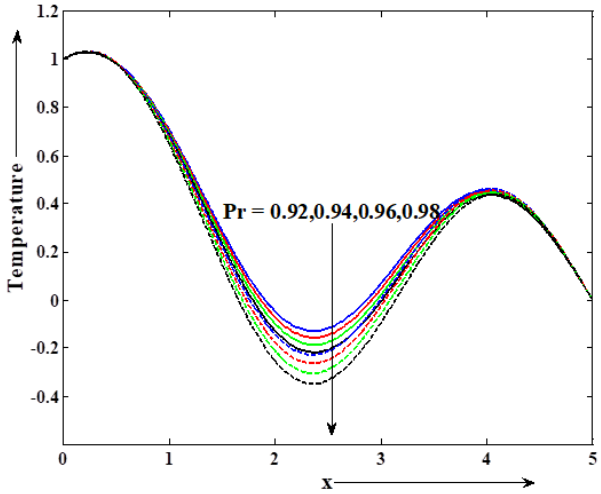


Fig. 7. Effect of Pr on temperature when Du = 7, $S_o = 2, M^2 = 9, K = 5, Gr = 4, Gc = 6, Ec = 0.07, R = 0.4, Q = 0.3, Kr = 5, S_c = 0.44$

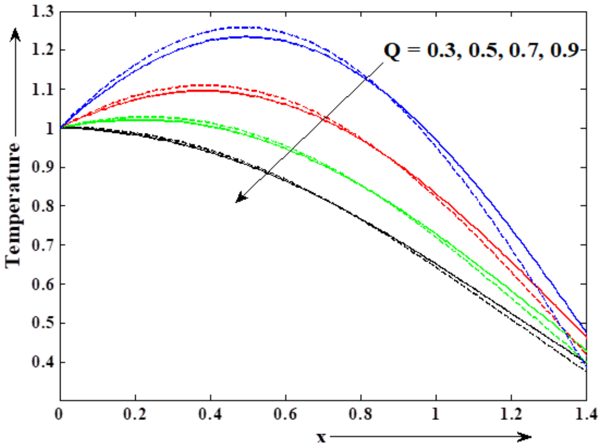


Fig. 8. Effect of Q on temperature when Du = 7, $S_o = 2, M^2 = 9, K = 5, Gr = 4, Gc = 6, Ec = 0.07, R = 0.4, Pr = 0.76, Kr = 5, S_c = 0.44$

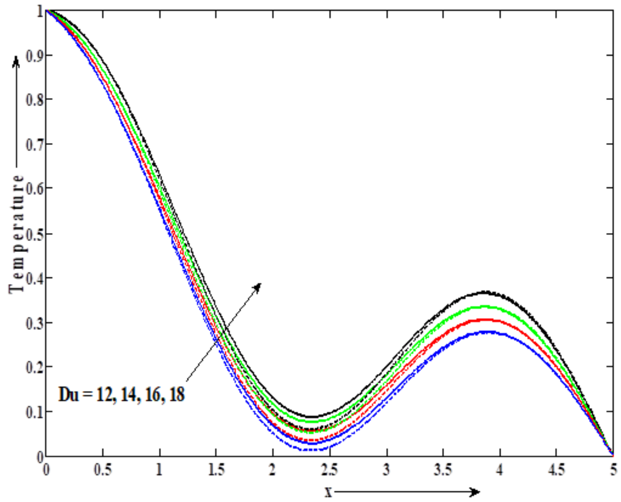


Fig. 9. Effect of Du on temperature when $\mathrm{Q} = 0.3, \mathrm{S_o} = 2, \mathrm{M}^2 = 9, \mathrm{K} = 5, \mathrm{Gr} = 4, \mathrm{Gc} = 6, \mathrm{Ec} = 0.07, \mathrm{R} = 0.4, \mathrm{Pr} = 0.76, \mathrm{Kr} = 5, \mathrm{S_c} = 0.44$

Fig. 9 displays the variation in temperature with respect to Dufour number (Du). Higher values of the Dufour number causes an increase in temperature. It is clear from this graph that the temperature profile increases with an increase in the Dufour number (Du), which shows an increase in heat transfer.

3.4. Effect of variation of different parameters on Nusselt number coefficients

The variation in the Nusselt number with respect to the Prandtl number is measured and given in Tab. 5. From the table, it is seen that the Nusselt number increases with an increase in the Prandtl number. Tab. 6 shows the effect of the heat source/sink parameter (Q) on the Nusselt number. As the heat source parameter increases, the value of the Nusselt number also increases. Tab. 7 shows the variation in the Nusselt number with respect to the Eckert number. It was observed that the Nusselt number increases with an increase in the Eckert number (Ec). Tab. 8 shows the variation in the Nusselt number with an increase in the Dufour number (Du). It is found that the Nusselt number shows a decrease with the increasing values of the Dufour number $(R = 5$ taken as constant). It reflects an increase in the heat transfer rate.

Tab. 5. Variation with Prandtl number

Pr	Du	S_o	S_c	Ec	M^2	K	Gr	Gc	Q	Kr	Nu
0.82	2	6	0.57	0.04	9	1.5	5	6	4	2	3.2663
0.84	2	6	0.57	0.04	9	1.5	5	6	4	2	3.2757
0.86	2	6	0.57	0.04	9	1.5	5	6	4	2	3.2850
0.88	2	6	0.57	0.04	9	1.5	5	6	4	2	3.2942

Tab. 6. Variation with heat source parameter

Q	Du	S_o	S_c	Ec	M^2	K	Gr	Gc	Pr	Kr	Nu
1	2	6	0.57	0.04	9	0.5	5	6	0.82	5	0.8377
2	2	6	0.57	0.04	9	0.5	5	6	0.82	5	2.2000
3	2	6	0.57	0.04	9	0.5	5	6	0.82	5	2.8508
4	2	6	0.57	0.04	9	0.5	5	6	0.82	5	3.4073

Tab. 7. Variation with Eckert number

Ec	Du	S_o	S_c	Q	M^2	K	Gr	Gc	Pr	Kr	Nu
0.1	2	6	0.57	1	3	0.5	5	6	0.82	5	0.4686
0.2	2	6	0.57	1	3	0.5	5	6	0.82	5	0.7026
0.3	2	6	0.57	1	3	0.5	5	6	0.82	5	0.7532
0.4	2	6	0.57	1	3	0.5	5	6	0.82	5	0.8377

Tab. 8. Variation with Dufour number

Du	S_o	S_c	Q	M^2	K	Gr	Gc	Ec	Pr	Kr	Nu
5	6	0.57	1	3	0.5	5	6	0.02	0.82	5	0.8704
6	6	0.57	1	3	0.5	5	6	0.02	0.82	5	0.8649
7	6	0.57	1	3	0.5	5	6	0.02	0.82	5	0.8559
8	6	0.57	1	3	0.5	5	6	0.02	0.82	5	0.8402

3.5. Effect of variation in different parameters on concentration field

Figs. 10–12 show the effect of various governing parameters on the concentration of fluid.

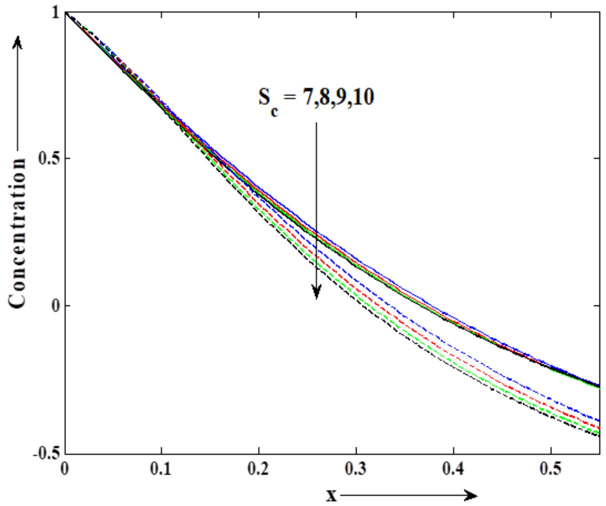


Fig. 10. Effect of $\mathrm{S_c}$ on concentration when $\mathrm{Du} = 3, \mathrm{S_o} = 4, \mathrm{S_c} = 3.5, \mathrm{Ec} = 0.02, \mathrm{M}^2 = 9, \mathrm{K} = 2, \mathrm{Gr} = 4, \mathrm{Gc} = 6, \mathrm{Pr} = 0.79, \mathrm{R} = 0.4, \mathrm{Q} = 1, \mathrm{Kr} = 5$

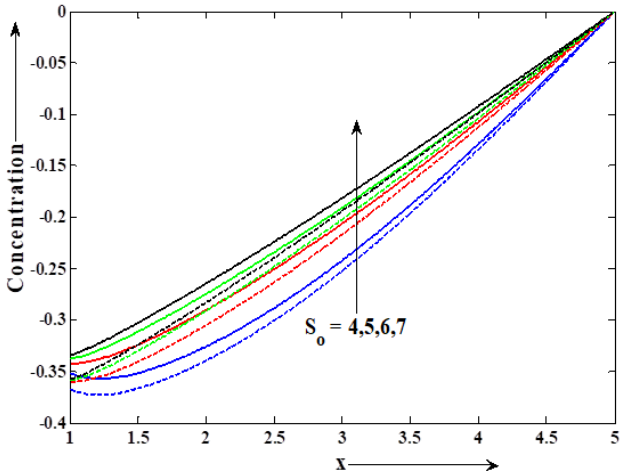


Fig. 11. Effect of $\mathrm{S_0}$ on concentration when $\mathrm{Du} = 2, \mathrm{S_c} = 1.5, \mathrm{Ec} = 0.02, \mathrm{M}^2 = 9, \mathrm{K} = 0.5, \mathrm{Gr} = 4, \mathrm{Gc} = 6, \mathrm{Pr} = 0.79, \mathrm{R} = 4, \mathrm{Q} = 1, \mathrm{Kr} = 5$

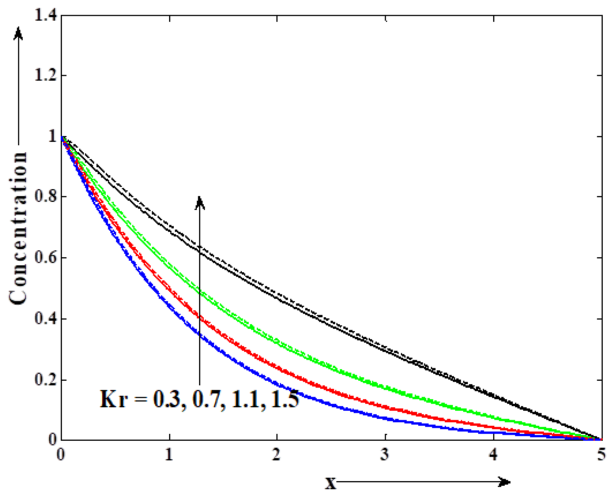


Fig. 12. Effect of Kr on concentration when $\mathrm{Du} = 2, \mathrm{S_0} = 4$, $\mathrm{S_c} = 0.44, \mathrm{Ec} = 0.02, \mathrm{M}^2 = 9, \mathrm{K} = 0.5, \mathrm{Gr} = 4, \mathrm{Gc} = 6$, $\mathrm{Pr} = 0.79, \mathrm{R} = 4, \mathrm{Q} = 1$

Fig. 10 depicts the variation in the concentration field subjected to Schmidt number (S_c). When the Schmidt number is high, the momentum is transported by molecular means across a liquid much more effectively than species. This figure reveals that an increase in S_c reduces the concentration profile. An increase in the Schmidt number results in weaker solute diffusivity, allowing a shallower penetration of the solutal effect. As a consequence, the concentration decreases with increasing S_c. From Fig. 11, it is evident that concentration profile increases with an increase in the Soret number (S_o). The higher values of the Soret number develops a higher convective flow. Fig. 12 describes the chemical reaction parameter (Kr) effect on nanofluid concentration profile. Small values of the chemical reaction parameter results in a decrease in the heat transfer coefficient and an increase in the mass transfer rate. As foreseen, the concentration increases with an increase in Kr. It reveals that with an increase in concentration under the influence of the chemical reaction parameter, more atoms move, which are more likely to collide with the reactant particles, causing the reaction to occur faster and reflecting the strengthened interfacial mass transfer by a chemical reaction.

3.6. Effect of variation of different parameters on Sherwood number coefficients

The effect of the Schmidt number (S_c) on the Sherwood number is given in Tab. 9.The Sherwood number (Sh) increased with increasing Schmidt number (S_c). Tab. 10 reflects the effect of the Soret number on the Sherwood number. Here it is noticed that the Sherwood number increases with increased Soret number (S_o). Tab. 11 shows the behaviour of the chemical reaction parameter with respect to the Sherwood number. Clearly, the Sherwood number is a decaying function of the chemical reaction parameter (Kr). ($R = 5$ taken as constant).

Tab. 9. Variation with Schmidt number

$\mathrm{S_c}$	Du	$\mathrm{S_o}$	Ec	M^2	K	Gr	Gc	Pr	Q	Kr	Sh
0.71	4	6	0.04	3	0.5	5	6	0.82	1	5	1.5742
0.72	4	6	0.04	3	0.5	5	6	0.82	1	5	1.5886
0.73	4	6	0.04	3	0.5	5	6	0.82	1	5	1.6029
0.74	4	6	0.04	3	0.5	5	6	0.82	1	5	1.6170

Tab. 10. Variation with Soret number

$\mathrm{S_o}$	Du	$\mathrm{S_c}$	Ec	M^2	K	Gr	Gc	Pr	Q	Kr	Sh
5	4	0.71	0.04	3	0.5	5	6	0.82	1	5	1.3136
6	4	0.71	0.04	3	0.5	5	6	0.82	1	5	1.5742
7	4	0.71	0.04	3	0.5	5	6	0.82	1	5	1.7744
8	4	0.71	0.04	3	0.5	5	6	0.82	1	5	1.9401

Tab. 11. Variation with chemical reaction parameter

Kr	Du	$\mathrm{S_o}$	Ec	M^2	$\mathrm{S_c}$	K	Gr	Gc	Pr	Q	Sh
2	4	6	0.04	3	0.71	0.5	5	6	0.82	1	1.7754
4	4	6	0.04	3	0.71	0.5	5	6	0.82	1	1.6477
6	4	6	0.04	3	0.71	0.5	5	6	0.82	1	1.4918
8	4	6	0.04	3	0.71	0.5	5	6	0.82	1	0.9127

3.7. Comparison of the results

The results of present study are validated by comparing them with results of published studies in the absence of chemical reaction, viscous dissipation and Dufour effect, as obtained by Reddy et al. [23]. The same can be interpreted from Figs. 13 and 14.

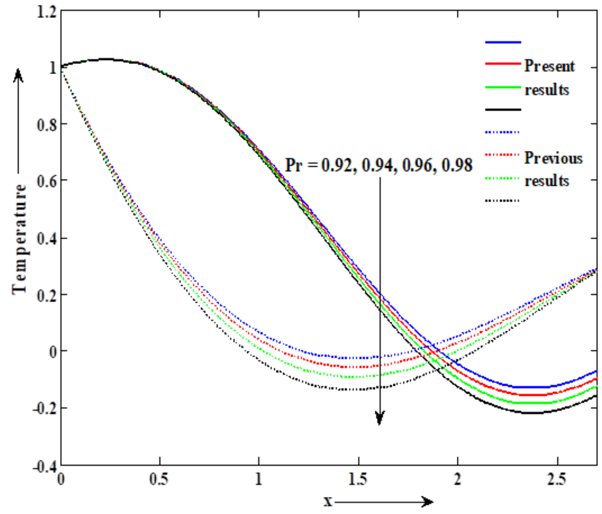


Fig. 13. Comparison of the results. Effect of Pr on temperature in the absence of Du, Ec and Kr

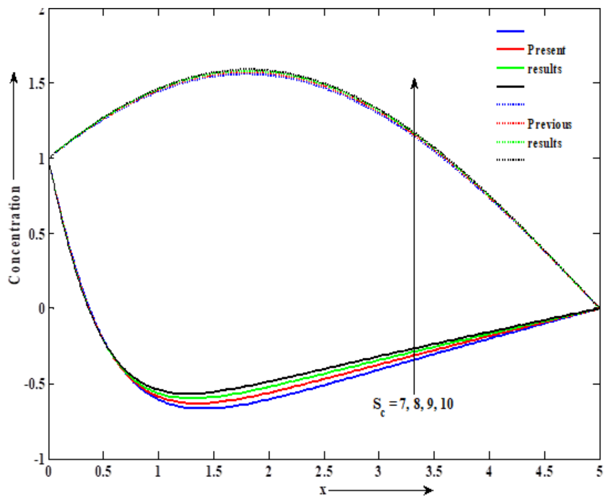


Fig. 14. Comparison of the results. Effect of $\mathrm{S_c}$ on concentration in the absence of Du, Ec and Kr

4. CLOSING REMARKS

In this study, the effect of viscous dissipation, heat source/sink, characteristics of heat and mass transfer on a chemically reacting magneto-nanofluid in a conducting field passing through a vertically moving porous plate in the presence of Soret and Dufour effects has been studied. By the using of the Laplace transform technique, Partial Differential Equations is converted to a set of Ordinary Differential Equations, and the numerical solutions are obtained by the MATLAB software package using the bvp4c function. From the numerical analysis, the following observations are derived:

- With an increase in the Grashof number and mass Grashof number, the velocity of the fluid is accelerated, while a reverse trend is observed in case of an increase in the magnetic parameter. In case of porosity parameter, velocity increases for the motion of the plate in forward direction, i.e., $\lambda = 1$, while the reverse trend is observed for the motion in backward direction, i.e., $\lambda = -1$.
- Temperature of the fluid is inversely related to the Eckert number, Prandtl number and heat source/sink parameter, whereas the Dufour number is directly related to it.
- The concentration profile declines with an increase in the Schmidt number, while it increases with an increase in the Soret number and chemical reaction parameter.
- An increase in the magnetic parameter leads to an increase in the skin friction coefficient, while an opposite effect is observed in case of Grashof number, mass Grashof number and porosity parameter.
- The local Nusselt number is enhanced with enhancement in the heat source/sink parameter, Eckert number and Prandtl number, but a reverse effect is observed in case of Dufour number.
- The Sherwood number increases with increasing Schmidt number and Soret number, whereas it decreases for the chemical reaction parameter.
- A comparison of results with those of previous studies validate the findings of the current study.
- In particular, the flow profile variations with respect to viscous dissipation, heat source/sink and chemical reaction in combination with Soret and Dufour effects have a great significance; thus, they find application in many engineering and industrial fields, particularly in separation of isotopes.
- The current study reveals that in the presence of chemical reaction, the mass transfer rate increase, causing the reaction to occur faster and also the viscous dissipation effect increases the heat transfer rate.

REFERENCES

1. Gupta M, Singh V, Kumar R, Said Z. A review on thermophysical properties of nanofluids and heat transfer applications. Review Sust Energy Rev. 2021; 74: 638-670. https://doi.org/10.1016/j.rser
2. Turkyilmazoglu M, Pop I. Heat and mass transfer of unsteady natural convection flow of some nanofluids past a vertical infinite flat plate with radiation effect. International Journal of Heat and Mass Transfer. 2013; 59:pp. 167171.
3. Dalir N, Nourazar S S. Solution of the boundary layer flow of various nanofluids over a moving semi-infinite plate using HPM. Mechanika. 2014; 20: pp. 57-63.
4. Ghalambaz M, Noghrehabadi A, Ghanbarzadeh A. Natural convection of nanofluids over a convectively heated vertical plate embedded in a porous medium. Brazilian Journal of Chemical Engineering. 2014; 31: 413-427.
5. Sulochana C, Samrat SP. Unsteady MHD Radiative flow of a Nano liquid past a permeable stretching sheet: An analytical study. Journal of Nanofluids. 2017; vol. 6: pp. 711-719.
6. Mishra A, Sharma BK. MHD Mixed Convection Flow in a Rotating Channel in the Presence of an Inclined Magnetic Field with the Hall Effect. Journal of Engineering Physics and Thermophysics 90, 2017; 1488–1499. https://doi.org/10.1007/s10891-017-1710-y
7. Ashwinkumar GP , Sulochana C, Samrat SP .Effect of the aligned magnetic field on the boundary layer analysis of magnetic-nanofluid over a semi-infinite vertical plate with ferrous nanoparticles. Multidiscipline Modeling in Materials and Structures. 2018; Vol. 14 No. 3: pp. 497-515. https://doi.org/10.1108/MMMS-10-2017-0128
8. Samrat SP, Sulochana, C, Ashwinkumar GP. Impact of Thermal Radiation on an Unsteady Casson Nanofluid Flow Over a Stretching Surface. International Journal of Applied and Computational Mathematics 5. 2019; 31. https://doi.org/10.1007/s40819-019-0606-2
9. Shaw S, Motsa, SS, Sibanda P. Nanofluid flow over three different geometries under viscous dissipation and thermal radiation using the local linearization method. Heat Transfer-Asian Research, 2019; 48(3). doi:10.1002/htj.21497
10. Kumar MA, Reddy YD, Goud BS, Rao VS. Effects of Soret, Dufour, hall current and rotation on MHD natural convective heat and mass transfer flow past an accelerated vertical plate through a porous medium. International Journal of Thermofluids. 2021; volume9: 100061.
11. Khan SA, Hayat T, Alsaedi A, Ahmad B. Melting heat transportation in radiative flow of nanomaterials with irreversibility analysis. Renewable and Sustainable Energy Reviews. 2021; Volume 140: 110739 ISSN 1364-0321, https://doi.org/10.1016/j.rser.2021.110739
12. Rasheed HUR, Islam S, Khan Z, Alharbi SO, Alotaibi H, Khan I. Impact of Nanofluid flow over an elongated moving surface with a uniform Hydromagnetic field and non-linear Heat reservoir. Hindawi Complexity. 2021; volume 2021: Article ID9951162.
13. Kumawat C, Sharma BK, M Al-Mdallal Q, Rahimi-Gorji M. Entropy generation for MHD two phase blood flow through a curved permeable artery having variable viscosity with heat and mass transfer. International Communications in Heat and Mass Transfer. 2022; Volume 133: 105954, ISSN 0735-1933. https://doi.org/10.1016/j.icheatmasstransfer.2022.105954
14. Tlili I, Baleanu D, Mohammad Sajadi S, Ghami F, Fagiry MA. Numerical and experimental analysis of temperature distribution and melt flow in fiber laser welding of Inconel 625. The International Journal of Advanced Manufacturing Technology. 2022; 121: 765–784. https://doi.org/10.1007/s00170-022-09329-3
15. Hejazi HA, Ijaz Khan M, Raza A, Smida K, Khan SU and Tlili I. Inclined surface slip flow of nanoparticles with subject to mixed convection phenomenon: Fractional calculus applications. Journal of the Indian Chemical Society, 2022; Volume 99: Issue 7, 100564, ISSN 0019-4522. https://doi.org/10.1016/j.jics.2022.100564
16. Anantha Kumar K, Sandeep N, Samrat SP, Ashwinkumar GP. Effect of electromagnetic induction on the heat transmission in engine oil-based hybrid nano and ferrofluids: A nanotechnology application. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering. 2022;0(0). doi:10.1177/09544089221139569
17. Veera Krishna M. Numerical investigation on steady natural convective flow past a perpendicular wavy surface with heat absorption/generation. International Communications in Heat and Mass Transfer. 2022; Volume 139: 106517, ISSN 0735-1933. https://doi.org/10.1016/j.icheatmasstransfer.2022.10657
18. Khanduri U, Sharma BK. Hall and ion slip effects on hybrid nanoparticles (Au-GO/blood) flow through a catheterized stenosed artery with thrombosis. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2022;0(0). doi:10.1177/09544062221136710

19. Sharma BK, Gandhi R, Mishra NK, Al-Mdallal QM. Entropy generation minimization of higher-order endothermic/exothermic chemical reaction with activation energy on MHD mixed convective flow over a stretching surface. Scientific Reports 12. 2022; 17688. https://doi.org/10.1038/s41598-022-22521-5
20. Khanduri U, Sharma BK. Entropy Analysis for MHD Flow Subject to Temperature-Dependent Viscosity and Thermal Conductivity. In: Banerjee, S., Saha, A. (eds) Nonlinear Dynamics and Applications. Springer Proceedings in Complexity. Springer, Cham. 2022. https://doi.org/10.1007/978-3-030-99792-2_38
21. Cramer KR, Pai SI. Magnetofluid Dynamics for Engineers and Applied Physicists. McGraw-Hill, New York, 1973.
22. Das S, Jana RN. Natural convective magnetonanofluid flow and radiative heat transfer past a moving vertical plate. Alexandria Engineering Journal. 2015; 54:55–64.
23. Reddy PC, Raju MC, Raju GSS. Free Convective Heat and Mass Transfer Flow of Heat-Generating Nanofluid past a vertical moving porous plate in a Conducting field. Special Topics & Reviews in Porous Media – An International Journal. 2016;7(2): 161-180. 10.1615/SpecialTopicsRevPorousMedia.2016016973

Nomenclature

u' = Velocity component along the coordinate axis $(\mathrm{m\ sec^{-1}})$
T' = Temperature (K)
T'_∞ = Ambient temperature (K)
T'_w = Temperature at the plate (K)
C' = Concentration $(\mathrm{kg\ m^{-3}})$
C'_∞ = Ambient species concentration $(\mathrm{kg\ m^{-3}})$
C'_w = Concentration at the plate $(\mathrm{kg\ m^{-3}})$
t' = Time (sec)
$\overrightarrow{B_O}$ = Magnetic field (Tesla)
C_p = Specific heat at constant pressure $(J\ K^{-1}\mathrm{kg}^{-1})$
ρ_{nf} = Density of the nanofluid $(\mathrm{kg\ m^{-3}})$
μ_{nf} = Dynamic viscosity of the nanofluid (Pa s)
$(\rho C_p)_{nf}$ = Heat capacitance of the nanofluid $(J\ K^{-1})$
k_{nf} = Thermal conductivity of the nanofluid (W $m^{-1}K^{-1}$)
σ_{nf} = Electrical conductivity of the nanofluid $(\mathrm{ohm^{-1}s^{-1}})$
ρ_f = Density of the base fluid $(\mathrm{kg\ m^{-3}})$
ρ_s = Density of the nanoparticles $(\mathrm{kg\ m^{-3}})$
σ_f = Electrical conductivity of the base fluid $(\mathrm{ohm^{-1}s^{-1}})$
σ_s = Electrical conductivity of the nanoparticles $(\mathrm{ohm^{-1}s^{-1}})$
μ_f = Viscosity of the base fluid nanoparticles $(\mathrm{Pa\ s})$
D = Mass diffusion coefficient $(m^2\ \mathrm{sec^{-1}})$
D_1 =Thermal diffusion coefficient $(\mathrm{m^2\ sec^{-1}})$
g = Acceleration due to gravity $(\mathrm{m\ sec^{-2}})$
ϕ = Solid volume fraction of the nanoparticles $(\mathrm{m\ g})$
β_{nf} = Thermal expansion coefficient (K^{-1})
β'_{nf} = Mass transfer coefficient $(\mathrm{m\ sec^{-1}})$
x' = Cartesian coordinate (m)
K'_p = Porosity medium permeability coefficient
q_o = Dimensional heat generation/absorption coefficient
K_1 = Chemical reaction rate

Superscript

′ Dimensional

Non-dimensional parameters

$Gr = \frac{g\beta\mu_f(T'_w-T'_\infty)}{u_0^3}$ (Grashof number)
$Gc = \frac{g\beta_f\mu_f(C'_w-C'_\infty)}{u_0^3}$ (modified Grashof number)
$M^2 = \frac{\sigma_f B_0^2 v_f}{\rho_f u_0^2}$ (magnetic parameter)
$K = \frac{K'_p u_0^2}{v_f^2}$ (porosity parameter)
$Pr = \frac{\mu_f C_p}{k_f}$ (Prandtl number)
$Ec = \frac{u_0^2}{C_p(T'_w-T'_\infty)}$ (Eckert number)
$Du = \frac{D}{v_f}\left(\frac{C'_w-C'_\infty}{T'_w-T'_\infty}\right)$ (Dufour number)
$S_c = \frac{v_f}{D}$ (Schmidt number)
$S_o = \frac{D_1}{v_f}\left(\frac{T'_w-T'_\infty}{C'_w-C'_\infty}\right)$ (Soret number)
$S_c = \frac{v_f}{D}$ (Schmidt number)
$Kr = \frac{K_1 v_f}{u_0^2}$ (chemical reaction parameter)
$Q = \frac{q_0 v_f^2}{k_f u_0^2}$ (heat source/sink parameter)

Appendix

Code for the numerical method: -
For differential equations:

```
global D s W C E M A K G B F A5 P R n Q
T
 dydx= [y (2)
        C*W*E*P*(y (5) ^2) +C*W*P*((n/P)
*Q-s) *y(1)-C*(s-T) *y(3)/(1-C*W*D*P)
        y (4)
        P*((-D*C*W*E*P*(y (6) ^2)
+D*C*W*P*(n/P*Q-s) *y(1)-D*C*(s-T)
*y(3)/(1-C*W*D*P))-E*(y (6) ^2) -((n/P)
*Q-s) *y (1))
        y (6)
        (s+(M^2) *A+1/K) *y (5)-G*B*y
(1)-
F*A5*y (3)];
```

For boundary conditions:

```
global L
res= [ya (1)-1
      ya (3)-1
      ya (5)-L
      yb (1)
      yb (3)
      yb (5)];
```

Aastha Aastha: https://orcid.org/0000-0002-1175-361X

Khem Chand: https://orcid.org/0000-0002-6360-7729

MULTIBODY DYNAMICS MODEL OF THE CYCLOIDAL GEARBOX, IMPLEMENTED IN FORTRAN FOR ANALYSIS OF DYNAMIC PARAMETERS INFLUENCED BY THE BACKLASH AS A DESIGN TOLERANCE

Roman KRÓL*, **Kazimierz KRÓL***

*Faculty of Mechanical Engineering, Department of Applied Mechanics and Mechatronics,
University of Technology and Humanities in Radom, ul. Stasieckiego 54, 26-600 Radom, Poland

r.krol@uthrad.pl, k.krol@uthrad.pl

received 10 October 2022 revised 16 January 2023 accepted 5 February 2023

Abstract: In this study, dynamical parameters of the cycloidal gearbox working at the constant angular velocity of the input shaft were investigated in the multibody dynamics 2D model implemented in the Fortran programming language. Time courses of input and output torques and forces acting on the internal and external sleeves have been shown as a function of the contact modelling parameters and backlash. The analysis results in the model implemented in Fortran were compared with the results in the 3D model designed using MSC Adams software. The values of contact forces are similar in both models. However, in the time courses obtained in MSC Adams there are numerical singularities in the form of peaks reaching 500 N for the forces at external sleeves and 400 N for the forces acting at internal sleeves, whereas in the Fortran model, there are fewer singularities and the maximum values of contact forces at internal and external sleeves do not exceed 200 N. The contact damping and discretisation level (the number of discrete contact points on the cycloidal wheels) significantly affect the accuracy of the results. The accuracy of computations improves when contact damping and discretisation are high. The disadvantage of the high discretization is the extended analysis time. High backlash values lead to a rise in contact forces and a decrease in the force acting time. The model implemented in Fortran gives a fast solution and performs well in the gearbox optimisation process. A reduction of cycloidal wheel discretisation to 600 points, which still allows satisfactory analysis, could reduce the solution time to 4 min, corresponding to an analysis time of 0.6 s with an angular velocity of the input shaft of 52.34 rad/s (500 RPM).

Keywords: cycloidal gearbox, backlash, contact modelling, multibody dynamics, output torque ripple, gearbox efficiency

1. INTRODUCTION

Cycloidal gearboxes find wide application in the drives of robotic systems and the winches of rescue helicopters or off-road vehicles. The cycloidal gearbox design analysed in this study is characterised by the vibration of the output shaft with significant amplitude. The variable angular velocity of the output shaft (Fig. 1) is the source of vibrations, which emerge from the given geometry of the gearbox.

Contemporary research in cycloidal gearbox engineering concerns the construction of discrete models, which contain springs and dampers [1], [2] or modelling of the cycloidal gearbox using engineering software [2], [3]. In the iterative analysis, the current loading state of the cycloidal gearbox depends on previous loading cycles. Excluding back iterations, applying the discrete models for the dynamic analysis gives low-accuracy results. Analysis of the influence of the various contact models on the torque at the output shaft requires constructing transient models based on multibody dynamics. Time courses of the forces and moments obtained in the analysis find their application in developing fault diagnosis methodologies and optimisation. Current research in this area includes fault diagnosis of planetary drives [4–6], while research that concerns cycloidal gearboxes is rare.

Studies in the field of cycloidal gearbox engineering deal with the subject of design issues [7–13], finite element analysis of the cycloidal gearbox parts [7,13–17], kinematical analysis [16,18], applications in robotics [12,16,17,19], friction, lubrication and machining [20–24], measurements [25], efficiency [26,27], optimisation [8,28], contact modelling [29–32], backlash and design tolerances [9,30,33–35] or vibration and dynamics [31,32,36–40].

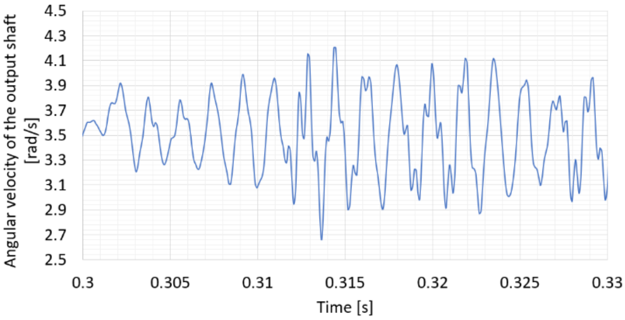


Fig.1. Variable angular velocity of the output shaft

Analysis of the backlash and design tolerances in the cycloidal gearboxes, which deal with the transient analysis, is scarce because it demands programming of the dynamic transient model based on multibody dynamics. Furthermore, high complexity and difficulties in setting up relevant values of the contact modelling parameters (contact damping and contact stiffness, which could be different for various parts of the gearbox being in contact) in these models allow simplified, static analysis methods to be more attractive.

The motivation to implement the presented transient model was a comprehensive study [32] describing methods of contact modelling in cycloidal gearboxes and showing the time courses of the dynamic entities in the analysed gearbox. The geometry of the cycloidal gearbox presented in Ref [32] is different than that presented in the current article. The difference is also in the adaptive time stepping used in the analysis [32]. The Fortran model described in this article uses the constant time step integrator, simplifying the analysis. Guidelines in contact modelling presented in Ref [32] are a good starting point in programming models of cycloidal gearboxes.

This article describes the transient model [41] programmed in Fortran for fast analysis of the cycloidal gearbox. The possible applications of this model are developing fault diagnosis methods, vibration analysis, contact stress optimisation or analysis of contact models. The authors analysed the backlash in the form of design tolerance, contact stiffness and relative velocity of the contact points in the scope of contact modelling methodology. In addition, the time courses of the forces and moments in the cycloidal gearbox were shown as a function of contact damping.

2. MODELS OF THE CYCLOIDAL GEARBOX

Two models of the cycloidal gearbox were built: the first model was designed using MSC Adams software, and the second model was programmed based on multibody dynamics in Fortran. The second model was also implemented using Matlab software. Unfortunately, the adaptive time step was time-consuming. The version implemented in Fortran, which is a fast structural programming language with computer algebra facilities, uses the 2nd-order Runge Kutta method with constant integration steps. Both models consisted of input and output shafts, two cycloidal wheels and 16 external and eight internal sleeves (Fig. 2). All bodies were mounted on bearings, which allowed relative rotational motions between the parts. The shape of the cycloidal wheels is described in the system of parametric Eq. (1); the parameters from the given system are presented in Tab. 1, the masses and moments of inertia of bodies are given in Tab. 2 and loads are given in Tab. 3.

In the model programmed in Fortran (Fig. 2b), visualisation software (programmed in Java language with the OpenGL library) does not show the input shaft with the eccentric cams on which cycloidal wheels are mounted. The analysed gearbox was designed with the dimensions presented in Fig. 3.

$$\begin{cases} u(\alpha) = \frac{e \cdot z_k}{m}\cos(\alpha) + e \cdot \cos(z_k \cdot \alpha) - q \cdot \cos(\alpha + \gamma) \\ v(\alpha) = \frac{e \cdot z_k}{m}\sin(\alpha) + e \cdot \sin(z_k \cdot \alpha) - q \cdot \sin(\alpha + \gamma) \\ \gamma = \operatorname{atan}\left[\frac{\sin(z_s \cdot \alpha)}{\frac{1}{m} + \cos(z_s \cdot \alpha)}\right] \end{cases}, \quad (1)$$

The input shaft rotates with a constant angular velocity, as given in Tab. 3. The output shaft is loaded by the constant output torque (Tab. 3). Despite the cycloidal gearbox being loaded by a constant output torque, the torque at the output shaft is variable due to the geometry of the gearbox. In Figs. 7b, 15 and 16, the output torque was solved based on the time courses of the forces acting on the internal sleeves and the time courses of the displacements of these sleeves.

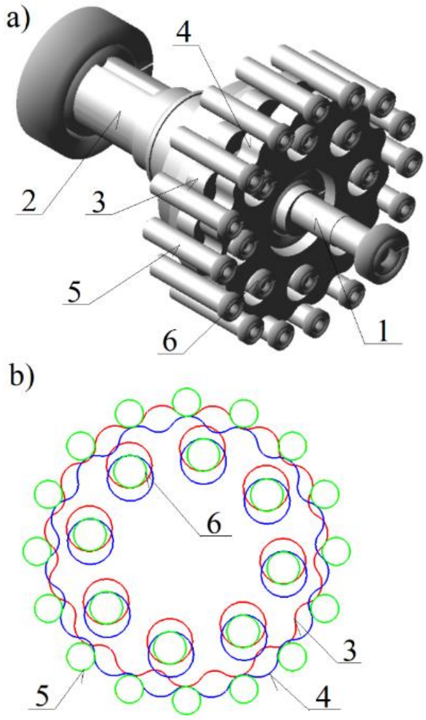


Fig. 2. Cycloidal gearbox models: designed in MSC Adams (a) and programmed in Fortran (b), where: 1 denotes the input shaft, 2 denotes the output shaft, 3 denotes the internal cycloidal wheel, 4 denotes the external cycloidal wheel, 5 denotes the external sleeve, 6 denotes the internal sleeve

Tab. 1. Parameters of the cycloidal wheels defined in parametric Eq. (1)

Parameter	Description	Value
u(α)	Horizontal coordinate	-
v(α)	Vertical coordinate	-
α [rad]	Equation parameter	0-2π
e [m]	Eccentricity	0.0028
z_k	Number of external sleeves	16
z_s	Number of lobes	15
m	Short-width coefficient	0.7
q [m]	Radius of the external sleeve	0.006

Tab. 2. Masses and moments of inertia relative to the local coordinate systems. The local coordinate systems are placed in the parts' centre of gravity. The parts are rigid bodies

Body	Model parameters	
	Mass [kg]	Moment of inertia [kg·m^2]
Input shaft	0.2341	$1.846 \cdot 10^{-6}$
Output shaft	1.6345	$1.373 \cdot 10^{-4}$
Cycloidal wheel	0.5998	$1.57538 \cdot 10^{-4}$
Internal sleeve	0.048	$1.1013 \cdot 10^{-5}$
External sleeve	0.048	$1.1013 \cdot 10^{-5}$

Tab. 3. Loads and excitations in the models

Entity	Value
Angular velocity of the input shaft [Hz; rad/s; RPM]	8.33; 52.34; 500
Torque applied at the output shaft [Nm]	22

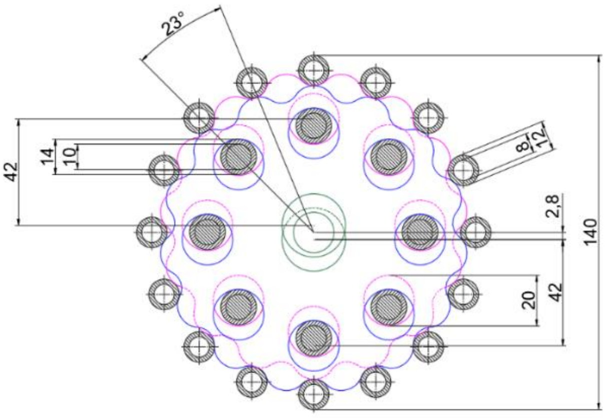


Fig. 3. Dimensions of the cycloidal gearbox analysed using MSC Adams and the model programmed in Fortran

3. MULTIBODY DYNAMICS METHODS USED IN THE MODEL PROGRAMMED IN FORTRAN

The following system of Eq. (2) is solved in each integration subroutine step. The model contains 28 bodies and 29 joints, one of which is a driver joint and the others are revolute joints. The joints appear in the system of Eq. (2) as constraint equations, which are components of the Jacobian matrix. The method of Jacobian matrix formulation is described in [42], [43].

$$\begin{bmatrix} M & -D^T \\ D & 0 \end{bmatrix} \begin{bmatrix} \ddot{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} F \\ \gamma \end{bmatrix}, \tag{2}$$

where M is the mass matrix (contains masses and moments of inertia of the parts), D is the Jacobian matrix, 0 is the zero matrix, $\ddot{u}$ denotes solved accelerations, λ denotes solved reaction forces, F is the vector of applied forces and moments and γ is the right-hand side vector of the acceleration equations.

In the described model, the body-coordinate formulation was used. The constraint equations for the revolute joint and the drive joint are given in Eqs. (3) and (4), respectively. The Jacobian matrix satisfies (5), where the components of this matrix depend on the type of joint. Submatrices of the Jacobian matrix for the revolute joint are given in Eqs. (6) and (7) and those of the driver joint in Eq. (8).

$$\Phi_{rev} = r_{P1} - r_{P2} = 0, \tag{3}$$

where Φ_{rev} is the revolute joint constraint equation and r_{P1} and r_{P2} are the position vector of the revolute joint attachment point on the 1st and the 2nd body, respectively. The position vectors are given in the global coordinate system.

$$\Phi_{driver} = \varphi - f(t) = 0, \tag{4}$$

where Φ_{driver} is the driver joint constraint equation, φ is the initial angle of a relative rotation of the connected parts and f(t) is the time course of the rotation angle between the connected parts (i.e. f(t)=ωt for the simulation of the motion with the constant angular velocity ω).

$$D = \frac{\partial \Phi}{\partial u}, \tag{5}$$

where D is the Jacobian matrix, Φ is the constraint equation and u is the vector of body coordinates.

$$D_1 = \begin{bmatrix} -1 & 0 & -\check{s}_{1X} \\ 0 & -1 & -\check{s}_{1Y} \end{bmatrix}, \tag{6}$$

$$D_2 = \begin{bmatrix} 1 & 0 & \check{s}_{2X} \\ 0 & 1 & \check{s}_{2Y} \end{bmatrix}, \tag{7}$$

where D1 and D2 are the Jacobian submatrix corresponding to the 1st and the 2nd body connected to the revolute joint, respectively; $\check{s}$ is the vector orthogonal to the joint attachment point local vector, relative to the centre of mass and X and Y are coordinates of the orthogonal vector.

$$D_{\mathrm{drv}} = [0 \quad 0 \quad 1], \tag{8}$$

where D_{drv} is the Jacobian submatrix corresponding to the driven body. The right-hand side acceleration equations vector contains 0 in the components corresponding to the driver-joint. For the revolute joint, the right-hand side acceleration vector is solved according to Eq. (9).

$$\gamma = \dot{\check{s}}_1 \dot{\phi}_1 - \dot{\check{s}}_2 \dot{\phi}_2, \tag{9}$$

where $\dot{\phi}_1$ and $\dot{\phi}_2$ are angular velocities of the 1st and the 2nd body attached to the revolute joint.

4. CONTACT MODELLING

Methods presented in the previous chapter were used to model three groups of parts in the cycloidal gearbox: (1) the input shaft with cycloidal wheels, (2) the output shaft with internal sleeves and (3) external sleeves. The gearbox should transfer the moment from the input shaft to the output shaft. Therefore contact should be set between mentioned parts. The Kelvin Voigt contact model based on the Hertzian theory with energy dissipation was programmed in four subroutines, which check the contact between two cycloidal wheels, external sleeves and internal sleeves. Subroutines that control the contact between the cycloidal wheels and the external sleeves solve the cycloidal wheel curvature in the contact point based on circular approximation from the three neighbouring points. Subroutines that check the contact between the cycloidal wheels and the internal sleeves use a constant stiffness coefficient because contact surfaces have a constant curvature radius.

Cycloidal wheels and their holes were discretised in the contact checking subroutines. The contact points on the cycloidal wheel boundary (Fig. 4) are placed with the constant angular increment α in Eq. (1). Convergence was obtained for 600 (rough analysis) and 4,000 (refined analysis) points in each cycloidal wheel and 20 (rough analysis) and 100 (refined analysis) in the cycloidal wheel holes in which the internal sleeves occur.

In the contact checking subroutine, the list of points is created whose distance to the centre of the sleeve is less than the sleeve's radius. The subroutine selects the point closest to the sleeve centre from the list of points being in contact. Two points, W and Q, detected by the contact checking subroutine are shown in Fig. 4.

Given the external sleeve centre of mass coordinates and point W (Fig. 4) coordinates, the contact normal vector is computed according to Eq. (10):

$$\vec{\boldsymbol{n}} = [W_x - P_{nX}, W_y - P_{nY}], \tag{10}$$

where W_X and W_Y are coordinates of the selected cycloidal wheel point being in contact with the external sleeve and P_{nX} and P_{nY} are coordinates of the given external sleeve centre of mass. To solve the friction force, the tangent contact vector (11) is solved using the rotation matrix R₉₀ (12):

$$\vec{t} = R_{90} \cdot \|\vec{n}\|, \tag{11}$$

$$R_{90} = \begin{bmatrix} \cos(90°) & -\sin(90°) \\ \sin(90°) & \cos(90°) \end{bmatrix}, \tag{12}$$

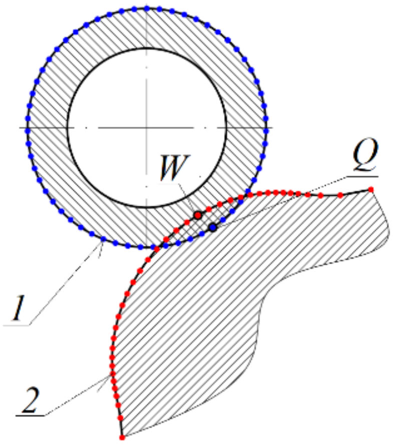


Fig. 4. Two points, W and Q, detected by the contact checking subroutine. 1 denotes the external sleeve, 2 denotes the cycloidal wheel, Q is a point on the external sleeve and W is a point on the cycloidal wheel

The vector r_{pQ} of the contact point Q position relative to the global coordinate system can be computed using Eq. (13). The linear velocity $\dot{r}_{cW}$ of the point W can be computed using Eq. (14) with the cycloidal wheel rotation matrix specified in Eq. (15):

$$r_{pQ} = r_p + R_{exsl} \cdot \|\vec{n}\|, \tag{13}$$

$$\dot{r}_{cW} = \dot{r}_c + \dot{A}_c \cdot s_{cW}, \tag{14}$$

$$A_c = \begin{bmatrix} \cos(\varphi_c) & -\sin(\varphi_c) \\ \sin(\varphi_c) & \cos(\varphi_c) \end{bmatrix}, \tag{15}$$

where r_p is the vector of the external sleeves' centre of mass localisation relative to the global coordinate system, R_{exsl} is the radius of the external sleeve, $\|\vec{n}\|$ is the contact normal vector with normalised length, A_C is the rotation matrix of the cycloidal wheel, φ_c is the angle of cycloidal wheel rotation, s_{cW} is the local vector of the contact point W position relative to the cycloidal wheel centre of mass and $\dot{r}_c$ is the linear velocity vector of the cycloidal wheels' centre of mass.

The normalised contact normal velocity is given in Eq. (16). The linear velocity (17) of the Q point can be computed by differentiating Eq. (13) and substituting Eq. (16). Contact normal velocity and contact tangential velocity can be solved using Eqs. (18) and (19) based on Eqs. (10), (11), (14) and (17):

$$\|\dot{\vec{n}}\| = \dot{r}_{cW} - \dot{r}_p, \tag{16}$$

$$\dot{r}_{pQ} = \dot{r}_p + R_{exsl} \cdot \|\dot{\vec{n}}\|, \tag{17}$$

$$v_n = (\dot{r}_{pQ} - \dot{r}_{cW}) \cdot \|\vec{n}\|, \tag{18}$$

$$v_t = (\dot{r}_{pQ} - \dot{r}_{cW}) \cdot \|\vec{t}\|, \tag{19}$$

where v_n is the contact normal velocity and v_t is the contact tangential velocity. The theory that concerns velocity computation is discussed in detail in Refs [32], [44].

The contact normal velocity is used for the computation of the contact normal force (20), and the contact tangential velocity is used to compute the friction force (21). Contact normal force can be computed using various methods presented in Ref [44]. In the Fortran model, the models of Lee and Wang or Herbert and McWhannel [44] were implemented. Unfortunately, the convergence was not obtained for the specified models designed for sphere-to-sphere contact and using initial contact velocity. Instead, the approach from MSC Adams documentation [45], [46] was used, and the Heaviside function was utilised in the modelling of the contact stiffness and contact damping (20). The Heaviside function S(t-t0) (22) is linear from 0 to t0.

$$F_n = S(\delta - 0.00001) \cdot K \cdot \delta^{1.5} + \\ S(\delta - 0.00001) \cdot C \cdot v_n, \tag{20}$$

$$F_t = -\mu \cdot S(v_t - 0.00005) \cdot F_n, \tag{21}$$

where S is the Heaviside step function (22), K denotes contact stiffness, C denotes contact damping, v_n is the contact normal velocity, v_t is the contact tangential velocity, μ is the static coefficient of friction and δ is the penetration depth (the distance between W and Q points):

$$S(t - t_0) = \begin{cases} 0\, for\, t < 0 \\ \frac{1}{t_0} \cdot t\, for\, t \in (0, t_0). \\ 1\, for\, t > t_0 \end{cases} \tag{22}$$

The stiffness coefficient (23) [32], [44] is solved based on the surface curvature of the contacting bodies and the parameters given in Eq. (24). The contact stiffness coefficient was multiplied by the H multiplier. The values of the H multiplier and contact damping C are presented in Tab. 4.

$$K = \frac{4H}{3\pi(h_c + h_p)} \sqrt{\frac{R_c \cdot R_p}{R_c \pm R_p}}, \tag{23}$$

$$h_c = \frac{1 - \nu_c^2}{\pi \cdot E_c}, h_p = \frac{1 - \nu_p^2}{\pi \cdot E_p}, \tag{24}$$

where K denotes contact stiffness; H is the contact stiffness multiplier; R_c is the radius of curvature of the cycloidal wheel surface; R_p is the radius of curvature of the external sleeve surface; v_c and v_p are Poisson ratios of the cycloidal wheel and the external sleeve material, respectively; and E_c and E_p are Young moduli of the cycloidal wheel and the external sleeve, respectively. The sign in the denominator between R_c and R_p values depends on the contact with the cycloidal wheel pit or the cycloidal wheel lobe (the negative or positive radius of curvature of the cycloidal wheel).

Tab. 4. Parameters used in contact modelling

Parameter	Values used by contact detection subroutines (external and internal sleeves)
v_c, v_p	0.3
E_c, E_p [N/m^2]	$2 \cdot 10^{11}$
C [Ns/m]	8
H	0.01

Computed values of the forces and moments acting on the external sleeves, internal sleeves and cycloidal wheels are substituted to the system of Eq. (2) in the F vector in each iteration of the 2nd order Runge Kutta algorithm. In the solution process, the constant integration step was used with 10^{-5} s. The analysis time was set to 0.6 s.

5. RESULTS OF THE ANALYSIS IN THE FORTRAN MODEL

One of the benefits of the transient model programmed in Fortran is the possibility to analyse a graph of every entity in the simulation process. MSC Adams allows the user to enter a constant value of the contact stiffness. In the Fortran model, the value of the contact stiffness depends on the curvature of contacting bodies at the point of contact. Contact stiffness at the moment of contact between the cycloidal wheel and one of the external sleeves is shown in Fig. 5. In Fig. 6 the relative velocity of the contact points is shown at the time of contact with the same sleeve. Figs. 5 and 6 show periodically changing values of contact stiffness and relative velocity, which can differ in detail for the following periods. These differences arise from the numerical solution process and its inaccuracies. The mentioned entities have been shown in the time range when the contact between the external sleeve and the cycloidal wheel is active. In the rest of the period, the parts do not interfere with each other.

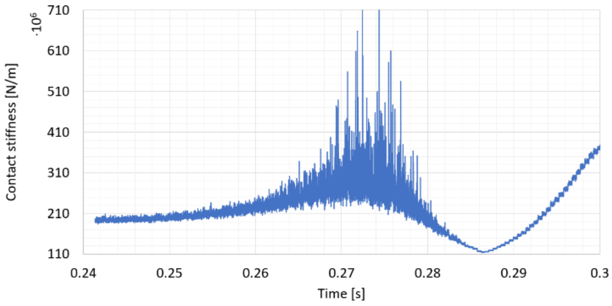


Fig. 5. Contact stiffness in the time range, when contact between the external sleeve and cycloidal wheel occurs. Analysis without backlash

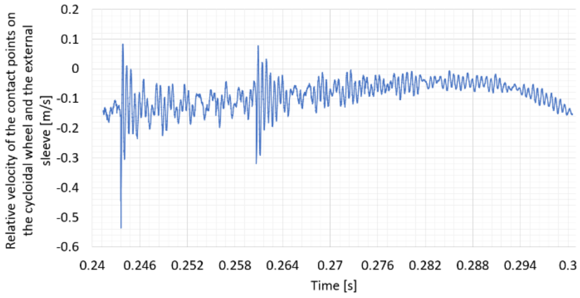


Fig. 6. Relative velocity of the contact points on the cycloidal wheel and the external sleeve. Analysis without backlash

Time courses of the input torque and output torque are shown in Figs. 7a and 7b respectively. It can be seen that despite the cycloidal gearbox being loaded by the constant output torque, the courses of the input and output torques oscillate due to the geometry. Forces acting on the internal and external sleeves are shown in Figs. 8a and 8b, respectively. Figs. 7 and 8 depict time courses of the dynamical entities for different values of the contact damping. This parameter significantly influences the accuracy of results. The presented time courses oscillate with high amplitude for low contact damping values. Structural damping for steel is considered to be less than 0.01 of the critical damping. The presented analysis modelled contact as the Kelvin Voigt model with spring and damper (viscous damping). For accurate contact modelling in the numerical analysis, according to Ref [46], the damping should be set to one per cent of the contact stiffness coefficient (23).

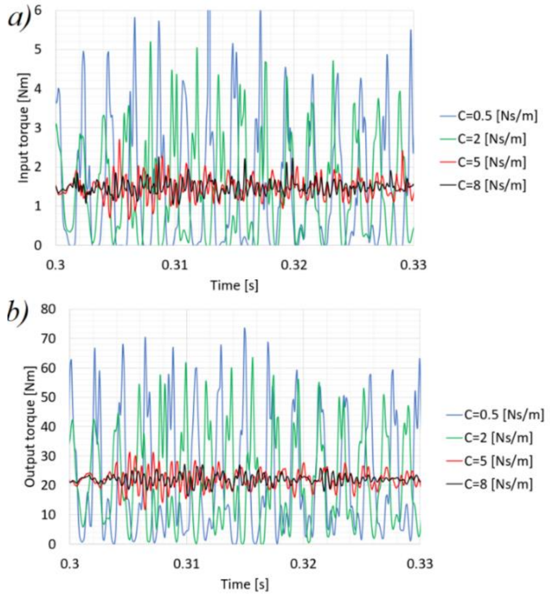


Fig. 7. Physical entities computed in the model programmed in Fortran: torque at the input shaft (a), torque at the output shaft (b)

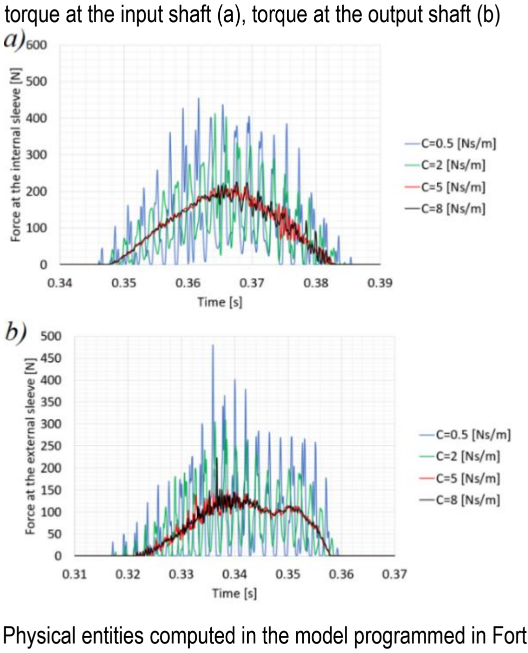


Fig. 8. Physical entities computed in the model programmed in Fortran: the force acting on the internal sleeve (magnitude) (a), the force acting on the external sleeve (magnitude) (b)

In the analysis presented in the current article, an increase in the contact damping above 8 Ns/m leads to non-convergence. The choice of the specified damping factor is motivated by the methodology of the contact modelling and is not related to the specific material.

Fig. 9 (a, b) presents a comparison of the force acting on the internal sleeve and the external sleeve obtained with rough discretisation (600 points for each cycloidal wheel) and refined analysis (4,000 points for each cycloidal wheel). The results obtained in MSC Adams were compared with the results from the transient model programmed in Fortran in Fig. 10 (a, b).

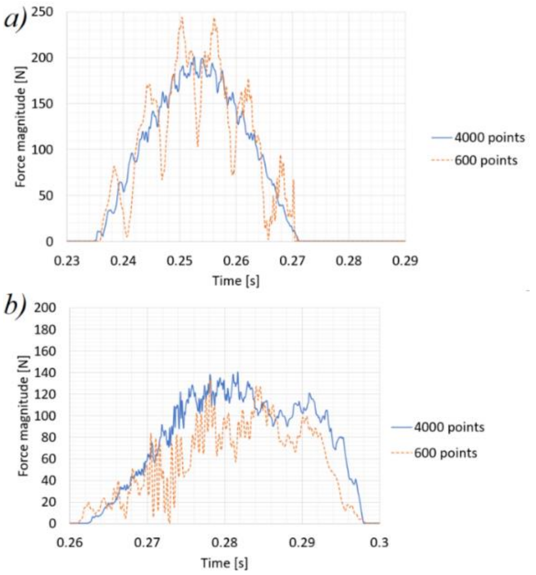


Fig. 9. Results computed in the model implemented in Fortran for various numbers of points in the cycloidal wheel (600 points and 4000 points): force acting on the internal sleeve (magnitude) (a), force acting on the external sleeve (magnitude) (b)

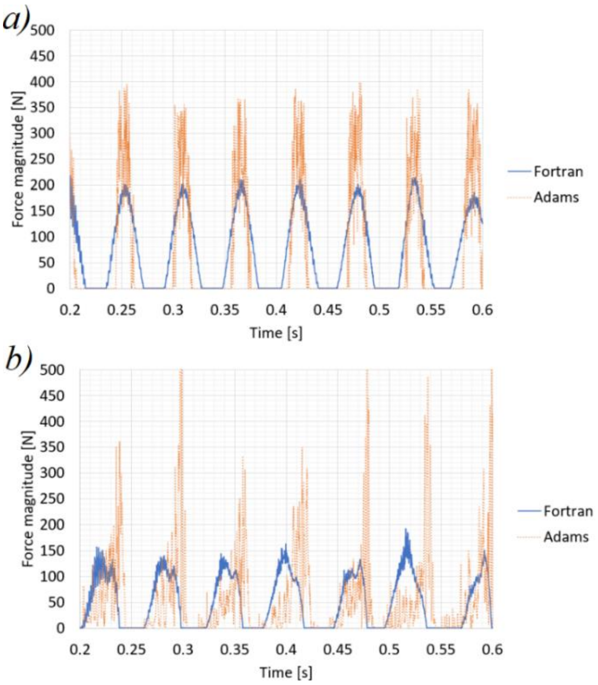


Fig. 10. Results from MSC Adams and the model implemented in Fortran: force acting on the internal sleeve (magnitude) (a) and force acting on the external sleeve (magnitude) (b)

6. BACKLASH INFLUENCE ON THE DYNAMIC PARAMETERS OF THE CYCLOIDAL GEARBOX

Analysis of backlash was performed in the model implemented in Fortran. In the analysed cycloidal gearbox, the artificial backlash was introduced in the form of design tolerance. The design tolerance value *d* increased the external sleeves' radial position. In the analysis without backlash, the radius of their localisation is *r*. After the introduction of backlash, it is *r+d* (Fig. 11).

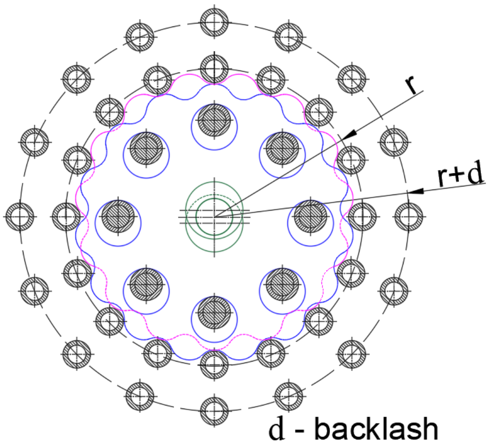


Fig. 11. Backlash in the cycloidal gearbox modelled as the design tolerance d of the radial position r of the external sleeves

The backlash has a decisive impact on the dynamic entities of the cycloidal gearbox (forces acting on the sleeves and torques at the shafts). An increase in backlash significantly influences the force at the external sleeve during cycloidal gearbox work. The higher the backlash, the shorter the contact duration between the external sleeve and the cycloidal wheel. High backlash ($0.6 \cdot 10^{-3}$ m) can increase the force magnitude by 50% relative to the force in the model without backlash (Fig. 12).

There is less influence of the backlash on the value of the force acting on the internal sleeve. However, for the high backlash, a higher amplitude of the force at the internal sleeve oscillations can temporarily arise (Fig. 13).

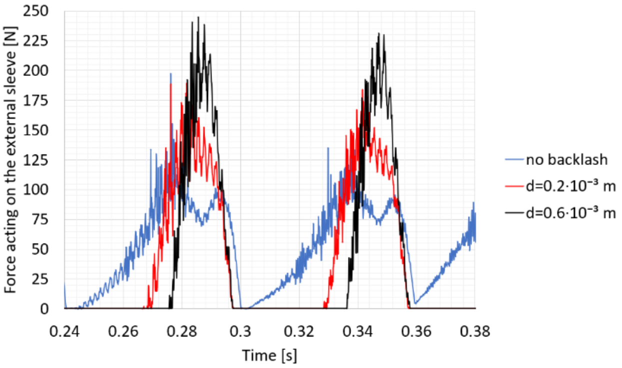


Fig. 12. Contact forces acting on the external sleeves as a function of backlash

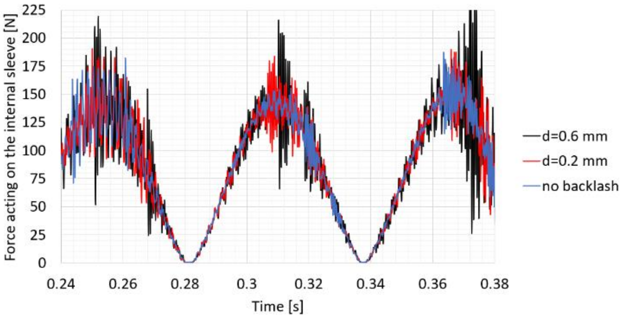


Fig. 13. Contact forces acting on the internal sleeves as a function of backlash

High backlash is unwanted in robotics applications, where reducing the amplitude of oscillations is desirable. Figs. 14 and 15 show an increase in the amplitude of the vibrations for the models with backlash. As mentioned before, in the cycloidal gearbox, the output torque's current state depends on the previous iterations of the analysis. Therefore, the output torque can significantly change its amplitude of oscillations in time. Figs. 14 and 15 present that the amplitude of the vibrations in the analysis without backlash is considerably lower than that with backlash, in general. Fig. 16 shows a higher range of the time in the torque at the output shaft. The torques shown in Figs. 14-16 are random time functions with an unpredictable amplitude, which does not reach a steady state.

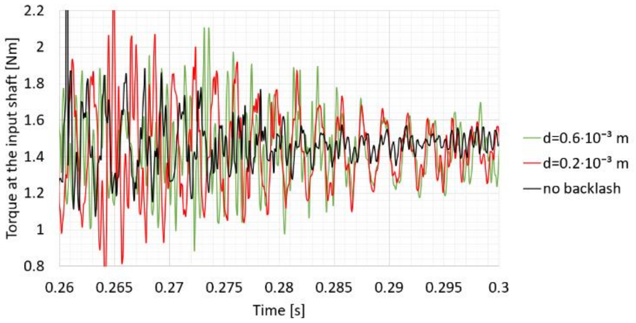


Fig. 14. Torque at the input shaft as a function of backlash

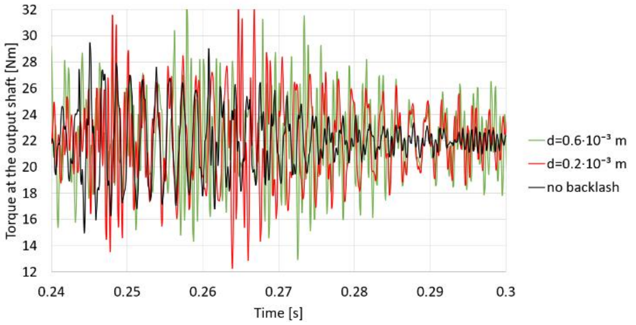


Fig. 15. Torque at the output shaft as a function of backlash

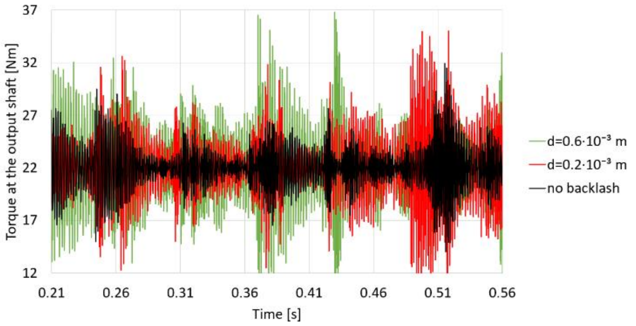


Fig. 16. Torque at the output shaft as a function of backlash. Over a wider time range, the function was shown to be random

In Fig. 17, the time courses of the contact stiffnesses are shown for various backlash values. For various backlashes, multiple shifts in time of the contact stiffness are obtained. It is related to the contact periods between the external sleeve and the cycloidal wheel.

Relative velocities of the contact points (Fig. 4, points W and Q) for various values of backlash are shown in Fig. 18. For each relative velocity, its time course starts with oscillations. However, the time courses for the multiple backlashes overlap in the corresponding time ranges. It shows that for high backlash, the contact between the cycloidal wheel and the external sleeve starts with a higher velocity, which leads to a higher impact and higher value of contact force (Fig. 12).

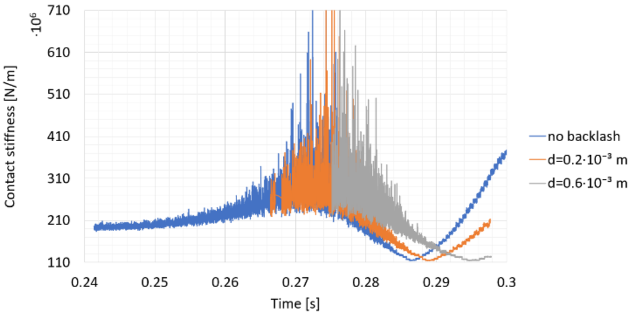


Fig. 17. Contact stiffness between the cycloidal wheel and the external sleeves as a function of backlash

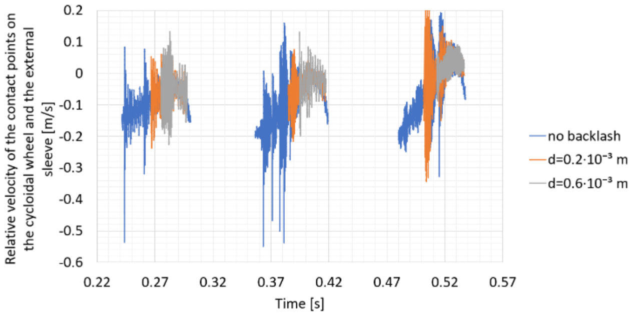


Fig. 18. Relative velocity of the contact points on the cycloidal wheel and the external sleeve as a function of backlash

7. DISCUSSION OF RESULTS

The presented model implemented in Fortran offers a short analysis time, which allows using this model in the optimisation processes. For rough discretisation, the solution time is 4 min, while for refined analysis it is 25 min. Analysis results from MSC Adams presented in Fig. 10 (a, b) require more than 1 h for the solution process.

In the model implemented in Fortran, various contact force models [44] were tested, but convergence was not obtained. These models were designed for two contacting spheres, while the contact in the analysed cycloidal gearbox is cylinder to cylinder. Therefore, the contact stiffness solved according to the Hertzian model should be multiplied by the coefficient H=0.01 to satisfy convergence conditions.

Analysis of various damping coefficients (Figs. 7 and 8) shows that contact damping significantly influences the quality of results. High oscillations emerge for low contact damping values, while more increased contact damping values guarantee smooth results. The MSC Adams results presented in Fig. 10 (a, b) show significant oscillations in the force diagrams for the internal and external sleeves. It can arise from the improper value of the contact damping coefficient, which was set to the same value in the model implemented in Fortran. Analysis of the 3D model in MSC Adams requires another set of contact parameters (contact stiffness, contact damping and the penetration depth at which the contact damping coefficient reaches maximum value).

The model implemented in Fortran can be used in the optimisation process. Therefore it is beneficial to reduce the solution time of the single simulation. Unfortunately, Fig. 9 (a, b) presents that application of the discretisation, which is too rough leads to inaccurate results. The results for the 600 points of discretisation show significant oscillations of the force acting on the internal sleeve in Fig. 9a. The level of discretisation depends on the optimisation demand.

The model programmed in Fortran allows for a more straightforward geometry parametrisation. It is much easier to modify the design tolerances of the model programmed in Fortran than to modify those in the engineering software (MSC Adams). Backlash introduced as the design tolerance of the position of the external sleeves has more impact on the forces acting on the external than on the internal sleeves. Time courses of the force acting on the internal sleeve in Fig. 13 for the high values of backlash have short-in-time oscillations at the peak values of force, which can be 25% higher than peak values in the analysis without backlash.

High backlash is unwanted in applications where the vibrations should be reduced. With increased backlash, the cycloidal wheel comes into contact with the external sleeve with a higher velocity (Fig. 18), which leads to more significant impacts and an increase in the values of contact forces. In addition, higher effects on the external sleeves lead to more remarkable changes in the dynamic parameters of the cycloidal gearbox, which in turn lead to higher oscillations of the input and output torques (Figs. 14 and 15).

Multibody dynamics models provide more accurate results than static calculations or discrete models. Unfortunately, these models contain drawbacks of numerical analyses. At the initial stage correction of the initial parameters is performed. However, the rise of external forces at the first iteration of the simulation leads to disruption of the results. The beginning part of the time course of almost every entity in this model is noised. In the other part of the analysis, the model undergoes stabilisation, and the results are more accurate. It is also an issue in using a proper contact model, including contact stiffness and contact damping solution. Rapid changes in external forces occur at the beginning of the contact between the cycloidal wheel and the external sleeve. On the time courses of relative contact velocity or contact stiffness, oscillations with high amplitude can be seen. Fig. 6 presents fluctuations in the range of 0.24-0.246 s at the beginning of the contact. The oscillations can also appear in the rapid changes of the curvature of the contacting bodies (Fig. 5, period 0.265-0.28 s), which is additionally influenced by the discretisation of the cycloidal wheel.

8. CONCLUSIONS

Summing up the results, the model implemented in Fortran guarantees a fast solution and can be successfully used in the optimisation process. The discretisation of the model should be considered when the accuracy of the optimization results is an essential factor. The contact stiffness should be multiplied by the experimentally adjusted coefficient to allow convergence of the solution process. The contact damping coefficient can significantly influence the accuracy of the results. The backlash has an essential impact on contact forces.

The main drawback of the numerical multibody dynamics analyses is oscillations, which could be caused by the initial rapid changes in the model parameters, discretisation and contact modelling methods.

REFERENCES

1. Wikło M, Król R, Olejarczyk K, Kołodziejczyk K. Output torque ripple for a cycloidal gear train. Proc Inst Mech Eng C J Mech Eng Sci 2019;233:7270–81. https://doi.org/10.1177/0954406219841656.
2. Król R. Resonance phenomenon in the single stage cycloidal gearbox. Analysis of vibrations at the output shaft as a function of the external sleeves stiffness. Archive of Mechanical Engineering 2021;68:303–20. https://doi.org/10.24425/ame.2021.137050.
3. Król R. Kinematics and dynamics of the two stage cycloidal gearbox. AUTOBUSY – Technika, Eksploatacja, Systemy Transportowe 2018;19:523–7. https://doi.org/10.24136/atest.2018.125.
4. Plöger DF, Zech P, Rinderknecht S. Vibration signature analysis of commodity planetary gearboxes. Mech Syst Signal Process 2019;119:255–65. https://doi.org/10.1016/j.ymssp.2018.09.014.
5. Lei Y, Han D, Lin J, He Z. Planetary gearbox fault diagnosis using an adaptive stochastic resonance method. Mech Syst Signal Process 2013;38:113–24. https://doi.org/10.1016/j.ymssp.2012.06.021.
6. Wang T, Han Q, Chu F, Feng Z. Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review. Mech Syst Signal Process 2019;126:662–85. https://doi.org/10.1016/j.ymssp.2019.02.051.
7. Naveen P, Kiran R, Siva Sankaram EVS, Bha-radwaj TM. Design, Analysis and Simulation of Compact Cycloidal Drive. Int J Sci Res Sci Eng Technol 2020;7:216–20. https://doi.org/10.32628/ijsrset207547.
8. Król R, Król K. Optymalizacja nieliniowa przekładni cykloidalnej z ograniczeniami równościowymi na wymiary obudowy. In: Pawliczek R, Owsinski R, Łagoda T, editors. Projektowanie, budowa i eksploatacja maszyn cz. 1, vol. 558, Opole: Politechnika Opolska; 2021, p. 95–108.
9. Li T, An X, Deng X, Li J, Li Y. A new tooth profile modification method of cycloidal gears in precision reducers for robots. Applied Sciences 2020;10. https://doi.org/10.3390/app10041266.
10. Kormin TG, Tsumbu JDB. Cycloidal reducer with rotation external ring gear. IOP Conf Ser Mater Sci Eng 2020;971. https://doi.org/10.1088/1757-899X/971/4/042072.
11. Huang X, Zhang J. Analysis of Geometric Characteristics of Cycloidal Transmission. IOP Conf Ser Mater Sci Eng 2020;751:12059. https://doi.org/10.1088/1757-899X/751/1/012059.
12. Huang JT, Li CW. The High-payload Manipulator Development Based on Novel Two-stage Cycloidal Speed Reducers and Hub Motors. J Phys Conf Ser 2020;1583:12002. https://doi.org/10.1088/1742-6596/1583/1/012002.
13. Blagojevic M, Marjanovic N, Djordjevic Z, Stojanovic B, Disic A. A new design of a two-stage cycloidal speed reducer. Journal of Mechanical Design 2011;133. https://doi.org/10.1115/1.4004540.
14. Olejarczyk K, Wiklo M, Król K, Kolodziejczyk K. Cycloidal disc calculation of cycloidal gear using finite element method. Logistyka 2015;6.
15. Blagojevic M, Marjanovic N, Stojanovic B, Blagojević M, Marjanović N, Đorđević Z. Stress And Strain State Of Single-Stage Cy-Cloidal Speed Reducer. The 7th International Conference Research And Development Of Mechanical Elements And Systems Irmes, 2011.
16. Strutynskyi S, Semenchuk R. Investigation of the accuracy of the manipulator of the robotic complex constructed on the basis of cycloidal transmission. Technology Audit and Production Reserves 2021;4:6–14. https://doi.org/10.15587/2706-5448.2021.237326.
17. Chavan U, Joshi A, Kolambe Y, Gwalani H, Chaudhari H, Khalate A, et al. Magnification of energy transmission ratio using miniature cycloidal gear box for humanoids. IOP Conf Ser Mater Sci Eng 2022;1272:012017. https://doi.org/10.1088/1757-899X/1272/1/012017.

18. Blagojevic M, Pantić I, Blagojević M. KINEMATIC ANALYSIS OF SINGLESTAGE CYCLOIDAL SPEED REDUCER. Machine Design 2015;7:113–8.
19. Al Kouzbary M, Al Kouzbary H, Liu J, Khamis T, Al-Hashimi Z, Shasmin HN, et al. Robotic Knee Prosthesis with Cycloidal Gear and Four-Bar Mechanism Optimized Using Particle Swarm Algorithm. Actuators 2022;11. https://doi.org/10.3390/act11090253.
20. Tonoli A, Amati N, Impinna F, Detoni G, Ruzimov S, Gasparin E, et al. Influence of dry friction on the irreversibility of cycloidal speed reducer. 5th World Tribology Congress, WTC 2013, 2013.
21. Luo SM, Liao LX, Mo JY. Prediction of surface roughness of end milling for cycloidal gears based on orthogonal tests. Engineering Transactions 2018;66:339–52. https://doi.org/10.24423/EngTrans.860.20180830.
22. Blagojevic M, Marjanovic N, Stojanovic B, Ivanovic L. Influence of the friction on the cycloidal speed reducer efficiency. Journal of the Balkan Tribological Association 2012;18:217–27.
23. Bo W, Jiaxu W, Guangwu Z, Rongsong Y, Hongjun Z, Tao H. Mixed lubrication analysis of modified cycloidal gear used in the RV reducer. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology 2016;230:121–34. https://doi.org/10.1177/1350650115593301.
24. Vasić M, Blagojević M, Dragoi M. Thermal stability of lubricants in cycloidal reducers. Engineering Today 2022;1:7–17. https://doi.org/10.5937/engtoday2202007v.
25. Zaręba R, Mazur T, Olejarczyk K, Bzinkowski D. Measurement of the Cycloidal Drive Sleeves and Pins. Mechanika 2021;27:505–12. https://doi.org/10.5755/J02.MECH.27815.
26. Petrovskiy AN. Increased efficiency of eccentric cycloidal engagement. Proceedings of Higher Educational Institutions Machine Building 2021:3–14. https://doi.org/10.18698/0536-1044-2021-9-3-14.
27. Olejarczyk K, Wikło M, Kołodziejczyk K, Król R, Król K. Theoretical and experimental verification of one stage cycloidal gearbox efficiency. Advances in Mechanism and Machine Science, vol. 73, Springer Science and Business Media B.V.; 2019, p. 1029–38. https://doi.org/10.1007/978-3-030-20131-9_102.
28. Król R, Wikło M, Olejarczyk K, Kołodziejczyk K, Zieja A. Optimization of the one stage cycloidal gearbox as a non-linear least squares problem. Advances in Mechanism and Machine Science, 2019, p. 1039–48. https://doi.org/10.1007/978-3-030-20131-9_103.
29. Sun X, Han L, Wang J. Tooth modification and loaded tooth contact analysis of China Bearing Reducer. Proc Inst Mech Eng C J Mech Eng Sci 2019;233:6240–61. https://doi.org/10.1177/0954406219858184.
30. Li T, Wang G, Deng X, An X, Xing C, Ma W. Contact Analysis of Cycloidal-pin Gear of RV Reducer under the Influence of Profile Error. J Phys Conf Ser 2019;1168:22095. https://doi.org/10.1088/1742-6596/1168/2/022095.
31. Xu LX. A dynamic model to predict the number of pins to transmit load in a cycloidal reducer with assembling clearance. Proc Inst Mech Eng C J Mech Eng Sci 2019;233:4247–69. https://doi.org/10.1177/0954406218809732.
32. Xu LX, Chen BK, Li CY. Dynamic modelling and contact analysis of bearing-cycloid-pinwheel transmission mechanisms used in joint rotate vector reducers. Mech Mach Theory 2019;137:432–58. https://doi.org/10.1016/j.mechmachtheory.2019.03.035.
33. Król R. Analysis of the backlash in the single stage cycloidal gearbox. Archive of Mechanical Engineering 2022;69:693–711. https://doi.org/10.24425/ame.2022.141521.
34. Csobán A. Impacts of a profile failure of the cycloidal drive of a planetary gear on transmission gear. Lubricants 2021;9. https://doi.org/10.3390/lubricants9070071.
35. Kostić N, Blagojević M, Petrović N, Matejić M, Marjanović N. Determination of real clearances between cycloidal speed reducer elements by the application of heuristic optimization. Transactions of Famena 2018;42:15–26. https://doi.org/10.21278/TOF.42102.
36. Blagojević M, Matejić M, Kostić N. Dynamic behaviour of a two-stage cycloidal speed reducer of a new design concept. Tehnicki Vjesnik 2018;25:291–8. https://doi.org/10.17559/TV-20160530144431.
37. Wikło M, Krzysztof O, Krzysztof K, Król K, Komorska I. Experimental vibration test of the cycloidal gearbox with different working conditions. Vibroengineering Procedia, vol. 13, EXTRICA; 2017, p. 24–7. https://doi.org/10.21595/vp.2017.19073.
38. Hsieh CF, Jian WS. The effect on dynamics of using various transmission designs for two-stage cycloidal speed reducers. Proc Inst Mech Eng C J Mech Eng Sci 2016;230:665–81. https://doi.org/10.1177/0954406215618984.
39. Xuan L, Xie C, Guan T, Lei L, Jiang H. Research on dynamic modeling and simulation verification of a new type of FT pin-cycloid transmission. Proc Inst Mech Eng C J Mech Eng Sci 2019;233:6276–88. https://doi.org/10.1177/0954406219861999.
40. Yang R, An Z. Theoretical calculation and experimental verification of the elastic angle of a cycloid ball planetary transmission based on the axial pretightening force. Advances in Mechanical Engineering 2017;9:1–17. https://doi.org/10.1177/1687814017734112.
41. Król R. Software for the cycloidal gearbox multibody dynamics analysis, implemented in Fortran. (Purpose: presentation of the results in the scientific article) 2022. https://doi.org/10.5281/ZENODO.7221146.
42. Nikravesh PE. Planar Multibody Dynamics. 2018. https://doi.org/10.1201/b22302.
43. Nikravesh PE. Planar multibody dynamics: Formulation, programming and applications. 2007.
44. Flores P, Lankarani HM. Contact Force Models for Multibody Dynamics. vol. 226. Cham: Springer International Publishing; 2016. https://doi.org/10.1007/978-3-319-30897-5.
45. MSC Software. MSC Adams Solver Documentation n.d.
46. MSC Software. MSC Adams View Documentation. n.d.

This research utilised the MSC Software National Scientific Software License, operated by the TASK Computer Centre in Gdańsk (Poland). This license was funded by a computational grant obtained by Kazimierz Pulaski University of Technology and Humanities in Radom, Poland.

Roman Król: https://orcid.org/0000-0002-6279-9562

Kazimierz Król: https://orcid.org/0000-0002-8131-468X

DOI 10.2478/ama-2023-0032

FORECASTING BIOGAS FORMATION IN LANDFILLS

Zbigniew KNEBA*, Jacek KROPIWNICKI*, Jakub HADRZYŃSKI**, Maciej ZIÓŁKOWSKI***

*Gdansk University of Technology, ul. Gabriela Narutowicza 11/12, 80-233 Gdańsk, Poland
**Eko Dolina, Aleja Parku Krajobrazowego 99, 84-207 Łężyce, Poland
***KO-Energia, ul. Klonowa 52 lok. 17, 83-330 Lniska, Poland

zkneba@pg.edu.pl jkropiwn@pg.edu.pl dt.ziolkowski@gmail.com

received 6 September 2022, revised 5 January 2023, accepted 23 January 2023

Abstract: The aim of the present research was to develop a mathematical model for estimating the amount of viscous gas generated as a function of weather conditions. Due to the lack of models for predicting gas formation caused by sudden changes in weather conditions in the literature, such a model was developed in this study using the parameters of landfills recorded for over a year. The effect of temperature on landfill gas production has proved to be of particular interest. We constructed an algorithm for calculating the amount of the produced gas. The model developed in this study could improve the power control of the landfill power plant.

Key words: landfill gas, waste usage, atmospheric conditions on landfills

1. INTRODUCTION

The landfill is primarily a source of two greenhouse gases: carbon dioxide and methane. In the face of threats posed by greenhouse gas emissions to the atmosphere, special attention needs to be paid to methane gas. It's Global Warming Potential (GWP) = 28–36 (100 is the maximum value) (1). Since non-collected methane appears a major greenhouse gas source (regarding its impact factor), all new landfills must have methane facilities.

In the European Union, the binding regulation in the field of industrial emissions (integrated pollution prevention and control) is DIRECTIVE 2010/75/EU of the European Parliament and of the Council of Europe 24 November 2010 (2). In the directive, the emission limits for burning landfill gas are set on the levels presented in Tab. 1.

Tab. 1. Emission limits for burning landfill gas in combustion engines

Pollutant	Limit mg/Nm3/day
NOx	350
SO2	200
Ash (solid particles of any kind)	50

The reduction of CH4 emission to the atmosphere in the years 1990–2010 reached a level of 30%, In next ten years 2010 – 2020 is was reduced of next 10 % (3) .

The phenomenon of methane formation in landfills has been analysed in numerous studies (4)(5) (6)(7). However, there are only few considerations on the impact of atmospheric conditions and the possibility of short-term prediction of changes in the amount of methane produced on the basis of the measurements of meteorological parameters (8).

Predicting and controlling the process of methane collec-tion and use are the objectives of this study. The future perspective of this study would be to change the control procedure of the combustion engines used in landfills.

2. LANDFILL GAS PRODUCTION TECHNOLOGY

Fig. 1 shows a schematic of using landfill gas for electricity and heat production.

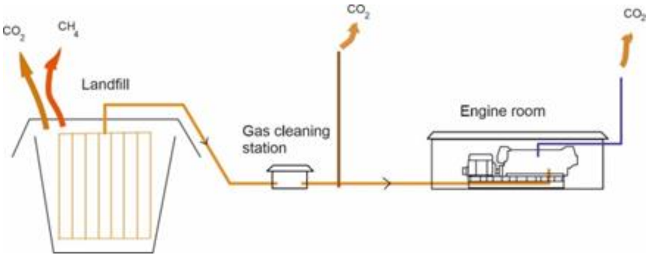


Fig. 1. Greenhouse gas emissions from landfill

There are several studies ([5,9,11,12,19] on landfill gas LFG production in the world. These studies mostly differ in waste types and ambient conditions, but the gas collecting technology is the same in most cases (Fig. 1).

The collection system is a combination of horizontal and vertical collectors. Horizontal collectors are used for removing condensates and rain water, but they are also used in collecting LFG. The system has three phases: I: active cells, where new wastes are delivered; II: closed cells with temporary covers; and III: closed cells with final covers. Collection efficiencies for phases I, II and III are mostly set at 50%, 75 % and 95 % of its potential [4], respectively.

In this study, the authors used data recorded in the course of

DOI 10.2478/ama-2023-0032

operation of two sections (phases I and II) of a working ECODOLINA landfill serving the city of Gdynia in Poland. The service life of LFG collecting facilities can reach up to 80 years. Initially, the amount of produced methane is rising (for about 36 years), then it slowly and exponentially drops. In the first stage, when methane generation increases, the landfill is not able to create anaerobic conditions. In optimal conditions, only 40% of waste can be used for LFG production.

To control odour and to meet emission regulations, landfill operators collect and burn LFG. LFG collection efficiency varies mainly due to the LFG collector operation, types of cover and time of the installation of covers. Landfill consists of many cells and is often managed manually by the staff on a cell-by-cell basis. This operation requires more-experienced staff.

The amount of degradable organic carbon factor (DOCF) is proportional to the mass of wet waste DOCF (Tab. 1) and equals 0.5 for landfills that actively produce methane (7). This factor is very uncertain and depends on the type of waste. However, it can be used to assume LFG mass from the mass of waste.

Tab. 2. Material degradable organic factors [5]

Material	DOCF
Paper	0.19 ÷ 0.54
Wood	0.02 ÷ 0.57
Food waste	0.36 ÷ 0.92
Trimming grass	0.09 ÷ 0.38

DOCF, degradable organic carbon factor.
The volume of methane in a landfill gas is 54% ÷ 73%. In a landfill in the USA, waste consisted of paper 14.3%, wood 8.1%, food 21.6% and yard trimming 7.9% (7).

Producing electricity from LFG can be regarded as distributing regional electricity, which indirectly avoids CO2 emissions. Electricity generated in landfills leads to a displacement credit of 550 g CO2/kWh (5). The oxidation factor in landfill cover is amounting to 36% (9).

There are models in the literature describing the amount of methane produced at a landfill.

The model assumes that the entire amount of organic carbon that undergoes decomposition forms CH4 and CO2 (10). The formula to calculate the final amount of the produced decomposed methane from organic carbon is given in Eq. (1):

$$V_{CH_4}{}^{Year} = M_{OV}{}^{Y-1} \cdot (1 - e^{-k}) \cdot sh \cdot \frac{16}{12}, \qquad (1)$$

where: $V_{CH_4}{}^{Year}$ is the amount of methane generated in year Y, M_{OV} is the mass of organic carbon containing waste, *k* is the reaction rate constant in the year, *sh* is the share of methane in landfill gas and $\frac{16}{12}$ is the mass ratio of CH_4 to C.

The Gaussian model is another model that can used for the simulation of LFG production when municipal solid waste (MSW) is pre-treated (11). The model assumes that the LFG production rate follows the normal distribution. The model is described by Formula (2):

$$V_{LFG}{}^{Day}(t) = a \cdot e^{-0.5 \cdot \frac{t-t_0}{b}}, \qquad (2)$$

where $V_{LFG}{}^{Day}(t)$ is the LFG production rate in m^3/(ton day), t is the time of digestion, *a* is the ultimate LFG production rate in m^3/(ton day), *b* is constant in day and t_0 is the time in day where the maximum LFG production rate occurred.

Next model was created as a result of research using an experimental reactor with a washed sample of landfill material by spreading the tested sample for a year (12).

More detailed models divide the deposited waste into fast, medium and slowly biodegradable waste. Their use is more difficult because most landfills does not register the types of waste. The equation that describes such method of calculation is given as follows (3):

$$V_{LFG}{}^{Day}(t) = \sum_{i=1}^{n} \sum_{j=0}^{m-1} A_{j+1}(t_i - t_j) e^{[-k_{j+1}(t_i - t_j)]}, \qquad (3)$$

where: $V_{LFG}{}^{Day}(t)$ is the LFG production at time [m^3/(t(MSW)*day)], *A* is the amplitude of LFG production at the day [m^3/(t(MSW)*day)], *k* is the reaction rate constant in the year, *n* is the total number of days, *m* is the number of biodegradable components of heterogeneous pre-treated MSW and t_j is the delay time, which is defined as a period between the beginning and the end of biodegradable components.

Based on the literature review of the models of landfill gas formation, it can be concluded that they are functions of time and waste composition. These models are built to predict slow changes in operation over the years. However, there are no models for forecasting changes caused by temporary weather conditions with a forecast for several days

3. EXPERIMENTAL SETUP

The subject of the research was the ECODOLINA landfill in Lezyce, near Gdynia, Poland. It is a MSW landfill. Ambient conditions influence the formation of favourable conditions for fermentation. Recorded weather conditions included pressure, temperature, rainfall in 10-min increments and landfill gas production: gas flow, temperature, pressure, methane concentration and oxygen concentration. As LFG is obtained from anaerobic digestion, high humidity and temperature, alkaline environment (pH = 7.5) with maximally limited oxygen access is preferred for the process.

Depending on the temperature of the anaerobic fermentation process, it is possible to distinguish the following:

- mesophilic fermentation, which takes place at a temperature of about 30–40°C and
- thermophilic fermentation, which occurs at a temperature of about 52–58°C.

Mesophilic fermentation takes place in the deposits utilized by EKODOLINA due to waste accumulation in the open air. There are three sites at the landfill: 1979–2002, 1990–2002 and 2002 – currently. The LFG is obtained from 105 wells – 75 from old sites and 40 from new sites. The wells and drain collectors are connected to two main collectors that carry gas to the power plant. The concentration of methane obtained from these quarters differs significantly: the oldest site produced only 10% of the concentration of methane in LFG, newer site provide about 50%, and current site produced 60% of concentration. The composition of the landfill gas is checked periodically mainly to prevent failures of gas powered engines. If hazardous chemicals such as hydrogen sulphide or siloxanes are detected, the gas is proceeded for treatment. Carbon dioxide and water also require treatment.

Parameters of the biogas composition were analysed by the "“Omnisfera”“. company Chemical analysis of the biogas sample of the EKODOLINA landfill is presented in Tab. 3.

Tab. 3. LFG parameters in EKODOLINA (13)

Parameter	Unit	Value	Error of measurement
Temperature	°C	21.5	0.1
Relative humidity	%	62.7	-
Density	kg/m3	0.83	-
Methane	%Vol	43.3	4.3
Carbon dioxide	%Vol	31.1	3.1
Oxygen	%Vol	2.0	0.3
Hydrogen	%Vol	0.0034	-
Nitrogen		8.92	-
Carbon monoxide	ppm	19	3
Hydrogen sulphide	%	0.1608	0.0053
Chlorine	mg/m3	94.83	28.45
Oil moisture	mg/m3	0.1	
Siloxanes	mg/m3	0.859	0.301
Silicon	mg/m3	0.314	0.110
Sulphur	mg/m3	2605	859
Ash PM10	mg/m3	0.03	-

As can be seen from the sample, the landfill gas is highly sulphated. Sulphur has a largely negative impact on all components of the technological line as well as on cooperating devices. The engines fail frequently because of sulphur as it reacts with all metal parts and the engine oil, significantly reducing the lubricating property of the oil. This, in turn, damages other components and causes a domino effect. Exhaust valves get damaged first. Due to the sulphur deposits on the valve plug and stem, the geometry of the valve is changed (17). Short-term prediction of changes in the gas composition as a function of raw material supply for its production and weather conditions would enable better control of engine operation.

The process of obtaining biogas is very simple. Through uniformly spaced gas wells that have an impact radius of 15 m, gas is collected into common collectors. Fig. 2 gives an example of the construction of a degassing well has been presented.

Through Φ 12-mm pores and drainage backfilling from 16 -to 32-mm gravel, biogas gets collected into a Φ160-mm pipe. The drainage pipe collects gas throughout the depth of the landfill. By doing so, it has only been possible to build a well comprising a 15-m-deep layer of waste. The landfill itself was closed with an average waste depth of 25 m. To make the wells most effective, it was decided to raise wells This construction proved to be unfavourable due to the deposit's topography. The site should be treated as morphologically unstable. This indicates that displacements and settlements occur inside, which irreversibly affects drainage; causes displacement, crushing and splitting at particular heights of collecting pipes; and limits the flow from the gas collection pipe to the drains towards the collecting stations. Leads are called connection collectors, and they run towards collecting stations.

In addition to vertical wells, it should be noted that horizontal drainage can also be carried out in a similar way. The horizontal wells exactly play the same role and help limit the release of biogas into the atmosphere. The system involving the two types of wells is more effective, thus doubling its functioning, as well as its efficiency and flow increase. This has a measurable effect in reducing the negative pressure that exists in the entire network. There are 150 wells in the plant, including 75 in the old site of the landfill and 40 in the new site. These are all connected to two main collectors that transport biogas to engines.

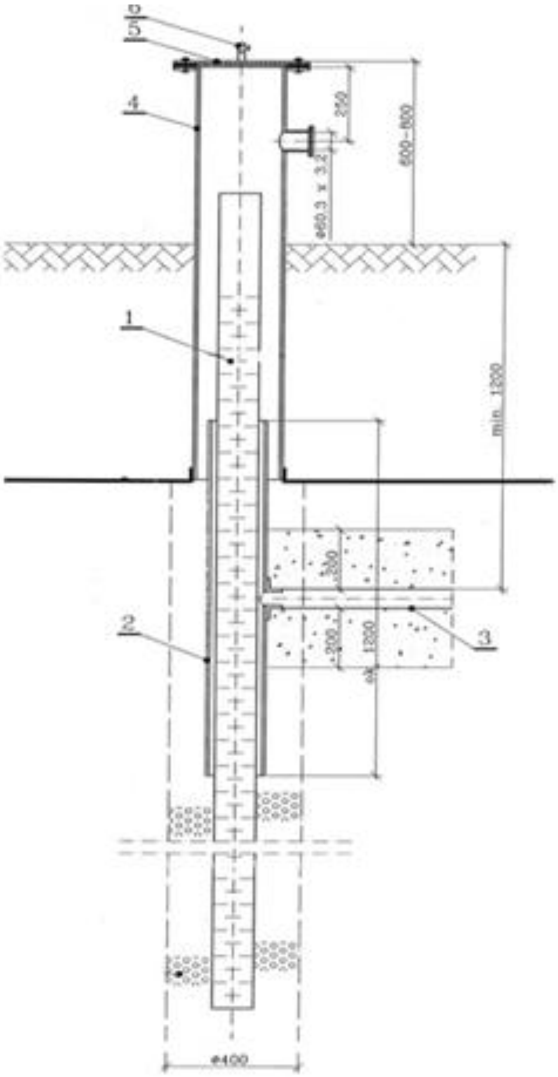


Fig 2. LFG well 1 - drainage pipe, 2 - casing pipe, 3 – water drainage pipe, 4 - degassing chimney 5 – cover, 6 – connection fitting (13)

The biogas intake network system includes the following:

- collective stations for individual water bodies with A (rich)/B (poor) gas selection,
- deep water intakes defined as gas wells,
- well connections to collecting stations made of PE Φ 63 SDR 110,
- transmission collectors for a technological building – a two-wire parallel system,
- drainage systems for gravity and pump networks and
- automatic biogas suction and pressure station.

From the storage sites, the collected biogas is first transported through collectors to the collective station, the construction of which is shown in Fig. 3.

The biogas is delivered to the compressor through two main pipelines. Drainage wells are installed along the entire length of the pipelines where the levels are broken. Due to the high moisture content of biogas and the differ-ences in temperature inside the deposit in relation to the temperature prevailing in the collecting stations and the pipelines, the condensate precipitates. To prevent clog-ging, pipelines should be discharged at the lowest points. An example of such a well is shown in Fig. 4.

DOI 10.2478/ama-2023-0032

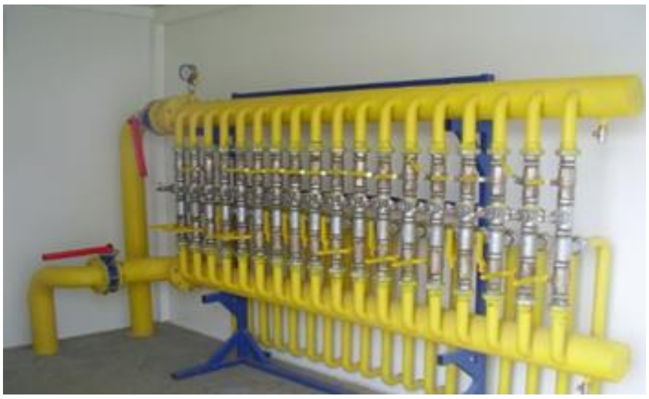


Fig. 3. Connection of gas wells to main pipelines (13)

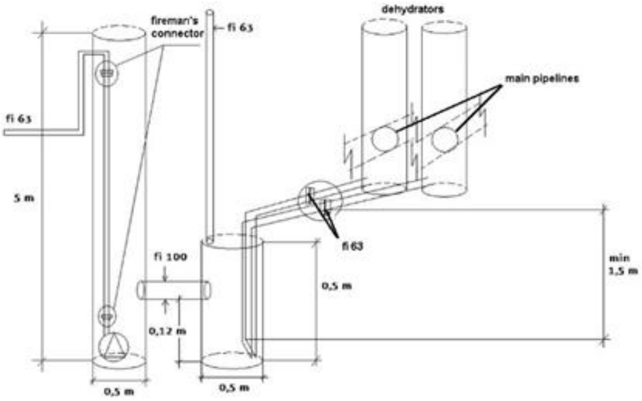


Fig. 4. Construction of a drainage well for the main gas pipelines (13)

The landfill gas flow rates and its chemical composition as well as weather conditions were registered at the ECODOLINA landfill. Existing equipment installed for monitoring the operation of the gas utilization system was used. The registered data were used to model the landfill gas generation in this study.

4. RESULTS OF THE MEASUREMENTS

The amount of landfill gas generated in the system is the source of electrical energy and thermal energy. The LFG and total energy production are presented in Fig. 5. Fig. 5 presents the streams of sucked gas registered in January 2018–December 2018.

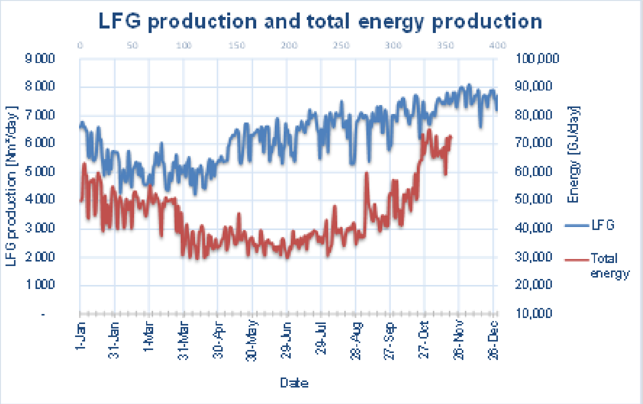


Fig. 5. LFG and total energy production in 2018

The total amount of energy produced in cogeneration increases during the heating season due to the use of thermal energy. Changes in the amount of energy produced over time, with greater amplitudes than changes in gas production in a landfill, result from fluctuations in the calorific value of landfill gas. These fluctuations are shown in Fig. 6. The largest changes were observes in August and September, which contributed to the enormous changes in energy production.

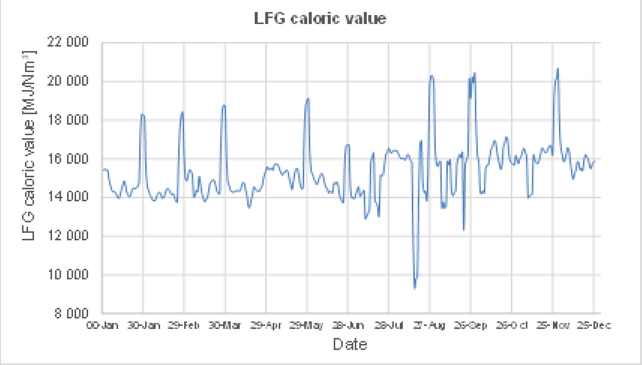


Fig. 6. LFG caloric value over the time in 1 year

The aim of this study was to discover the relationship between atmospheric conditions and LFG production. After changing the ambient temperature, the heating or cooling of the sites at the landfill takes place slowly. This is typical of the old, enclosed quarters. They have a layer of soil insulating from the environment, and before the heat penetrates this layer, gas production does not change. While working on the data recorded at the ECODOLINA landfill, the time shift when considering the effect of temperature is 1 week. To illustrate the process of gas production based on temperature, the charts shown in Fig. 7 were prepared. The graphs presented in the charts take into account the 1-week delay of the change in the production of landfill gas in relation to the moment of the temperature change.

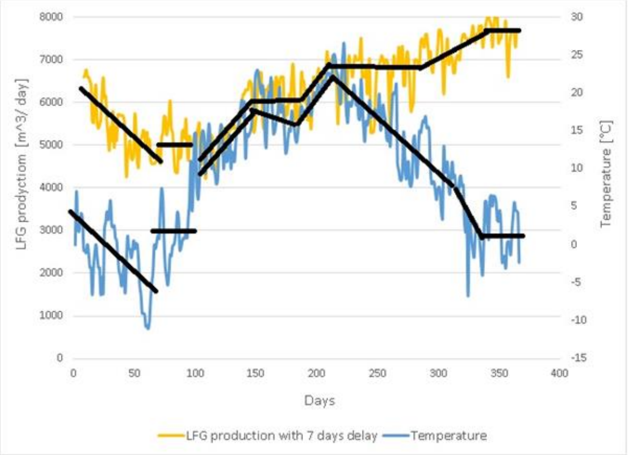


Fig. 7. LFG production (with 7 days delay) – yellow line and ambient temperature on the landfill area – blue line. 0 on the abscissa represents the first day of 2018

The analysis of the graphs plotted for the whole year allows drawing the following conclusions. The calendar year is divided into periods:

- the landfill sites heat up from the surroundings;
- they are in thermal equilibrium with the surroundings; and
- they are cooled by the surroundings.

After reaching sufficient temperatures for the development of methanogenic bacteria (in our measurement, 24°C ambient air), an increase in ambient temperature does not increase the production of landfill gas.

During summer, and more specifically from April 1 to July 30 (from 90th to 212th day of the year), landfill gas production is correlated with ambient temperature. After summer, the microbial population is so high that gas production in the sites is constant or even increases significantly regardless of the ambient temperature decrease at that time. Constant production is maintained by saturating the bed with bacteria, and in the absence of new food, their number does not increase. However, the decrease in temperature worsens living conditions of the old and the new bacteria that have multiplied and began their work. During winter, until a critical point at the turn of the year, production is stable due to the sufficient microbial population maintained in the bed.

We noted the following relationships between the amount of gas produced and the ambient temperature:

- Analysis of the temperature changes 1 week in advance to calculate the actual gas production
- An increase in ambient temperature results in an average directly proportional increase in gas production until the ambient temperature reaches 24°C.
- After 24°C is reached, methane production becomes constant.
- With a slow temperature drop, the production remains constant due to the heat accumulated in the landfill and a large microbial population. This phenomenon can last for even 50 days.
- Only a temperature drop to below 5°C causes a directly proportional decrease in methane production.

The gas production dynamics can be described using time to predict future changes in production, relative to the change in temperature, which is mostly rising during the summer season. This 100-day seasonal change is shown in the graph in Fig. 8.

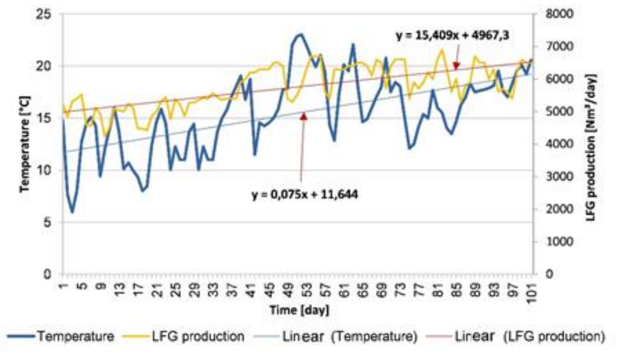


Fig. 8. Illustration of the dynamics of the process during summer season (100 days)

The derivative of the temperature as a function of time dT/dt is 0.07 [°C/day], and the derivative of the daily gas production dQ/dt is 15.4 [Nm^3/day].

We propose a method to predict the amount of gas produced in a current week based on the average ambient temperature in the period of time including previous 10 days. Since the relationships between the velocities turned out to be logarithmic, it is possible to formulate an algorithm for planning the amount of gas produced in summer, knowing the temperatures in the past. Different periods were tested for LFG production prediction, but the 10-day period seems to be the most suitable time. Various tested periods for the process are shown in Fig. 9.

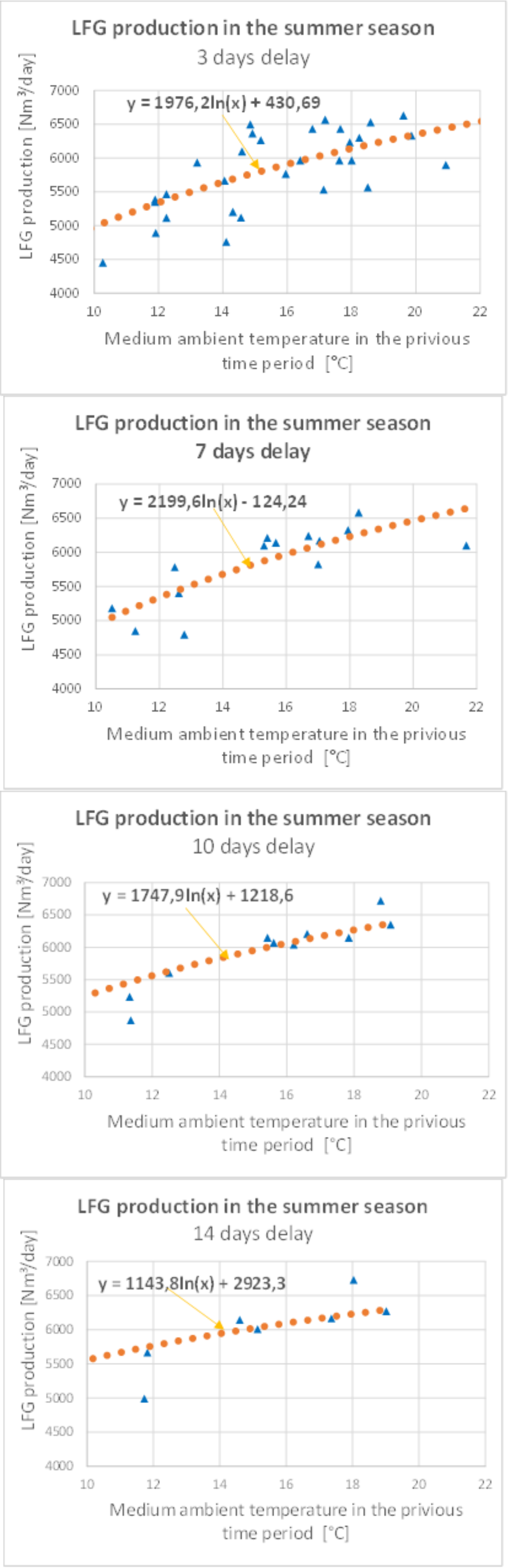


Fig. 9. LFG production model verification process for several days of the observation process

The algorithm for calculating the future gas production in the summer season is shown in Fig. 11. Fig. 10 shows-Correlation coefficient between model formula and experimental data.

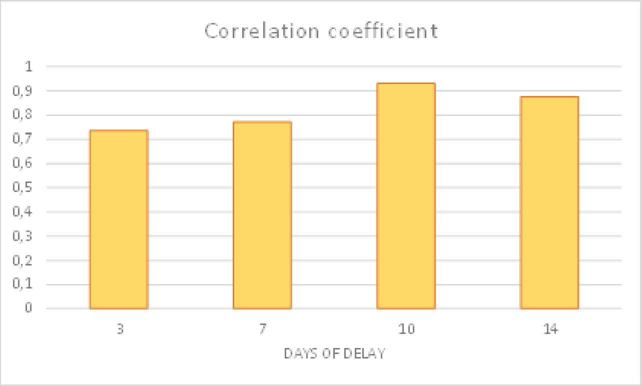


Fig. 10. Correlation coefficient between model formula and experimental data

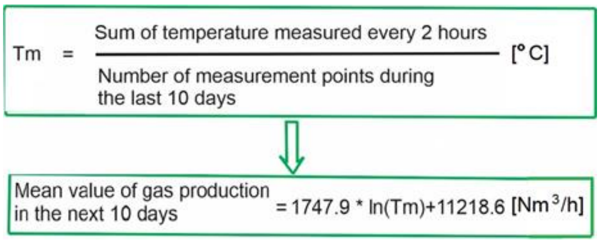

Tm =
Sum of temperature measured every 2 hours
Number of measurement points during the last 10 days
[° C]
Mean value of gas production in the next 10 days
= 1747.9 * ln(Tm)+11218.6 [Nm³/h]

Fig. 11. Algorithm for calculating the future gas production in the summer season

The aforementioned remarks should be used in planning the process of using landfill gas for energy production.

A previous study (14) showed that changes in the methane content in biogas follow atmospheric pressure changes.
Changes in the methane content depending on atmospheric pressure are analysed. After changing the atmospheric pressure upwards, the methane content in the landfill gas begins to increase with a 2-day delay. This situation is related to the volatility of the methane–carbon dioxide mixture. Methane is less dense than air and carbon dioxide and is more dense than air.

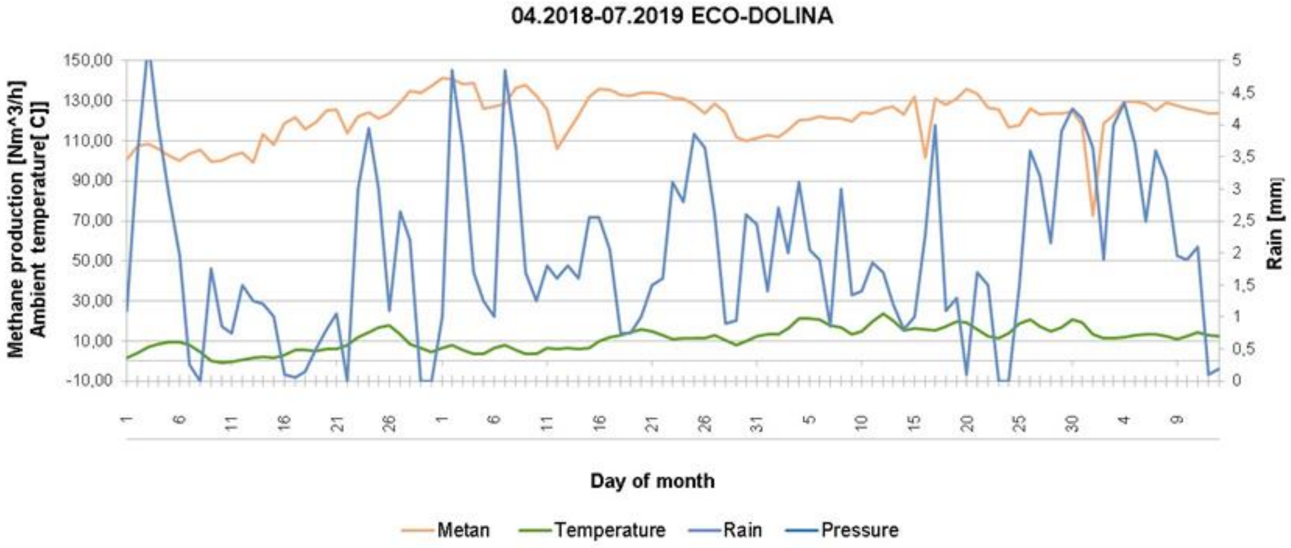


Fig. 12. Production of methane in relation to the intensity of rainfall and ambient temperature

Probably, during periods of higher pressure, the atmospheric part of carbon dioxide is not pulled out of the well (prisms) by the system suction of biogas, which results in obtaining biogas with a higher methane content. If atmospheric pressure decreases, carbon dioxide quantity increases, causing the reduced methane content. No statistically significant correlation was found in the effects of barometric pressure on LFG quality in wells. (15) (16) (16).

The current trend leads towards landfill bioreactor technology (LBT) systems, which augment the amount of water contacting the waste to stabilize it rapidly. This technique can produce large initial LFG generation rates, which will decrease sharply after a few years.

Rainfall is conducive to the production of landfill gas; a previous study (17) found that the methane generation coefficient k increases with a wet bed from 0.02 to 0.065. The authors put forward a similar conclusion, observing the increase in methane production in the diagram shown in Fig. 12. The winter months of 2018 were dry, while the spring months of 2019 were rainy, which affected the amount of methane produced.

The effect of rainfall is significantly lower during the summer season than during winter.

5. SUMMARY

There are difficulties in taking into account many factors affecting landfill gas production. Weather phenomena can change quickly. For example, in Poland, temperature drops by 28° have been recently recorded. Rapid tem-perature changes do not affect the amount of LFG pro-duced immediately. The effect of temperature changes is only visible after approximately 10 days. After a long period of summer temperatures, the amount of live methane bacteria is sufficient and gas production does not increase with increasing temperature. Due to the high heating effect of the landfill site in the summer, the production of the LFG maintains a high level in winter.

After testing various periods of several days of taking into account the effect of temperature on gas production, a 10-day observation period was selected.

Since the relationships between the velocities turned out to be linear, it is possible to formulate an algorithm for estimating the amount of gas produced in summer, based on the derivative of the temperature change in the last 10 days.

The impact of atmospheric pressure was more evident in gas production, after 2 days or 3 days.

There was no clear correlation between the amount of rainfall and the production of landfill gas.

This work contributes to building short-term mathematical models of landfill gas production. The developed model could be used to control the operation of a cogeneration heat and power CHP plant, such as preparations for switching on further power generators to increase the efficiency of energy production.

REFERENCES

1. EPA. Municipal Solid Waste Landfills What is a Municipal Solid Waste Regulations for Municipal Solid Waste Landfills Publications and Guidance for Municipal Solid Waste [Internet]. 2023. Available from: http://epa.gov/landfills
2. European Council. Directive 2010/75/EU Industrial Emissions. Off J Eur Union [Internet]. 2010;L334:17–119. Available from: http://europa.eu/legislation_summaries/environment/air_pollution/ev0027_en.htm
3. Van Dingenen R, Crippa, M. J, Anssens-Maenhout G, Guizzardi D, Dentener F. Global trends of methane emissions and their impacts on ozone concentrations. Vol. EUR29394EN, JRC Science for Policy Report. 2018.
4. Sarptaş H, EKER S, Seyfioglu R, Boyacioglu H, Dolgen D, Alpaslan N. Models for the Prediction of Landfill Gas Potential – A Comparison. 2012.
5. Benato A, Macor A. Italian biogas plants: Trend, subsidies, cost, biogas composition and engine emissions. Energies. 2019;12(6): 1–31.
6. Vakalis S, Moustakas K. Applications of the 3T method and the R1 formula as efficiency assessment tools for comparing waste-to-energy and landfilling. Energies. 2019;12(6):1–11.
7. Barlaz MA, Chanton JP, Green RB. Controls on landfill gas collection efficiency: Instantaneous and lifetime performance. J Air Waste Manag Assoc. 2009;59(12):1399–404.
8. Meres M, Szczepaniec-Cieciak E, Sadowska A, Piejko K, Szafnicki K. Operational and meteorological influence on the utilized biogas composition at the Barycz landfill site in Cracow, Poland. Waste Manag Res. 2004;22(3):195–201.
9. Manyuchi MM, Mbohwa C, Mpeta M, Muzenda E. Methane generation from landfill waste as a resource recovery strategy. Proc Int Conf Ind Eng Oper Manag. 2019;2019(MAR):200–24.
10. Negm AM, Shareef N. Waste Management in MENA Regions. Springer Water Ser. 2019;(June).
11. Mahar RB, Sahito AR, Yue D, Khan K. Modeling and simulation of landfill gas production from pretreated MSW landfill simulator. Front Environ Sci Eng. 2016;10(1):159–67.
12. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Vol 5 Chapter 3 Solid Waste Disposal. 2006 IPCC Guidel Natl Greenh Gas Invent [Internet]. 2006;4:6.1-6.49. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20604432.
13. Handrzyński J. Design of modification of control system of engines powered by biogas fuel, diploma thesis. Polish Naval Academy; 2015.
14. Zi M, Kropiwnicki J. Identification analysis of dynamic changes in the composition of biogas and their impact on the operation of internal combustion engines. 2011;1–6.
15. LMOP. Landfill Gas Energy Basics Landfill Gas Modeling Project Technology Options Project Economics and Financing Landfill Gas Contracts and Regulations Best Practices for Landfill Gas Collection System Design and Installation Best Practices for Landfill Gas Co. 2021.
16. Czepiel PM, Shorter HR, Mosher B, Allwine E, McManus JB, Harriss R, et al. The influence of atmospheric pressure on landfill methane emissions. Waste Manag.; 2003.
17. Leone J. The Effects of Atmospheric Pressure Changes on Landfill Gas Collection Efficiency and Quality. 2007.

Zbigniew Kneba: https://orcid.org/0000-0002-6956-2330

Jacek Kropiwnicki: https://orcid.org/0000-0001-7412-7424

NUMERICAL INVESTIGATION OF PRODUCTION-RELATED CHARACTERISTICS REGARDING THEIR INFLUENCE ON THE FATIGUE STRENGTH OF ADDITIVELY MANUFACTURED COMPONENTS

Michaela ZEIßIG*, Frank JABLONSKI**

*Faculty of Production Engineering, Bremen Institute for Mechanical Engineering (bime), University of Bremen,
Am Biologischen Garten 2, 28359 Bremen, Germany
**Faculty 5, City University of Applied Sciences Bremen, Neustadtswall 30, 28199 Bremen, Germany

mzeissig@uni-bremen.de, Frank.Jablonski@hs-bremen.de

received 10 October 2022, revised 1 January 2023, accepted 5 January 2023

Abstract: In order to further enhance the application of additive manufacturing (AM) processes, such as the laser powder bed fusion (L-PBF) process, reliable material data are required. However, the resulting specimen properties are significantly influenced by the process parameters and may also vary depending on the material used. Therefore, the prediction of the final properties is difficult. In the following, the effect of residual stresses on the fatigue strength of 316L steel, a commonly used steel in AM, is investigated using a Weibull distribution. The underlying residual stress distributions as a result of the building process are approximated for two building directions using finite element (FE) models. These imply significantly different distributions of tensile and compressive residual stresses within the component. Apart from the residual stresses, the impact of the mean stress sensitivity is discussed as this also influences the predicted fatigue strength values.

Key words: additive manufacturing, fatigue strength, Weibull distribution, FEM

1. INTRODUCTION

Additive manufacturing (AM) shows great potential across various industries ranging from aerospace to medical applications. Some of its main advantages are moldless manufacturing, production of near net-shape geometries and great geometrical freedom. There are different processes which belong to the group of AM techniques. One of them is the widespread laser powder bed fusion (L-PBF) process. The component is built using repetitive cycles of powder layer distribution and selective laser melting of the powder. This happens according to a previously defined and virtually sliced geometry [1]. The process itself has numerous configurable process parameters. Among them are the building direction of the component relative to the building platform, the laser parameters like speed, power and scanning path and powder-related parameters such as the particle size, temperature and the powder layer height [1]. This multitude of parameters makes predictions of the final properties of the component difficult, especially as general assumptions for different AM materials are not possible. This means that each material needs to be investigated separately considering the process and material-related characteristics. In the following, the focus is on the fatigue strength of the austenitic steel 316L (1.4404).

2. CHARACTERISTICS OF ADDITIVELY MANUFACTURED COMPONENTS WITH RESPECT TO FATIGUE STRENGTH

Due to the manufacturing process, L-PBF components show some characteristics which are in line with the known influencing parameters on the fatigue strength. These include tensile residual stresses, high surface roughness, a specific grain structure and stress raisers such as pores and other defects [2]. These in combination with external parameters such as the loading direction and potential mean stress determine the fatigue life of a specific component [3]. Each of the characteristics has an effect on the fatigue strength; however, their influences may not be easily assessed independently because of their simultaneous occurrence in the components. In the following, the AM characteristics and their general effects on the fatigue strength will be briefly described.

2.1 Residual stresses

The repetitive process of melting and solidification in line with heating and cooling of the component leads to the development of significant tensile residual stresses, at least in the outer region of as-built components [4,5]. These have a negative impact on the fatigue strength. Postprocessing, e.g., in terms of heat treatments or hot isostatic pressing may reduce these residual stresses (for some materials) as reported in the study by Leuders et al. [6]. Knowledge about the residual stress distribution within components is required to decide whether these postprocessing procedures are necessary to achieve the required properties. Knowledge is also required if these treatments shall be avoided by minimisation of residual stresses and distortion via optimised building parameters (e.g. [7]).

2.2 Pores and other defects

The porosity of L-PBF components is usually very low. In Zhang et al. [8], the density of such components is given as around

95%, while in the study by Hatami et al. [5], the porosity determined via image analysis was even well below 0.5%. However, there are still some pores and other defects present in the component. These result from the manufacturing process and may be divided by their origin, which in turn is related to their morphologies. In the study by Zhang et al. [9], three different defect types are distinguished: pores, lack-of-fusion (LOF) defects and cracks. While pores are usually small (up to about 100 $\mu\mathrm{m}$) and rather spherically shaped, LOF defects are rather irregularly shaped [10,11]. The pores are attributed to gas bubbles while the LOF defects are attributed to insufficient melting of the particles [11]. The defects correspond to stress concentrations and, therefore, may be potential crack initiation sites.

2.3 Surface state

Unprocessed, so-called as-built surfaces show high surface roughness. For 316L, surface roughness values are given in the study by Wang et al. [12] for Ra as between 5 and 25 µm, while the value depends on the manufacturing parameters. Furthermore, the surface roughness varies with the orientation of the surface relative to the building direction. In addition to the mere surface roughness, near-surface porosity, meaning small pores in a zone below the surface as reported in the study by Hatami et al. [5], may need to be included in the assessment of the surface state. In general, high surface roughness has a detrimental effect on the fatigue strength as it may, e.g., induce stress peaks.

2.4 Microstructure

Depending on the building parameters and the material, different types of microstructures may be found in the final component. For 316L, a cellular microstructure is typical while for other materials like Ti-6-Al-4V elongated, columnar grains are reported [13].

3. FATIGUE STRENGTH EVALUATION OF L-PBF COMPONENTS

As stated above, various factors influence the fatigue strength of a component built via the L-PBF process. In the following, a Weibull distribution [14,15] which is often applied for fatigue assessment will be used. Here, it is applied in such a way that it explicitly incorporates the effect of residual stresses and inherently the effect of defects present in the component.

3.1 Weibull distribution

The Weibull distribution is a probability density function based on the assumption that every component initially includes randomly distributed defects. Furthermore, fatigue failure is defined as crack initiation. This in turn is supposed to take place if the local stress exceeds the local material endurance.

For an individual volume element ΔV of a component with the total volume V_0, the survival probability is calculated according to the following equation:

$$P_{\mathrm{survival},\Delta\mathrm{V}} = 2^{-(\Delta V/V_0)(\sigma_{\mathrm{Mises}}/\sigma_{\mathrm{WV}})^{m_{\mathrm{V}}}} \tag{1}$$

with σ_{Mises} being the equivalent stress amplitude and σ_{WV} representing the local fatigue limit. The Weibull exponent m_{V} is used to fit the effect of size and scattering of the defects present in the component [16].

The residual stresses $\sigma_{\mathrm{i}}^{\mathrm{RS}}$ are accounted for as part of the material resistance [17,18], not as part of the material loading. Consequently, part of the exponent of Eq. (1) can be rewritten as follows:

$$\frac{\sigma_{\mathrm{Mises}}}{\sigma_{\mathrm{WV}}} = \frac{\sigma_{\mathrm{Mises}}}{\sigma_{\mathrm{WV}} - M\cdot(\sigma_1^{\mathrm{RS}} + \sigma_2^{\mathrm{RS}} + \sigma_3^{\mathrm{RS}})} \tag{2}$$

The impact of the residual stresses is influenced by the residual stress sensitivity M. According to [3], it may be assumed equal to the mean stress sensitivity following the assumption that the impact of residual stresses is comparable to that of mean stresses. For non-welded components, M can be determined based on the ultimate tensile strength R_m as follows:

$$M = a_M \cdot 10^{-3} \cdot \frac{R_m}{\mathrm{MPa}} + b_M \tag{3}$$

with a_M and b_M being material constants. For steel, the constants take on the values 0.35 and −0.1, respectively [19]. Using the values from Tab. 1, M is calculated to be about 0.14 for the horizontal and 0.11 for the vertical building direction. In addition, a common value used for M is 0.3. It corresponds to the value given in the FKM guideline [19] for the state of low residual stresses within a welded component. However, as the values measured in the study by Hatami et al. [5] and the values obtained via the approximations using finite element (FE) models below both suggest residual stresses above the threshold of $0.2 \cdot R_p$ (see Tab. 1), the assumption of low residual stresses seems not to be valid. However, in order to assess the impact of the residual stress sensitivity on the fatigue predictions using Weibull distributions, it will be included in the calculations below.

Moreover, the application of Eq. (2) requires the local fatigue limit σ_{WV}. This value refers to the fatigue limit without influencing factors such as residual stresses. Therefore, the value is approximated using the relation between the ultimate tensile strength and the fatigue strength. A factor of 0.4 is assumed [19]. Furthermore, the Weibull exponent m_{V} is set to 20. The required input data regarding the residual stress distribution is obtained via FE calculations as described in the following.

3.2 Approximation of residual stress distribution

The unnotched fatigue specimen used has a length of about 50 mm and a smallest diameter of 4 mm. The material used is 316L steel, an austenitic steel commonly used in AM due to its wide applicability and its comparably easy handling. The main material parameters used for the Weibull approach are given in Tab. 1. As one can see, the material data have a dependency from the building direction of the component. This applies to the static values and may also be valid for the fatigue strength.

The residual stress data were obtained via a rough approximation using the Abaqus Welding Interface (AWI) [21], an extension to Abaqus [22]. For the calculations, Poisson’s ratio was assumed as 0.3, and material data implemented in the AWI [21], similar to that of 316L steel, were supplemented. Sequentially coupled calculations have been carried out, neglecting the scanning strategy, combining several component layers of typical height around 50 $\mu\mathrm{m}$ [5] and adding them as a whole. The thermal history is obtained via the first calculation and subsequently used as input for the second calculation in order to obtain the stress distribution. Afterwards, the component is removed from the building platform, and the axial stress distributions along a path through the middle of the specimen as shown in Fig. 1 are obtained.

Tab. 1. Material data for 316L steel with respect to building direction taken, adapted from Gläßner et al. [20].

Characteristics	Horizontal manufacturing direction	Vertical manufacturing direction
Young's modulus	167 GPa	152 GPa
R_m	681 MPa	612 MPa
$R_{p0.2}$	609 MPa	490 MPa

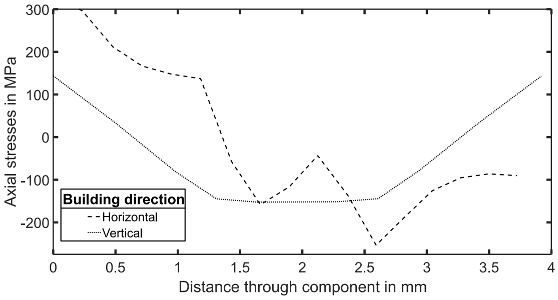


Fig. 1. Axial residual stress approximation along a path through the smallest cross section of the fatigue specimen for the vertical and horizontal building directions.

Although the FE calculations using the AWI [21] are only a rough approximation and are also related to the specimen geometry and the process parameters, which are mostly not included in the approximations, the values are in general agreement with those reported for the near surface area of a fatigue specimen made of 316L steel in the study by Hatami et al. [5]. The stresses in longitudinal direction reported in the study by Hatami et al. [5] for an as-built specimen are given as about 300 MPa and 600 MPa, depending on the process parameters. They were measured up to a depth of around 100 $\mu\mathrm{m}$. Due to the mesh size, data in comparable steps are not available; however, the general size seems to agree. For the vertically built specimen, the axial stresses seem to be quite symmetric with respect to the longitudinal axis, while the stresses in the horizontally built specimen change from tensile to compressive from the top to the bottom (orientation during the building process). Comparing both curves, the stresses within the horizontally built component are predicted higher, both tensile and compressive, and the component also shows larger areas of compressive stresses than the vertically built specimen. Reasons for the change of direction of the horizontal curve around 2.5 mm and its stepwise shape might be the removal from the platform, the chosen layer height and mesh size, although a more symmetrical shape would be expected. Again, the overall FE results seem to be in general agreement regarding the tensile amount near the surface; however, experimental validation is required.

3.3 Results and discussion

According to Eq. (1), the fatigue limit for a 50% survival probability of the whole component will be determined. Although the absolute values of residual stresses and calculated fatigue strength may not be used for direct comparison and prediction, the influence of the residual stress sensitivity can be evaluated. Fig. 2a shows the survival probability distribution of the vertically built specimen under consideration of the approximated residual stress distribution. Furthermore, different stress sensitivity values are evaluated. In Fig. 2b, the results for the horizontally built specimen are shown. The predicted fatigue limits for the vertically built specimens are lower than those for their horizontally built counterparts. This is based on the lower material properties as given in Tab. 1 and the different residual stress distribution.

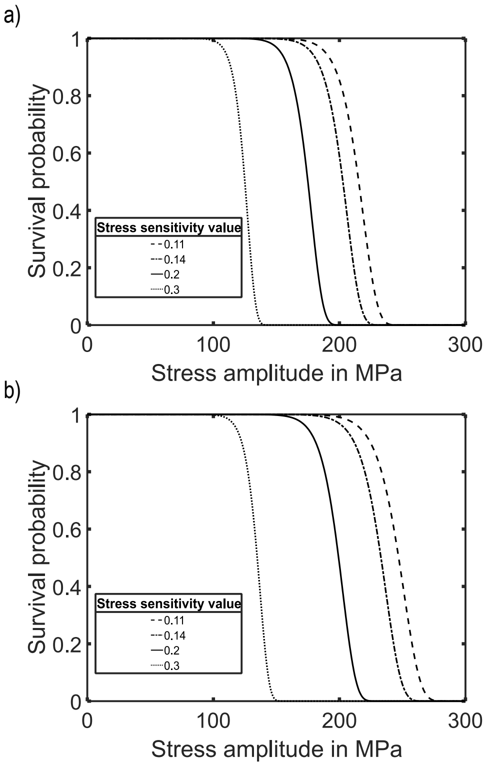


Fig. 2. Effect of different residual stress sensitivity values on the predicted fatigue survival probabilities: (a) vertically built; (b) horizontally built.

As expected, the curves are aligned according to the underlying residual stress sensitivity value from left to right starting with the highest value of 0.3. Furthermore, the curves are parallelly shifted while the shape is not affected by the stress sensitivity. The shape of the curves is determined via the Weibull exponent as shown, e.g., in the study by Zeißig and Jablonski [23]. According to Fig. 2, choosing a value of 0.3, in the absence of experimental data for this parameter, might be a very conservative approach and might unnecessarily restrict the application of components. The differences between the predicted endurable load stresses for a survival probability of 50% for the two curves on the left, with stress sensitivity values of 0.3 and 0.2, are 65 MPa and 50 MPa, respectively. Further investigations regarding the residual stress sensitivity of L-PBF components, therefore, seem to be advisable.

Fig. 3 shows the calculated fatigue stress amplitude values given in Fig. 2 in relation to the underlying fatigue strength which is taken as residual stress-free. It should be noted that according to Tab. 1, these values are different for the two manufacturing directions. Overall, the predicted reduction of the fatigue limit in terms of the base values is in the same order for both directions. However, the difference in the calculated fatigue limit for the lowest and highest chosen residual stress sensitivity value is significant and ranges between about 50% and about 90% of the base value. Comparing this to the results stated in the study by Leuders et al. [6] where the effect of different heat treatments on the fatigue limit of 316L was investigated, the impact on the fatigue limit seems to be overestimated in the calculations.

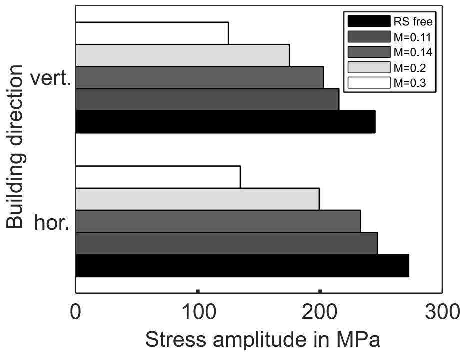


Fig. 3. Effect of different residual stress sensitivity values on the predicted fatigue stress amplitudes for vertical and horizontal building direction and considering residual stress-free state. (Survival probability of 50%).

4. SUMMARY AND OUTLOOK

As has been shown, in fatigue calculations, the effect of the residual stress sensitivity on the results should not be neglected. Experimental validation in terms of this parameter as well as for residual stress distributions of entire components seems to be necessary.

Predictions of the fatigue strength of components manufactured via L-PBF are difficult as there are numerous influencing parameters due to the building process, and their variations have an impact on the final results. Furthermore, results have no general validity for all materials as the material characteristic governing the fatigue limit may not be the same for every material; hence, individual investigations regarding each material seem to be necessary. Therefore, with respect to the shown approach, in the future, the microstructure and local deviations of the material properties should also be included in order to make it more versatile and enhance its applicability.

REFERENCES

1. Pelleg J. Additive and Traditionally Manufactured Components: A Comparative Analysis of Mechanical Properties. Amsterdam (NL): Elsevier; 2020.
2. Kruth J-P, Badrossamay M, Yasa E, Deckers J, Thijs L, Van Humbeeck J, Zhao W. Part and material properties in selective laser melting of metals. Proceedings of the 16th International Symposium on Electromachining (ISEM XVI), 2010, 3-14.
3. Radaj D, Vormwald M. Ermüdungsfestigkeit. 3rd ed. Berlin (DE): Springer-Verlag, 2007.
4. Mercelis P, Kruth J-P. Residual stresses in selective laser sintering and selective laser melting. Rapid Prototyp. J. 2006; 12(5): 254-265.
5. Hatami S, Ma T, Vuoristo T, Bertilsson J, Lyckfeldt O. Fatigue Strength of 316 L Stainless Steel Manufactured by Selective Laser Melting. J. of Materi Eng and Perform 2020; 29(5): 3183-3194.
6. Leuders S, Lieneke T, Lammers S, Tröster T, Niendorf T. On the fatigue properties of metals manufactured by selective laser melting – The role of ductility. J. Mater. Res. 2014; 29(17): 1911-1919.
7. Keller N. Verzugsminimierung bei selektiven Laserschmelzverfahren durch Multi-Skalen-Simulation [dissertation]. Bremen: University of Bremen, 2017 [cited 6 July 2017]. Available from: http://nbn-resolving.de/urn:nbn:de:gbv:46-00105808-15
8. Zhang Y, Jung Y-G, Zhang J. Multiscale Modeling of Additively Manufactured Metals: Application to Laser Powder Bed Fusion Process. Amsterdam (NL): Elsevier; 2020.
9. Zhang B, Li Y, Bai Q. Defect Formation Mechanisms in Selective Laser Melting. Chin. J. Mech. Eng. 2017; 30(3): 515-527.
10. Nadot Y, Nadot-Martin C, Kan WH, Boufadene S, Foley M, Cairney J, Proust G, Ridosz L. Predicting the fatigue life of an AlSi10Mg alloy manufactured via laser powder bed fusion by using data from computed tomography. Addit. Manuf. 2020; 32(3): 100899.
11. Mertens A, Reginster S, Paydas H, Contrepois Q, Dormal T, Lemaire O, Lecomte-Beckers J. Mechanical properties of alloy Ti–6Al–4V and of stainless steel 316L processed by selective laser melting: Influence of out-of-equilibrium microstructures. Powder Metall. 2014; 57(3): 184-189.
12. Wang D, Liu Y, Yang Y, Xiao D. Theoretical and experimental study on surface roughness of 316L stainless steel metal parts obtained through selective laser melting. Rapid Prototyp. J. 2016; 22(4): 706-716.
13. Vrancken B. Study of Residual Stresses in Selective Laser Melting [dissertation]. Leuven (BE): KU Leuven, 2016 [cited 10 May 2017]. Available from: https://lirias.kuleuven.be/1942277
14. Weibull W. A statistical theory of the strength of materials. Ingeniörsvetenskapsakademiens handlingar 151. Stockholm (SE): Generalstabens Litografiska Anstalts Förlag, 1939.
15. Weibull W. A Statistical Distribution Function of Wide Applicability. J. Appl. Mech. 1951; 18(3): 293-297.
16. Bomas H, Mayr P, Schleicher M. Calculation method for the fatigue limit of parts of case hardened steels. Materials Science and Engineering: A 1997; 234: 393-396.
17. Macherauch E, Kloos K-H. Bewertung von Eigenspannungen. Härterei-Technische Mitteilungen, Beiheft Eigenspannungen und Lastspannungen, Moderne Ermittlung – Ergebnisse – Bewertung 1982; 175-194.
18. Jablonski F. Rechnerische Ermittlung von Dauerfestigkeitskennwerten an einsatzgehärteten Proben aus 16 MnCrS 5 unter Berücksichtigung von Mittel- und Eigenspannungen [dissertation]. University of Bremen. Aachen (DE): Shaker Verlag, 2001.
19. FKM Forschungskuratorium Maschinenbau e.V. FKM-Richtlinie; Rechnerischer Festigkeitsnachweis für Maschinenbauteile. 6th ed. Frankfurt am Main (DE): VDMA-Verlag, 2012.
20. Gläßner C, Blinn B, Burkhart M, Klein M, Beck T, Aurich JC. Comparison of 316L test specimens manufactured by Selective Laser Melting, Laser Deposition Welding and Continuous Casting. In: Schmitt RH, Schuh G, editors. 7. WGP-Jahreskongress. 2017 5-6 Oct; Aachen, Germany. Aachen (DE): Apprimus Verlag, 2017; 45-52.
21. Abaqus Welding Interface 2017, User Manual, AWI Version AWI_2017-5. Dassault Systems Simulia Corp., 2018.
22. Abaqus/CAE 2017. Dassault Systemes Simulia Corp., 2016.
23. Zeißig M, Jablonski F. Comparison of different approaches to model fatigue for additively manufactured specimens considering production related characteristics. Procedia Struct. Integr. 2022; 38(5): 60-69.

The authors gratefully acknowledge the financial support by the Deutsche Forschungsgemeinschaft (DFG - German Research Foundation) under contract no. 275999847.

Michaela Zeißig: https://orcid.org/0009-0001-8249-9495

Frank Jablonski: https://orcid.org/0000-0002-9670-0540

DOI 10.2478/ama-2023-0034

ABILITY OF BLACK-BOX OPTIMISATION TO EFFICIENTLY PERFORM SIMULATION STUDIES IN POWER ENGINEERING

Lukas PETERS*, Rüdiger KUTZNER*, Marc SCHÄFER, Lutz HOFMANN*****

*University of Applied Sciences and Arts Hannover, Faculty I – Electrical Engineering and Information Technology,
Ricklinger Stadtweg 120, 30459 Hannover, Germany
**Siemens Energy Gas and Power Combustion Systems, Mellinghofer Str. 55, 45473 Mülheim a.d. Ruhr, Germany
***Leibniz University Hannover, Institute of Electric Power Systems, Electric Power Engineering Section,
Appelstraße 9a, 30167 Hannover, Germany

lukas.peters@hs-hannover.de, ruediger.kutzner@hs-hannover.de, schaefermarc@siemens-energy.com, hofmann@ifes.uni-hannover.de

received 14 November 2022, revised 9 February 2023, accepted 12 February 2023

Abstract: In this study, the potential of the so-called black-box optimisation (BBO) to increase the efficiency of simulation studies in power engineering is evaluated. Three algorithms ("Multilevel Coordinate Search" (MCS) and "Stable Noisy Optimization by Branch and Fit" (SNOBFIT) by Huyer and Neumaier and "blackbox: A Procedure for Parallel Optimization of Expensive Black-box Functions" (blackbox) by Knysh and Korkolis) are implemented in MATLAB and compared for solving two use cases: the analysis of the maximum rotational speed of a gas turbine after a load rejection and the identification of transfer function parameters by measurements. The first use case has a high computational cost, whereas the second use case is computationally cheap. For each run of the algorithms, the accuracy of the found solution and the number of simulations or function evaluations needed to determine the optimum and the overall runtime are used to identify the potential of the algorithms in comparison to currently used methods. All methods provide solutions for potential optima that are at least 99.8% accurate compared to the reference methods. The number of evaluations of the objective functions differs significantly but cannot be directly compared as only the SNOBFIT algorithm does stop when the found solution does not improve further, whereas the other algorithms use a predefined number of function evaluations. Therefore, SNOBFIT has the shortest runtime for both examples. For computationally expensive simulations, it is shown that parallelisation of the function evaluations (SNOBFIT and blackbox) and quantisation of the input variables (SNOBFIT) are essential for the algorithmic performance. For the gas turbine overspeed analysis, only SNOBFIT can compete with the reference procedure concerning the runtime. Further studies will have to investigate whether the quantisation of input variables can be applied to other algorithms and whether the BBO algorithms can outperform the reference methods for problems with a higher dimensionality.

Key words: black-box optimisation, power plant engineering, derivative-free optimisation, simulation studies in power engineering

1. INTRODUCTION

Many engineering tasks are, at their core, worst- or best-case analyses and thus optimisation problems in a mathematical sense. Furthermore, for many of these analyses, the use of simulation models is necessary, and simulation models can be expensive to compute. In some cases, the inner structure of models is too complex to directly investigate the influence of certain parameters, and the only way to obtain the needed information is the simulation. Calculations showing this behaviour can be characterised as black-boxes [1] (p. 1). In other applications, the term black-box could also be used for real physical experiments, where the correlation of the conditions of a system and the result of the experiment is unknown. Referring to simulations, it is also possible that models from external sources are given as black-box codes to protect intellectual property. For all these cases, optimisation methods that do not need any information about the structure of the function to be optimised are necessary. These methods are called black-box optimisation (BBO).

To avoid any confusion, the nomenclature used in this article is explained at this point: The term BBO is connected to the terms derivative-free optimisation (DFO) and simulation optimisation. Even though the terms emphasise on different aspects, all three might be applicable to most of the methods categorised with one of them. DFO emphasises the absence of information about the derivatives for the optimisation algorithm, whereas simulation optimisation is linked to the optimisation of experiments with simulation models. In the field of the project, the most important aspect is the universality of the methods for different use cases or objective functions. To emphasise the characteristic of not needing further information about the actual optimised function, the term BBO is the most appropriate for this study and will therefore be used.

There exist various reviews in the field of BBO [2–6], and beneath a high number of applications in biology, medicine, logistics or operations [5] (p. 358), these methods have also been used in engineering, e.g., for shape optimisation [7–12], electronic components [13, 14], control design [15] and power engineering [16, 17]. More recently, different methods have been applied in fluid dynamics [31–33], in chemical engineering [34, 35] or to tune parameters of other algorithms [36].

First algorithms that can be used for BBO were developed in the 1960s. One of those algorithms, the Nelder–Mead Simplex

Algorithm [27], is still used because of its straightforward principle that makes it easily applicable in different scenarios. After the first use of trust region methods [28] in 1969, in the 1970s and 1980s, genetic [29] and stochastic [30] algorithms emerged. Due to the increasing computational capabilities from then on, the number of algorithms and applications has increased enormously. Rios and Sahinids [3] have vividly illustrated the progress in their work.

While the performance comparison of optimisation algorithms is mostly based on academic test functions (see, e.g., [3]), the applications listed previously do focus more on the specific problem to solve and less on the evaluation of different algorithms. Using an algorithm for a single problem allows for tuning this algorithm to efficiently solve this problem, which might lead to a worse performance of this algorithm for other applications.

In this project, a variety of BBO methods were considered for different applications in the field of power engineering and evaluated concerning their potential to simplify engineers' daily work. For this simplification, the key is to find optimisation methods that are easy to implement and use as well as reliable. Following this goal, the present study introduces this topic by evaluating different methods mentioned in the corresponding reviews and answer the question whether BBO methods are competitive to manual procedures that are currently applied.

For the models used to develop and analyse gas and steam turbine control systems, models with this characteristic are used, e.g., to investigate damping capabilities of the control system for dynamic events of the electric grid, to define the highest possible rotational speed in case of a load shedding event or to parametrise the controllers.

Traditionally, mathematical optimisation is performed using information about the function's derivatives. Unfortunately, there are functions whose derivatives are very or even prohibitively expensive to obtain – as it is the case for the field mentioned previously. Consequently, the field of BBO has been driven forwards in mathematics over the past decades as computational power has massively increased.

Literature research and first attempts are promising that BBO can be easily applied to different problems and that BBO algorithms could improve the efficiency of simulation studies, in this case in turbine control system engineering.

In this study, different algorithms were applied to a variety of exemplary problems from the aforementioned sphere of action. The study will answer to which extend BBO can simplify the corresponding studies. The study also shows that BBO methods can – considering both of the algorithms and the accuracy of the found solution – efficiently solve optimisation problems in the field of power plant control engineering. Nevertheless, running parallel functions evaluations and limiting the resolution of the input variables are necessary features for an algorithm to be competitive even for computationally cheap objective functions.

2. SIMULATION STUDIES IN POWER PLANT CONTROL ENGINEERING

In 1999, Lu wrote in "Dynamic modelling and simulation of power plant system" [18] that in the previous decades, different models of power plants had been used to compute steady-state operation as well as to predict dynamics in case of events, like failures or incidents, and – of course – for usage in hardware training simulators.

Currently, the so-called digital twins of turbines or even whole power plants mirror all signals to indicate and forecast a plant's behaviour in real time [19]. Simulation studies are inevitable not only in the development of different components, such as turbine blades, but also in testing logics for control systems.

While there exist models in every complexity for different simulation scenarios to serve the most suitable computation of the plant's behaviour, the efficiency of the simulation studies themselves is not always optimal. In many cases, studies are performed manually and/or by computing large grids of input parameters. These studies might be led by an engineer's experience and therefore miss potentially interesting aspects, and additionally, the duration of executing these studies can be rather long.

Semi-automating parts of the studies might significantly decrease the time consumption of the studies. One idea to automate parts of such studies is using optimisation algorithms specially designed to solve problems with computational expensive functions, which will be referred to as BBO. Serving the aim to solve – at best – all possible problems with a single algorithm in practice, the following two use cases are chosen for a first evaluation of different BBO methods. The two use cases differ enormously in their computational costs and complexity and therefore cover a large percentage of the scale of problems that are subjects for this project.

2.1. Use Case 1: gas turbine overspeed simulation

Fig. 1 shows a load shedding event of an unspecified gas turbine power plant.

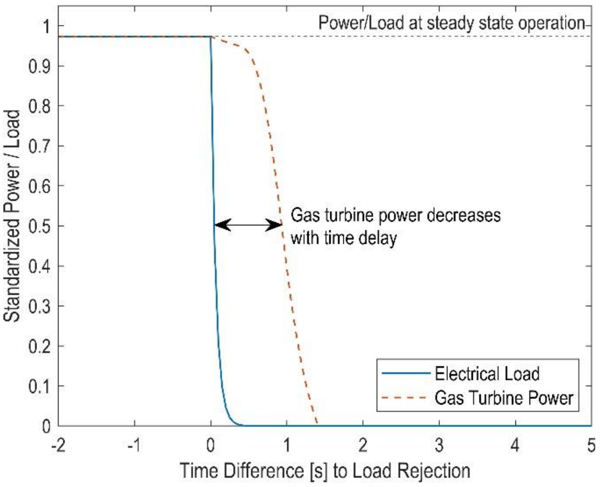


Fig. 1. Electrical load and gas turbine power output during a load shedding event

Here, the load from the electric grid and the effective power generated by the gas turbine are plotted over time. For steady-state operation with a constant rotor speed, both quantities are close to equal, which means that not only the power the gas turbine provides is approximately the same as the power of the electric grid, but also that the torque the gas turbine transmits M_{turb} on the shaft has the same amount as the counter-torque necessary to operate the compressor of the gas turbine M_{comp} and the counter-torque from the electric load M_{load}:

$$M_{turb} + M_{comp} + M_{load} = 0 = J_{shaft} \cdot \dot{\omega} \qquad (1)$$

where J_{shaft} is the moment of inertia of the rotating components of the turbine and the generator and $\dot{\omega}$ is the change in the angular velocity of the rotor. Losses resulting, e.g., from friction within the bearings are included within M_{turb}. For any difference in the torques, the shaft will be accelerated or decelerated, as shown in Eq. 1, until the power provided by the gas turbine P_{turb} is reduced or a higher angular velocity is reached at an equilibrium of the torque from the turbine and the counter-torque from the compressor, which depend on the rotational speed:

$$P_{turb} = (M_{turb} + M_{comp}) \cdot \omega \qquad (2)$$

Following Eq. 2, it is clear that whenever the load changes, the power of the turbine has to be adjusted to keep the angular velocity *ω* and the rotational speed *n=ω/2π* constant.

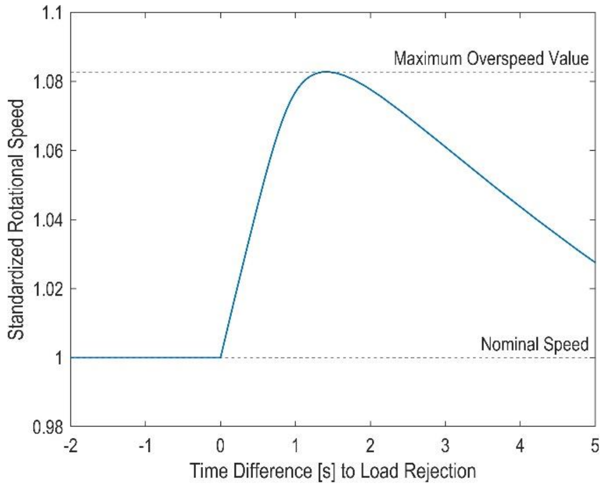


Fig. 2. Rotational speed of a gas turbine during a load shedding event

When the generator is disconnected from the grid, the electrical load does no longer provide a counter-torque, and the power of the turbine will – following Eq. 2 – result in an acceleration of the rotor until the control system has reduced the fuel supply and therefore the power.

A load shedding event (see the blue line in Fig. 1) is a highly dynamic process, and the turbine control systems are not able to immediately adjust the supplied power via the supplied fuel mass flow. Firstly, the fuel gas in the piping volume between the control valve and the combustion chamber (CC) (see Fig. 3) will expand into the CC and therefore be still converted into thermal power, which results in a delay between lowering the power setpoint of the gas turbine and the decrease in the actual power. Secondly, grid code requirements and the cycle time of the control system result in a maximum reaction time of up to 250 ms before adjusting the positions of the fuel gas control valves. Both delays have an influence on the maximum rotational speed (see Fig. 2), but the reaction time of the control system is a fixed parameter. Once the control system has adjusted the corresponding parameters, the power of the gas turbine decreases and the rotor speed as the effective power output of the gas turbine turns negative (see Figs. 1 and 2).

The increase in the rotational speed must – of course – not exceed the mechanical limits of the shaft. Therefore, load shedding events are simulated with different scenarios (e.g., normal load shedding, clamping of the control valve and freezing of the controller) with every update of a gas turbine to identify worst-case scenarios, i.e., finding the maximum overspeed value for all possible combinations of the corresponding parameters within the respective ranges, and ensure that the mechanical integrity is not violated.

The resulting maximum rotational speed for a load shedding event is influenced by several parameters (see Eq. 3), e.g., the state of the fuel supply system, ambient conditions and – of course – the current mechanical power of the turbine in the moment of the event. For a complete overspeed analysis, around 10 parameters can be considered. As mentioned previously, the simulation deals with a complex model. The overspeed as an objective function is therefore described as a black-box function with the following characteristics:

$$n_{max} \\ = f_{blackbox}(p_{gas}, H_{ux}, t_{gas}, P_{setpoint}, n_{trip}, n_0, t_a \qquad (3)$$

where p_{gas} is the fuel gas supply pressure, H_{ux} is the lower heat value of the fuel gas, t_{gas} is the fuel gas temperature, $P_{setpoint}$ is the load setpoint for the turbine, n_{trip} is the trip limit for rotational speed, n_0 is the net frequency at the time of load shedding *and* t_{amb} is the ambient air temperature.

Models for the simulation of the turbine operation include the thermodynamics and the control system and are therefore complex. The simulation for a single set of parameters takes up to 20 min. Computing a reasonably high number of variations of influencing parameters for an overspeed analysis may require several thousand simulations and therefore requires an enormous runtime.

Applying methods of BBO to this problem, the required number of simulations can be reduced because the methods will search for the actual optimum, instead of covering the whole range of variables in a constant resolution. Therefore, the analysis can be executed faster. For this work, a simplified overspeed analysis was performed by BBO methods. The resulting maximum speed of the gas was computed only for varying conditions of the fuel itself, namely, lower heat value H_{ux}, fuel gas temperature T_{gas} and fuel gas supply pressure p_{gas}. Furthermore, only normal load shedding was simulated without additional failure scenarios.

2.2. Use Case 2: transfer function identification

As a second use case, transfer function coefficients for a delayed ramp function (ramp function applied to a lag element) are to be identified. The behaviour of a delayed ramp is used to approximate the mass flow supplied to a gas turbine during a load shedding event when the fuel gas valves are closed instantly.

The physical background is simple: to reduce the fuel gas mass flow, the corresponding valve is moved from a certain position for the steady-state operation to its closed position. This movement can me modelled as a ramp. Between the valve and the CC, where the fuel gas is finally converted into thermal energy, the fuel gas must flow through a piping volume (see Fig. 3). This volume dampens the reduction of the mass flow rate. Furthermore, this volume is under a higher pressure than the CC. Due to the pressure compensation, fuel gas will flow into the CC from the piping for a short time after the valve is completely closed. This results in the dynamics of the mass flow rate, as shown in Fig. 4.

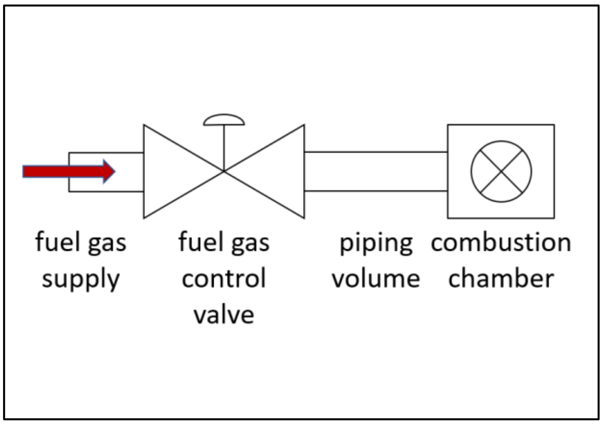


Fig. 3. Schematic fuel gas piping

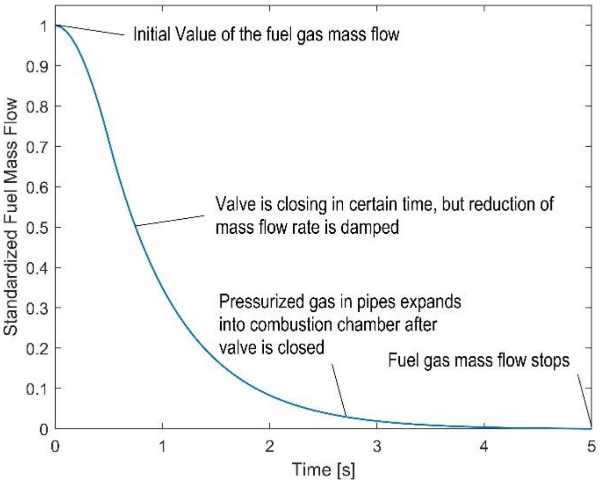


Fig. 4. Decrease in the supplied fuel mass flow during and after valve closing

As a reference value used to compare the results of the BBO methods, a minimum error can be manually determined as follows:

A lag element can be described with the following differential equation:

$$T \cdot \dot{y}(t) + y(t) = K \cdot u(t). \tag{4}$$

where *u* is the input of the lag element and *y* is the output of the lag element.

Integration and rearranging of this equation give the following (with *K* assumed to be included in *u(t)*):

$$T[(y(t) - y_0)] = U(t) - Y(t) \tag{5}$$

where *U* and *Y* are antiderivatives of *u* and *y* for *U*(0) = *Y*(0) = 0.

The input of the lag element is separated into different functions: a ramp function from *t = 0* to *t* = T_r. At T_r, the ramp has reached its final level and is from here on constant (in the case discussed here, it is 0). A ramp is defined as follows:

$$u_1(t) = (y_1 - y_0) \cdot \frac{t}{T_r} + y_0 \ \text{ for } 0 \leq t \leq T_r \tag{6}$$

where y_0 is the start value of a ramp, y_1 is the end value and T_r is the time between the start and the end.

As mentioned, the value of the input function after T_r is constant with a value of y_1:

$$u_2(t) = y_1 \ \text{ for } t \geq T_r \tag{7}$$

Given a set of *n* measured points of the real behaviour [t_i, $y(t_i)$] for *i = 1 to n*, a system of linear equations can be set up and solved using the least-squares method. Given a set [*T*, T_r], an approximated value y_{ap} for the corresponding lag element can be calculated for each t_i and the error between the measured and the calculated value can be taken as the sum of squares of the errors for each point t_n

$$\Delta y(T_r, T) = \sum_{i=1}^{n} (y_{ap}(T_r, T, t_i) - y(t_i))^2 \tag{8}$$

for any given set of pairs [t_i, $y(t_i)$].

For this application, it is also possible to calculate the time coefficients T_r and *T* in a more direct manner, with the following vectors:

$$\vec{y} = (y_1, y_2, \ldots, y_n)^T \\ \vec{U} = (U_1, U_2, \ldots, U_n)^T \\ \vec{Y} = (Y_1, Y_2, \ldots, Y_n)^T \tag{9}$$

Eq. 5 provides a value for *T* with a given value T_r. With this, it is possible to minimise Eq. 8 only as a function of T_r and determine *T* afterwards.

Identifying functions like this is necessary to develop simplified models of gas turbine power plants that are, e.g., used in studies to analyse the interaction with the electrical grid.

Basically, this is not a black-box function as these values can be calculated numerically, and furthermore, in terms of computational expenses, it is very cheap. Nevertheless, this problem can serve as a simple version of the identification of more complex models and shows how the performance differs between certain levels of complexity.

3. BLACK-BOX OPTIMISATION

When explaining BBO, it is necessary to define two terms: black-boxes and optimisation.

Black-box. A black-box is, according to Kimiaei and Neumaier, "[…] an oracle that returns for a given x ∈ $\mathbb{R}^n$ the function value f(x)" [1] (p. 1). The main characteristic is the non-visibility of the inner computational process of the function. Applying a set of input variables to a black-box function, it will return a (set of) output variable(s), and there is no way to investigate the procedure of computing this output with reasonable effort.

Optimisation: From a mathematical point of view, optimisation is simply the determination of – local or global – minima and maxima of a given function. Commonly, this calculation is driven by derivatives of the function.

Combining these two definitions, we need to find those optima without having a real insight into the function and therefore no information about its derivatives.

Over the past decades, various methods for BBO have been introduced for a variety of applications, but a universal algorithm that can be applied to any problem is not developed yet [7] (pp. 123–124).

There are several categorisations for BBO or DFO methods, and in this study, the categorisation suggested by Rios and Sahinidis [3] was used. They divide algorithms based on the following properties:

- Direct methods or model-based methods
  The (expensive) objective function can either be evaluated itself or an approximation (or surrogate) model can be used for optimisation.
- Global or local methods
  Local methods are designed to quickly find the next local optimum but miss the capability to check if there exist other optima, whereas global methods will have this capability and will determine an absolute optimum.
- Stochastic or deterministic methods
  Categorisation by a deterministic or randomised determination of the algorithm's search steps.

In addition to Rios and Sahinidis [3], Larson, Menickelly and Wild [7] have also published a profound review of methods for DFO and BBO, where examples for each category are given. Furthermore, methods differ in their ability to deal with boundaries of the search space. They might require boundaries in the form of equality or inequality constraint, while others cannot process any limitation of the input variables.

Of course, there also exist numerous methods that are hybrids and combine the categories mentioned before.

4. SELECTED BBO ALGORITHMS

For the use cases to be addressed in this study, there are several conditions of the objective functions that the optimisers will have to cope with. These are listed as follows:

- Noise
  The models used might contain stochastically influenced functions, and therefore, the same input can provide a slightly different simulation output.
- Boundaries
  Input variables will mostly have lower and upper limits. The highest or lowest values might therefore lie on these boundaries, instead of forming an optimum within the search region.
- Local Optima
  Due to the interdepending influence of different parameters on the objective function, multiple local optima might occur. For most cases, a global optimum or the highest/lowest function value within a search region is important to be found, and therefore, the optimisation should not stop at a local optimum.

The existence of noise is applicable for both use cases mentioned: in the operation simulation for the overspeed analysis, the simulation of measurements within the gas turbine model is modified with artificial noise as they would be noisy in the operation of a real plant. For use case 2, we have to assume that the measured flow rates used to identify the transfer function are provided with usual measurement errors. The fuel gas parameters for the overspeed analysis are limited to the fuel definition of a contract for a plant. The boundaries for the parameters of the transfer function are, in this case, chosen by the previous numerical calculation of the problem. For a totally unknown transfer function, the upper limits would have to be estimated by the user, while the lower limits could always be set to 0. Concerning the existence of multiple local optima, the overspeed analysis could have multiple local tallest points due to the interdependency of the variables. For use case 2, the previous evaluation has shown that only one optimum exists, but considering that the methods should be applied to black-box functions with an unknown structure, the existence of local optima should always be considered.

Taking these conditions into account and based on the nomenclature provided by Rios and Sahinidis [3], only global optimisation algorithms are able to appropriately solve these problems. Additionally, all methods that do not accept boundaries cannot be considered.

Furthermore, in terms of usability, only few algorithmic parameters of a BBO method should be needed to be adjusted for any new problem.

To evaluate how well BBO methods perform applied to the exemplary problems (Section 2), methods that are available as open source and, at best, as MATLAB® software will be tested preferably. Two algorithms that meet the conditions outlined previously are the Multilevel Coordinate Search (MCS) algorithm [20] and the Stable Noisy Optimization by Branch and Fit (SNOBFIT) algorithm [21] by Huyer and Neumaier. A different algorithm that is a classically model-based approach is the blackbox algorithm [22] by Knysh and Korkolis. This algorithm is available as a Python® package and translated to MATLAB® for this study by the authors.

The MCS algorithm is a direct evaluation method, whereas SNOBFIT uses different local models, and blackbox uses a global model of the objective function; therefore, they cover a broad range of the algorithmic spectrum. Additionally, MCS and SNOBFIT showed satisfactory results in the review of Rios and Sahinidis [3] (pp. 1267–1287).

The concepts of the algorithms are briefly presented in the following paragraphs; for a complete description of all details, we refer to the corresponding papers noted previously.

4.1. Global optimiSation by MCS

The so-called MCS algorithm was developed by Waltraud Huyer and Arnold Neumaier in 1998 [20]. The search space is continuously divided into boxes that have a base point and a corresponding function value given by the objective function. Partitioning of the boxes is executed along the coordinate axis, and with each splitting of a box, a level s is increased. Any box can be split until this level reaches the level $s = s_{max}$, which indicates that a box is too small for another division [3] (p. 1253).

Fig. 5 shows a two-dimensional search space that has already been split several times (black lines). The point in the middle (green-yellow) provides the best function value. The rectangle highlighted in magenta is chosen to be split next, and the green-highlighted point (approximately at [0.5; 1.5]) marks the position for the next split.

The algorithm selects the boxes to be divided in the next iteration by the values of the objective functions of the base points (see Figs. 5 and 6). There are two options for splitting the boxes:

Splitting by rank: Boxes already having a high level of s can be chosen for splitting by rank. This means that the coordinate along which they will be divided will be chosen by how many splits have already been performed in each coordinate. Selecting the coordinate with the lowest numbers helps decrease unexplored spaces in the corresponding direction.

Splitting by expected gain: The second method for splitting is used for lower levels of s. Then, a local quadratic model of the objective function is built using points from previous iterations, and the splitting coordinate is chosen by optimisation of this model.

Additionally, local searches can be executed when a box has already reached the splitting level s_{max}.

As seen in Fig. 6, the algorithm has reached a higher density within a region of low function values (dark green), which means that in this case, the boxes were split by the expected gain. While proceeding the search, promising areas will be explored more and more. The convergence to a global minimum is ensured when s_{max} approaches to infinity, which means that the number of functions evaluations is not limited [20] (pp. 347–349).

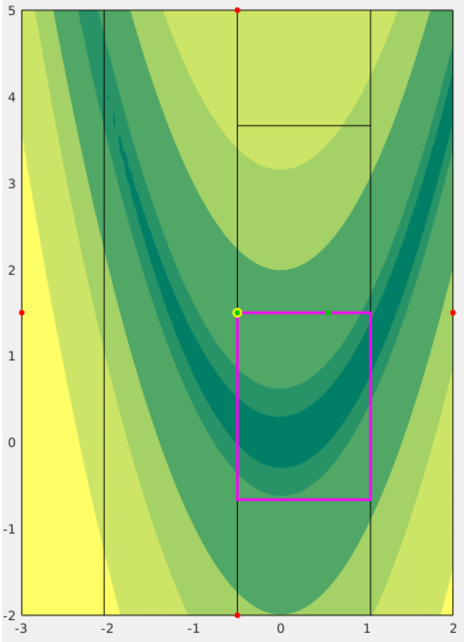


Fig. 5. Rectangle chosen to be split by the MCS algorithm marked in magenta [Figure: "MCS algorithm" by Najko32. License: CC BY-SA 4.0] [26]. MCS, multilevel coordinate search

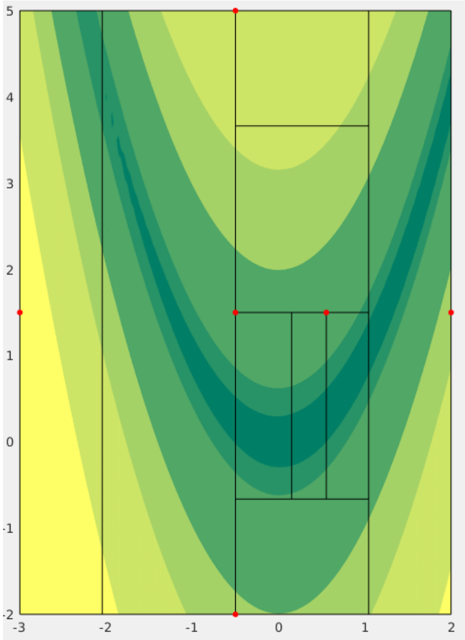


Fig. 6. Result of the iteration of the MCS algorithm based on Fig. 5 [Figure: "MCS algorithm" by Najko32. License: CC BY-SA 4.0] [26]. MCS, multilevel coordinate search

4.2. Stable noisy optimisation by branch and fit

Huyer and Neumaier [21] have published another promising algorithm for BBO: SNOBFIT. The sequence of the SNOBFIT algorithm is as follows:

Initialisation: Given a set of points within the search space and the objective function, the initial setup for the algorithm is built.

Consecutive iterations: Based on this set, SNOBFIT generates the user-defined number of required new evaluation points n_{req}. The search space is split into boxes so that each contains exactly one point. A smallness is computed for each box to have a measure on how often a box has already been split, and a local fit around each point is computed (or updated for old points). The split is performed along the coordinate axis that shows the highest variance of the points in the box to be split. The current best point and the corresponding function value are evaluated, and a local quadratic fit around this point will be computed. Points to be evaluated are prioritised in different classes:

- Class 1: By optimisation of the aforementioned fit, a new potentially best point can be generated.
- Class 2: Points determined by a trust region estimation around the so-called local point that shows a significantly lower function value than its nearest neighbours.
- Class 3: Points determined by a trust region estimation about non-local points.
- If less than n_{req} points have been chosen, two more classes of points can generated.
- Class 4: Points that split large boxes to explore uncovered areas of the search space.
- Class 5: Points determined by a randomised space-filling set of the missing number of points are created in a way that these points have the maximum distance to all other points.

All generated points are evaluated by the objective function, and a new iteration will start.

Stopping. The algorithm stops if for a defined number of consecutive iterations, no new point of class 1 has been generated, and the current best point is therefore assumed to be the global optimum.

4.3. Blackbox: a procedure for parallel optimisation of expensive black-box functions

Knysh and Korkolis [22] have provided a method for BBO that utilises a global response surface with radial basis functions (RBFs). The method is originally provided as a Python® package [23] and was implemented in MATLAB® for this study.

For the procedure, all variables are rescaled to a range of [0, 1] so that the search space equals a unit hypercube. Furthermore, after initial sampling, the function values are also rescaled to a range from 0 (best or lowest function value) to 1 (worst or highest function value).

The initial sampling set is generated by quasirandom sequences [24]. After the first approximation of the response surface is calculated, further sampling candidates are provided by an adapted version of the Constrained Optimization using Response Surfaces (CORS) algorithm [25]:

DOI 10.2478/ama-2023-0034

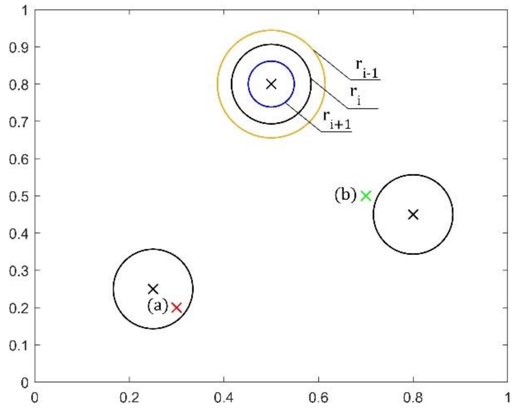


Fig. 7. Demonstration of the blackbox algorithm

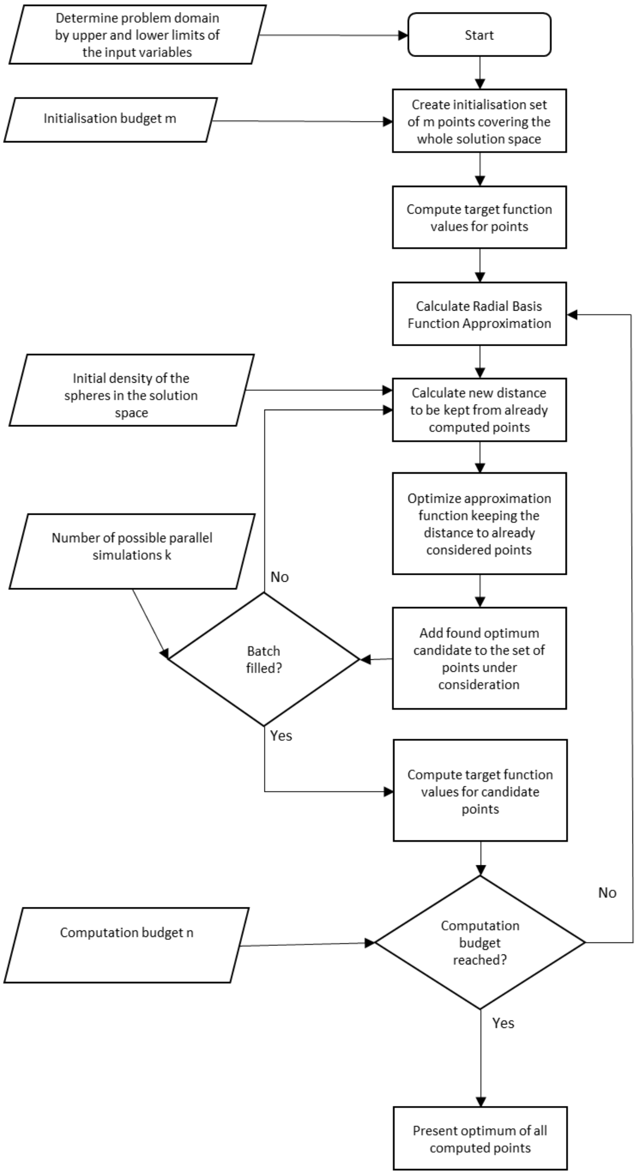

Determine problem domain by upper and lower limits of the input variables
Start
Initialisation budget m
Create initialisation set of m points covering the whole solution space
Compute target function values for points
Calculate Radial Basis Function Approximation
Initial density of the spheres in the solution space
Calculate new distance to be kept from already computed points
Optimize approximation function keeping the distance to already considered points
Number of possible parallel simulations k
No
Batch filled?
Add found optimum candidate to the set of points under consideration
Yes
Compute target function values for candidate points
No
Computation budget n
Computation budget reached?
Yes
Present optimum of all computed points

Fig. 8. Flowchart of the blackbox algorithm

Basically, the RBF approximation is updated in each iteration step by adding candidate points for the global optimum that minimise the fitted function under the following constraint: the new candidate must lie outside of a (hyper-)sphere around all other points (already sampled and current candidates, marked with a black x in Fig. 7), which is also shown in Fig. 7: the point (a) highlighted in red lies within the current radius r_i and is therefore not used as a new candidate, whereas the green point (b) meets the distance constraint. The radius of the spheres (i.e., the distance a new point must keep to all other sample points) decreases over the number of iterations (see r_{i-1}, r_i and r_{i+1} in Fig. 7). By this, the algorithm first explores the whole search space and then proceeds to a more local search. The speed of decreasing (see Fig. 8) is set by the user.

The algorithm stops after a user-specified number of function evaluation is reached (Fig. 8). Even though the authors do not mention any proven convergence of the algorithm, it will cover the search space densely, if the number of function evaluations is not limited and the values for the initial density (defining the starting value for the spheres' radii) and its decrease are chosen to focus on a global search. Thereby, the algorithm will determine a global optimum.

5. APPLICATION OF THE METHODS TO CHOSEN USE CASES

All algorithms are applied to both use cases. As the first of both applications (Section 2) – the gas turbine overspeed analysis – is expensive in terms of computation, only one run per algorithm is performed, whereas the identification of the transfer function is solved 20 times to see whether the quality of the solution found after varies between the executions of the code or not. To compare the algorithms, the following measures are considered:

- Number of function evaluations used to determine the global maximum
- Execution time for a single run of the algorithm
- Quality of the found global optimum function value
- Quality of the found optimum input variables

Each method is limited to about 200 function evaluations. As a benchmark for the real maximum of the overspeed value, a grid search is performed, that is – compared to how the grid search is performed in practice – considerably denser. In this manner, we aim to obtain a point as close as possible to the actual maximum for the overspeed value. To validate how the BBO algorithm performs in comparison to the standard procedure, another grid search[1] based on 216 equally distributed evaluation points was executed. This run will be referred to as "sparse grid search," whereas the formerly mentioned will be called "dense grid search."

If parallelisation is possible, the number of parallel function evaluations is set to 8. By this limit, the RAM of the testing computer works almost at capacity during the compilation process but does not exceed it, which would result in out-of-memory errors.

Additionally, the resolution vector for SNOBFIT was set to [H_{ux}: 1 MJ/kg, T_{gas}: 1°C, p_{gas}: 1 bar] (see Section 2.1). The resolution is based on the practical use: a) for each variation of the heat value, a new compilation of the model is necessary. For this reason, the number of possible values needs to be limited and

[1] In a grid search, a set of certain values to be considered for simulation is defined for each input variable. Combining the sets of all variables provides a grid covering the whole search space in a certain resolution. All points of the grid will be used to compute a value of the objective function and the optimum of these values is returned as the determined global optimum.

b) the overspeed value does not change substantially within the space between the values allowed by the resolution vector, and furthermore, the values will not be stated in higher accuracy in the documentation for customers.

The results and corresponding parameters are given as percentages of the values for the benchmark results to simplify the comparison.

The objective function is covered in a small MATLAB® function that manages the parallelisation of simulations, if necessary, and returns the overspeed values for the points requested for an iteration step by the algorithms.

For use case 2 (parameter identification of a transfer function, Section 2.2), the averaged values over all 20 runs are taken for evaluation of the performance. The variation of the results is compared separately. SNOBFIT and blackbox were able to compute four function values at the same time. Furthermore, the resolution vector for SNOBFIT was set to [t_r: 10^{-6} s; T: 10^{-6} s] as the standard search has shown that further quantisation does not substantially improve the found minimum but – of course – increases the number of calculations needed. Other parameters do not differ between use cases 1 and 2. Again, to run the algorithms, a simple function that returns the error values corresponding to the requested points and managing the parallelisation is sufficient.

5.1. Results for gas turbine overspeed analysis

For the gas turbine overspeed analysis, the objective functions are defined as a black-box function, as shown in Eq. 3. For the evaluation of the capability of BBO to solve problems based on computationally expensive simulations, the objective was reduced on the following parameters:

$$n_{max} = f_{blackbox}(p_{gas}, H_{ux}, t_{gas}) \tag{10}$$

Obviously, the most important question to answer is whether the methods are capable of providing a value sufficiently close to the reference value (dense grid search) for the objective function value (namely, the maximum overspeed value). Fig. 9 displays how close the computed optimum values are to the reference value determined by a dense grid search with multiple thousand simulations covering the search space. All methods provide a value that differs less than 0.15% from the reference optimum and therefore meet the previously mentioned requirement. Among these overall satisfactory results, blackbox provides an optimum value that is – considering noise – equal to the reference value.

Nevertheless, it is also relevant if the suggested optimum is close to the actual optimum in terms of coordinate values. Fig. 10 shows the deviation of the found optimum from the reference values for all methods. Noticeably, the divergence in the fuel gas supply temperature T_{gas} is enormous. By contrast, SNOBFIT, blackbox and the sparse grid search deliver values for the heat value H_{ux} and the fuel gas supply pressure p_{gas} that are close to the actual optimum coordinates. It is thereby clear that T_{gas} has only a minor influence on the maximum rotor speed.

Additionally, the points suggested by SNOBFIT and blackbox are both only insignificantly different from points within the top 5 points (best 0.15%) of the dense grid search and therefore extremely close to the reference optimum point. As MCS provides the worst (even if still good) value for the maximum rotor speed, the determined values for the input variables are also the farthermost from the reference optimum from the dense grid search.

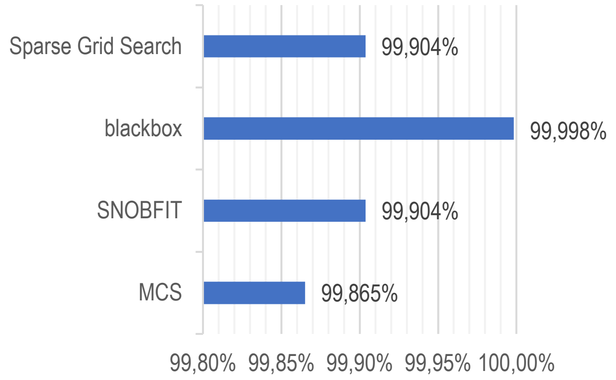
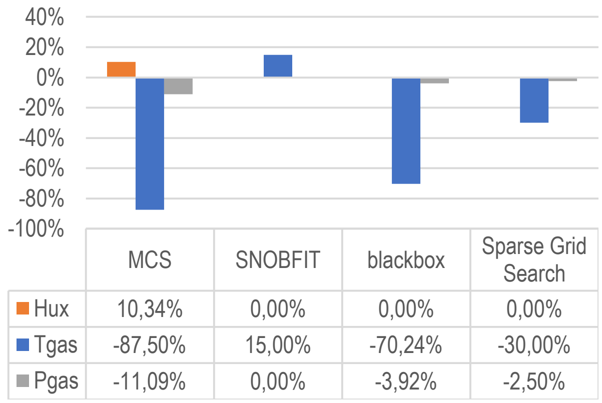


Fig. 9. Accuracy of determined optimum for BBO methods for use case 1. BBO, black-box optimisation; MCS, multilevel coordinate search; SNOBFIT, Stable Noisy Optimization by Branch and Fit



	MCS	SNOBFIT	blackbox	Sparse Grid Search
🟧 Hux	10,34%	0,00%	0,00%	0,00%
🟦 Tgas	-87,50%	15,00%	-70,24%	-30,00%
⬜ Pgas	-11,09%	0,00%	-3,92%	-2,50%

Fig. 10. Accuracy of optimum coordinates of BBO methods for use case 1. BBO, black-box optimisation; MCS, multilevel coordinate search; SNOBFIT, Stable Noisy Optimization by Branch and Fit

Provided that all methods compute adequate values for the value of interest, it is of same importance how efficient they find the corresponding solution. In Fig. 11, the numbers of function evaluations needed to determine the optimum are compared for the three tested BBO methods and the Sparse Grid Search as a reference to the currently used procedure. As already mentioned, the BBO methods were limited to about 200 iterations. Therefore, the grid search was designed to have a comparable computational effort.

While blackbox and MCS use all possible iterations (or due to local searches even more), SNOBFIT stops the iterations after 64 function evaluations after it has not been able to improve the current optimum for five iterations.

In this application, the computational expense between two simulations might differ significantly. The used Simulink® model is compiled to C-code, and an .exe file is built to run the actual simulation. Depending on which parameters change, the .exe file can be used again, and the model does not need to be compiled again. The actual simulation runs only few minutes on the system used, whereas the compilation might take up to 15 min. In use case 1, a change of the heat value requires a new compilation.

Fig. 12 shows the runtime of each method and, in this case, also for the dense grid search. Most remarkable is the runtime of MCS: it finishes the run with 225 function evaluations after more than 1 day and 8 h. By contrast, blackbox does only need about 7 h, whereas SNOBFIT needs a runtime of 1 h and 44 min, which is very close to the reference method, which needs 1 h and 14 min. The explanation for this is simple:

Firstly, MCS changes the coordinates of requested points only slightly, which results in a very high number of recompilations. Contrarily, both grid searches and SNOBFIT have a limited resolution of the axes and therefore provoke only few compilations. As the sparse grid search does only compile seven times, whereas SNOBFIT and the dense grid search need 21 compilations, it is faster than SNOBFIT even though it takes many more iterations.

The second obstacle for MCS is its seriality: it cannot be parallelised, while all other algorithm work with eight simulations at each iteration. This is the main reason that blackbox performs better than MCS even if it needs the same number of compilations.

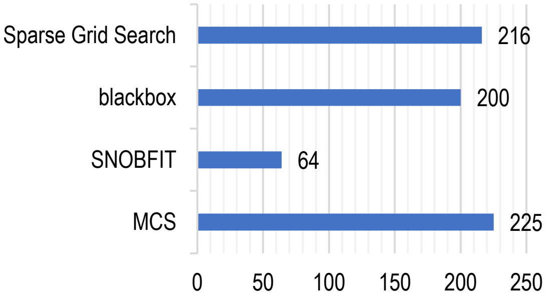


Fig. 11. Number of functions evaluations needed to solve use case 1. MCS, multilevel coordinate search; SNOBFIT, Stable Noisy Optimization by Branch and Fit

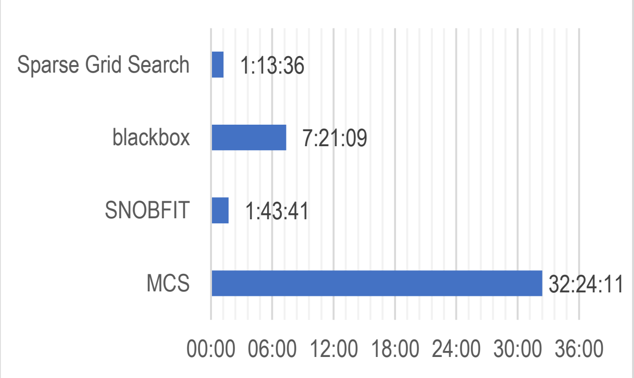


Fig. 12. Runtime of BBO methods needed to solve use case 1. BBO, black-box optimisation; MCS, multilevel coordinate search; SNOBFIT, Stable Noisy Optimization by Branch and Fit

This use case with a complex or computational expensive objective function brings up three important learnings:

- All methods provide sufficient solutions for the given problem.
- The performance of an algorithm strongly depends on the degree of parallelisation.
- A reasonably coarse variation of input parameters is extremely beneficial.

The following second use case will show if these aspects are also relevant for less complex and expensive objective functions.

5.2. Results for transfer function identification

Again, the first point of interest is how precise the identified optima are compared to the reference value. As mentioned earlier, for this application, the actual solution can be computed numerically. Tab. 1 shows the results of the three algorithms:

Tab. 1. Results for use case 2

Name	Standard	MCS	SNOBFIT	blackbox
Minimum error determined	0.229974	0.229977	0.230002	0.230126
Deviation from standard value	-	0.0011%	0.0123%	0.0658%
***Tr* for minimum error**	0.11893	0.11897	0.11895	0.11884
Deviation from standard value	-	0.0326%	0.0167%	-0.0731%
***T* for minimum error**	0.426205	0.426207	0.426212	0.426029
Deviation from standard value	-	0.0005%	0.0015%	-0.0414%

MCS, multilevel coordinate search; SNOBFIT, Stable Noisy Optimization by Branch and Fit.

It is clearly visible that all algorithms determine the optimum extremely close to the actual values in terms of the least error as well as in terms of the variables tr and T. Nevertheless, it is remarkable that for this use case, MCS delivers the most accurate optimum value and also very precise parameters. Even though both provided solutions are still more than acceptable, MCS is about 10 times more precise than SNOBFIT and about 60 times more precise than blackbox.

Having such close results, it is even more important to relate these results to the computational performance, which is shown in Figs. 13 and 14.

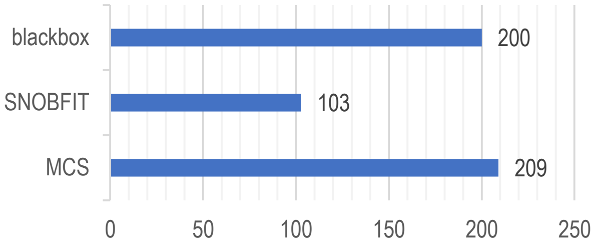


Fig. 13. Number of function evaluations needed to solve use case 2. MCS, multilevel coordinate search; SNOBFIT, Stable Noisy Optimization by Branch and Fit

Fig. 12 shows that while blackbox and MCS consume the whole computational budget, SNOBFIT finds the sufficient solution only using half the function evaluations.

Fig. 14 shows the average time the methods needed to execute use case 2 in the test. A clear ranking can be derived from these numbers: SNOBFIT also outperforms the other algorithms by solving the problem in only 5 s, while MCS needed about 12 s, and blackbox is even more slow, with 30 s.

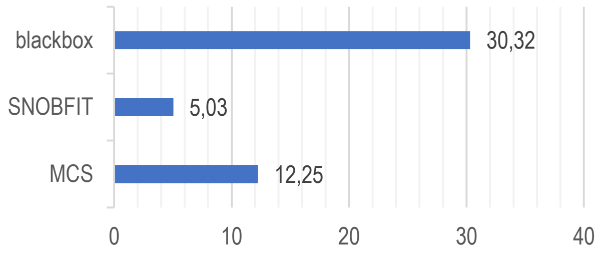


Fig. 14. Runtime of BBO methods [s] needed to solve use case 2. BBO, black-box optimisation; MCS, multilevel coordinate search; SNOBFIT, Stable Noisy Optimization by Branch and Fit

In terms of replicability of the results, MCS and blackbox only differ in the runtime over the 20 runs (MCS: minimum runtime: 11.61 s / maximum runtime: 13.78 s; blackbox: 29.15 s / 37 96 s), and SNOBFIT's performance is more volatile: the least number of function evaluations to find the optimum for the transfer functions parameters was 71, and the highest number of evaluations was 147 and the runtimes vary accordingly between 3.35 s and 6.81 s. Nevertheless, the quality of the found minimum error is only varying in the range about ±0.1%, and the actual parameters do not fluctuate more than 0.4% around the average values.

Even though SNOBFIT varies in its efficiency, the worst performance is still significantly above the best performances of blackbox or MCS.

For this use case, all methods are fully competitive to the reference value. SNOBFIT has clear advantages in terms of computational and time efficiency, whereas MCS is more precise. The significantly higher runtime of blackbox might not be dramatic in absolute terms but is still a remarkable disadvantage.

6. CONCLUSIONS

Three BBO methods (MCS, SNOBFIT and blackbox) were tested on two use cases with very high and very low computational expense. The application of the methods to our exemplary problems have been proven to be highly user-friendly: setting up a sufficient MATLAB® function to be called and the maximum number of function calls, and all methods run stable and provide satisfactory results. Only for SNOBFIT, the input of the resolution vector of the input variables is an additional value to be specified.

All methods provided optima for the objective function close to the set reference values and do therefore meet the most basic requirement.

Additionally, all coordinate values determined to be the optimum point were close to the actual points from the reference optimum. Even though MCS cannot compete in accuracy for the overspeed analysis, it is the most precise algorithm for the identification of the transfer function parameters.

In summary, the accuracy of the found optimum is satisfying for all algorithms in both applications out of common power engineering tasks, and therefore, methods of BBO can be considered for solving these kinds of problems as questioned in Chapter 1.

To be competitive with the procedures used currently, the computational efficiency (actual runtime of the algorithms) is nearly as important as the quality of the found solution. The runtimes for the identification of transfer function parameters (use case 2, Section 2.2) differ significantly but are still short in absolute terms. Still, it is to mention that SNOBFIT is significantly faster than MCS, and blackbox is – compared to this – slow in this use case. It is – of course – still much more efficient than the manually performed procedure to find the best fitting parameters.

Of course, the time efficiency is far more important when dealing with more expensive objective functions as the overspeed analysis. Here, MCS and blackbox cannot compete with the sparse grid search and SNOBFIT. The foundation for that discrepancy is the compilation of the models to be simulated, as described in Section 5.1.

Furthermore, the results point out that parallelisation is necessary for complex problems.

Referring to the questions addressed at the beginning of this article, it was shown that BBO methods are easily applicable to simulation studies and that they do have the potential to meet the tasks arising in the context of gas turbine control simulation. However, for computationally expensive functions, only SNOBFIT is, in general, competitive to a knowledge-based grid search as currently performed as a standard procedure, and it also shows the best performance for a cheap objective function.

7. OUTLOOK

While the methods tested for this study already show promising results, there are still several questions that should be answered in future work.

Firstly, the whole range of BBO algorithm has not been tested yet, and especially the field of genetic or evolutionary algorithms, which has seen a significant development in the past years is not included and should be considered in further studies. Currently, there are ongoing studies on the implementation of additional algorithm, which will be tested soon. The experience portrayed in this work is promising that the adaption of further methods will be successful.

Secondly, it was clearly shown that for the gas turbine overspeed analysis, the high number of necessary compilations is limiting the performance of the BBO methods, blackbox and MCS. The performance of SNOBFIT for this application is far more effective, which is based on less compilations due to the resolution vector. Therefore, we aim to adapt this resolution vector to all methods.

For a non-simplified overspeed analysis with more input variables, the BBO methods might be significantly more competitive to the standard grid-based procedure as it will perform far less efficient as the number of points increases exponentially with the dimensions. This results either in a worse accuracy or in a far longer runtime.

Considering that most parameters to adjust the algorithms' behaviour have been kept to standard values, as provided in the corresponding literature, the performance might be improvable. Combining this with the idea of the resolution vector for the input variables, the performance of all methods will possibly increase. Still, it needs to be ensured that an adjustment of the algorithmic parameters works properly for all use cases and does not provide benefits just for one problem.

Furthermore, the actual usability and universality must be tested on more applications. Studies testing the methods for the identification of a complex gas turbine model with real-world data and the adjustment of control parameters will follow.

REFERENCES

1. Kimiaei M, Neumaier A. Efficient Global Unconstrained Black Box Optimization. Mathematical Programming Optimization. 2022;*14*: 365-414. https://doi.org/10.1007/s12532-021-00215-9
2. Custódio AL, Scheinberg K, Vicente LN. Methodologies and Software for Derivative-free Optimization. In Advances and Trends in Optimization with Engineering Applications (SIAM). 2017: 495-506. https:///doi.org/10.1137/1.9781611974683.ch37
3. Rios LM, Sahinidis NV. Derivative-free optimization: a review of algorithms and comparison of software implementations. Journal of Global Optimization. 2013;*56:* 1247-1293. https://doi.org/10.1007/s10898-012-9951-y
4. Larson J, Menickelly M, Wild SM. Derivative-free Optimization Methods. Acta Numerica. 2019;28*:* 287-404. https://doi.org/10.1017/S0962492919000060
5. Amaran S et al. Simulation Optimization: A Review of Algorithms and Applications. Ann Oper Res. 2016;240: 351–380. https://doi.org/10.1007/s10479-015-2019-x
6. Ammeri A, Hachicha W, Chabchoub H, Masmoudi F. A comprehensive literature review of mono-objective simulation optimization methods. Advances in Production Engineering & Management. 2011;6(4): 291–302.
7. Walton S, Hassan O, Morgan K. Selected Engineering Applications of Gradient Free Optimisation Using Cuckoo Search and Proper Orthogonal Decomposition. Archives of Computational Methods in Engineering. 2013**;**20: 123-154. https://doi.org/10.1007/s11831-013-9083-7
8. Yang XS, Deb S. Engineering Optimisation by Cuckoo Search. International Journal of Mathematical Modelling and Numerical Optimisation. 2010**;**1(4): 330–343. https://doi.org/10.48550/arXiv.1005.2908
9. Xing XQ, Damodaran M. Assessment of Simultaneous Perturbation Stochastic Approximation Method for Wing Design Optimization. Journal of Aircraft. 2002;39: 379–381. https://doi.org/10.2514/2.2939
10. Xing XQ, Damodaran M. Application of Simultaneous Perturbation Stochastic Approximation Method for Aerodynamic Shape Design Optimization. AIAA Journal. 2005;43(2): 284–294. https://doi.org/10.2514/1.9484
11. Xing XQ, Damodaran M. Inverse Design of Transonic Airfoils Using Parallel Simultaneous Perturbation Stochastic Approximation. Journal of Aircraft. 2005;42(2): 568–570. http://dx.doi.org/10.2514/1.10876
12. Kothandaraman G, Rotea MA. Simultaneous-Perturbation-Stochastic-Approximation Algorithm for Parachute Parameter Estimation. Journal of Aircraft. 2005;42(5): 1229–1235. http://dx.doi.org/10.2514/1.11721
13. Prakash P et al. Design Optimization of a Robust Sleeve Antenna for Hepatic Microwave Ablation. Physics in Medicine and Biology. 2008;53: 1057–1069. https://doi.org/10.1088/0031-9155/53/4/016
14. Li Y. A Simulation-based Evolutionary Approach to LNA Circuit Design Optimization. Applied Mathematics and Computation. 2009; 209(1): 57–67. http://dx.doi.org/10.1016/j.amc.2008.06.015
15. Radac MB et al. Application of IFT and SPSA to Servo System Control. IEEE Transactions on Neural Networks. 2011;22(12): 2363–2375. https://doi.org/10.1109/tnn.2011.2173804
16. Ernst D et al. The Cross-Entropy Method for Power System Combinatorial Optimization Problems. Power Tech. IEEE. 2007: 1290–1295. https://doi.org/10.1109/PCT.2007.4538502
17. Kowalczyk Ł, Elsner W, Niegodajew P. The Application of Non-Gradient Optimization Methods to New Concept of Power Plant. 6th IC-EpsMsO; 2015 Jul 8-11; Athens.
18. Lu S. Dynamic modelling and simulation of power plant systems. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy. 1999;*213(1)*: 7-22. https://doi.org/10.1243/0957650991537392
19. Huan J et al. The Application of Digital Twin on Power Industry. IOP Conf. Series: Earth and Environmental Science. 2021;647. https://doi.org/10.1088/1755-1315/647/1/012015
20. Huyer W, Neumaier A. Global Optimization by Multilevel Coordinate Search. Journal of Global Optimization. 1999;14(2): 331-355. https://doi.org/10.1023/A:1008382309369
21. Huyer W, Neumaier A. SNOBFIT - Stable noisy optimization by branch and fit. ACM Transactions on Mathematical Software. 2008; *35(2):* Article No.: 9, 1-25. https://doi.org/10.1145/1377612.1377613
22. Knysh P, Korkolis Y. blackbox: A procedure for parallel optimization of expensive black-box functions. *arXiv (cs.MS)*. preprint submitted 2016, https://doi.org/10.48550/arXiv.1605.00998
23. Knysh P. blackbox: A Python module for parallel optimization of expensive black-box functions [Internet]. [place unknown]; [publisher unknown]; 2016 Feb 19 [updated 2022 Sep 5; cited 2021 Oct 17]. Available from: https://github.com/paulknysh/blackbox
24. Roberts M. Extreme Learning. The Unreasonable Effectiveness of Quasirandom Sequences [Internet]. [place unknown]; [publisher unknown]; 2018 Apr 25 [cited 2022 June 2]. Available from: http://extremelearning.com.au/unreasonable-effectiveness-of-quasirandom-sequences/
25. Regis RG, Shoemaker CA. Constrained Global Optimization of Expensive Black Box Functions Using Radial Basis Functions. Journal of Global Optimization. 2005**;***31*: 153-171. https://doi.org/10.1007/s10898-004-0570-0
26. Najko32. MCS algorithm [Internet]. [place unknown]; [publischer unknown]; 2018 August 6 [cited 2023 February 9]. Available from: https://commons.wikimedia.org/wiki/File:MCS_algorithm.gif
27. Nelder JA, Mead R. A Simplex Method for Function Minimization. The Computer Journal. 1965;7: 308-313. https://doi.org/10.1093/comjnl/7.4.308
28. Winfield D. Function Minimization by Interpolating in a Data Table. IMA Journal of Applied Mathematics. 1973;12: 339-347. https://doi.org/10.1093/imamat/12.3.339
29. Holland JH. Genetic Algorithms and Adaptation. Adaptive Control for Ill-Defined Systems. 1984: 317-333. https://doi.org/10.1007/978-1-4684-8941-5_21
30. Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and Analysis of Computer Experiments. Statistical Science. 1989;4: 409-423. https://doi.org/10.1214/ss/1177012413
31. Jung I et al. Computational Fluid Dynamics Based Optimal Design of Guiding Channel Geometry in U-Type Coolant Layer Manifold of Large-Scale Microchannel Fischer–Tropsch Reactor. *Ind. Eng. Chem. Res.* 2016;55: 505–515. https://doi.org/10.1021/acs.iecr.5b03313
32. Hemmat Esfe M, Hajmohammad M, Moradi R, Abbasian Arani AA. Multi-objective optimization of cost and thermal performance of double walled carbon nanotubes/water nanofluids by NSGA-II using response surface method. *Applied Thermal Engineering.* 2017;112: 1648–1657. https://doi.org/10.1016/j.applthermaleng.2016.10.129
33. Abdollahi A, Shams M. Optimization of heat transfer enhancement of nanofluid in a channel with winglet vortex generator. *Applied Thermal Engineering.* 2015;91: 1116–1126. https://doi.org/10.1016/j.applthermaleng.2015.08.066
34. Arora A, Bajaj I, Iyer SS, Hasan MMF. Optimal synthesis of periodic sorption enhanced reaction processes with application to hydrogen production. *Computers & Chemical Engineering*. (2018);*115:* 89–111.https://doi.org/10.1016/j.compchemeng.2018.04.004
35. Iyer SS, Bajaj I, Balasubramanian P, Hasan MMF. Integrated Carbon Capture and Conversion To Produce Syngas: Novel Process Design, Intensification, and Optimization. *Industrial & Engineering Chemistry Research.* (2017);*56*(30): 8622–8648.
36. https://doi.org/10.1021/acs.iecr.7b01688
37. Liu J, Ploskas N, Sahinidis NV. Tuning BARON using derivative-free optimization algorithms. *Journal of Global Optimization.* 2019;*74*(4): 611–637. https://doi.org/10.1007/s10898-018-0640-3

Lukas Peters: https://orcid.org/0009-0005-8195-8177

Rüdiger Kutzner: https://orcid.org/0009-0005-7171-024X

Marc Schäfer: https://orcid.org/0009-0009-2265-7023

Lutz Hofmann: https://orcid.org/0000-0002-3688-6136

DOI 10.2478/ama-2023-0035

ESTIMATION OF THE REGENERATIVE BRAKING PROCESS EFFICIENCY IN ELECTRIC VEHICLES

Jacek KROPIWNICKI*, Tomasz GAWŁAS**

*Faculty of Mechanical Engineering and Ocean Technology, Gdańsk University of Technology,
ul. Gabriela Narutowicza 11/12, 80-233, Gdańsk, Poland
**BMG Goworowski Gdynia, ul. Łużycka 9, 81-537, Gdynia, Poland

jkropiwn@pg.gda.pl, tomasz.gawlas@mazdagdynia.pl

received 2 September 2022, revised 3 February 2023, accepted 18 February 2023

Abstract: In electric and hybrid vehicles, it is possible to recover energy from the braking process and reuse it to drive the vehicle using the batteries installed on-board. In the conditions of city traffic, the energy dissipated in the braking process constitutes a very large share of the total resistance to vehicle motion. Efficient use of the energy from the braking process enables a significant reduction of fuel and electricity consumption for hybrid and electric vehicles, respectively. This document presents an original method used to estimate the efficiency of the regenerative braking process for real traffic conditions. In the method, the potential amount of energy available in the braking process was determined on the basis of recorded real traffic conditions of the analysed vehicle. The balance of energy entering and leaving the battery was determined using the on-board electric energy flow recorder. Based on the adopted model of the drive system, the efficiency of the regenerative braking process was determined. The paper presents the results of road tests of three electric vehicles, operated in the same traffic conditions, for whom the regenerative braking efficiency was determined in accordance with the proposed model. During the identification of the operating conditions of the vehicles, a global positioning system (GPS) measuring system supported by the original method of phenomenological signal correction was used to reduce the error of the measured vehicle's altitude. In the paper, the efficiency of the recuperation process was defined as the ratio of the accumulated energy to the energy available from the braking process and determined for the registered route of the tested vehicle. The obtained results allowed to determine the efficiency of the recuperation process for real traffic conditions. They show that the recuperation system efficiency achieves relatively low values for vehicle No. 1, just 21%, while the highest value was achieved for vehicle No. 3, 77%. Distribution of the results can be directly related to the power of electric motors and battery capacities of the analysed vehicles.

Key words: electric vehicle, urban conditions, regenerative braking, energetic efficiency

1. INTRODUCTION

Electricity delivered by the grid or range extenders [21] to an electric car is used by the propulsion system with various efficiencies, depending on the performance of the charging process [26], energy storage process in battery [28], electric engine control system, drive motor [10] and driveline system [23, 27]. It is estimated that the efficiency of all the above-mentioned processes is approximately 77%. Nowadays, electric drive systems of vehicles additionally enable the recovery of energy during braking and its reuse while driving the vehicle [11, 15]. Recently developed technical solutions that use recuperative braking energy allow the conversion of mechanical energy into electricity with high efficiency, even for very high power [7, 29, 31]. Control strategies for hybrid and electric vehicles in the field of regenerative braking energy are relatively complex [15, 17, 22] and depend strongly on the state of charge of the batteries. In modern electric cars, battery durability is a very important point of interest. This criterion is heavily impacted by the accurate estimation of the state of charge and state of health. In order to improve the accuracy of state of charge or state of health estimation, new researches are provided [24]. However, as a rule, manufacturers are trying to use the regenerative braking energy for charging batteries to the highest extent possible. Tests performed on a Toyota C-HR in urban conditions showed that the share of energy recovery from regenerative braking is over 50% of all energy supplied to the battery while driving [25].

It is necessary to consider the possibility of regenerative braking when planning the route of such vehicles, which is not implemented with the use of today's car navigation systems. A model of a regenerative braking process may be helpful for this purpose. On one hand, it will enable providing information on how traffic conditions affect the amount of energy dissipated (wasted) into the atmosphere, while on the other hand, it will help to optimize the route of vehicles with regenerative braking systems. A number of studies have been carried out to achieve this objective, while the methods used to predict electricity consumption are characterized by a very diverse level of complexity and accuracy [9, 19, 22].

Typically, the efficiency of a regenerative braking process is defined as a regenerative braking energy delivered to battery and back to a drive system, for an assumed driving cycle, divided by the energy achievable from braking process, which in conventional drive system is consumed by a hydraulic brake system. The achieved efficiency depends on the configuration of the drive system, applied control strategy and the driving cycle. For example, for a China Typical City Regenerative Driving Cycle (CTCRDC) driving cycle, one vehicle was tested and different strategies were applied; the calculated efficiency of a regenerative braking process was 31%–42% [25]. According to the results of simulations performed in work [15], for pure braking process, starting from the vehicle's initial speed of 80 km/h, the efficiency of a regenerative braking process for the tested control strategy was calculated at the level of 86%. Another study [16] shows that up to 50% of the total brake energy can be recycled in the urban driving cycle.

One of the most important factors directly affecting the aforementioned complexity and accuracy is the method used to identify car operating conditions [4, 10, 20]. Using the car operating conditions and model connecting the operating conditions with electricity consumption, it is possible to predict electricity consumption for the selected vehicle and route [5, 8, 9]. Due to the constantly increasing number of roads in city centres, in particular building new express roads within an agglomeration, drivers usually have more than one route to choose from. Their choice is most often determined by the travel time, but from an ecological point of view, more attention should be paid to the energy consumption or the emission of toxic compounds.

Currently, there are several platforms for the effective collection of data on the instantaneous traffic; however, as a rule, they determine the fastest or the shortest route. The development of a tool for determining the route due to the energy consumption is a primary need. This, in turn, requires the construction of a universal model, which enables the parametric description of traffic conditions to be linked with the amount of energy used to drive the vehicle [1, 2, 18]. This work focuses on the initial stage of building such a model. The paper presents a comprehensive method of identifying vehicle traffic conditions, and the method of estimation of the efficiency of the regenerative braking process using an on-board energy consumption measurement system for electric vehicles. Vehicle tests have been carried out in regular traffic for urban, rural and motorway conditions, following the operating conditions specified in the Real Driving Emissions (RDE) test [3, 6, 30]. Such a test represents the real operating conditions of vehicles as opposed to the previously used synthetic tests performed on a chassis dynamometer.

2. ENERGY CALCULATION OF THE REGENERATIVE BRAKING PROCESS

Vehicle operating conditions are identified in this work with the use of the specific energy consumption (SEC) that takes into account both an influence of external conditions and a driver's style of driving [13, 14]. The factors mentioned above affect the amount of mechanical energy transmitted to the drive wheels, which is one of the parameters constituting the SEC. Information on SEC for the considered road section can be directly used to calculate electricity consumption in the case of battery-powered vehicles. The value of the parameter for assumed cycle duration may be calculated using the following equation:

$$SEC = \frac{E}{m \cdot L} \tag{1}$$

where *SEC* is the specific energy consumption, *E* is the mechanical energy delivered by drive system to the wheels, *L* is the distance covered by the car and *m* is the gross vehicle mass.

Mechanical energy transmitted to the drive wheels can be calculated using the following equation:

$$E = \int_{t=0}^{t=t_c} \left(k_p \cdot F_t \cdot V\right) dt \tag{2}$$

where k_p is the positive traction force factor:

$$k_p = \begin{cases} 1 \ for\ powered\ wheels \\ 0 \ for\ idlling\ or\ braking \end{cases} \tag{3}$$

F_t is the traction force, calculated for recorded speed and altitude change,

$$F_t = m \cdot a \cdot \delta + m \cdot g \cdot sin(\alpha) + \rho_{air} \cdot A_f \cdot C_D \cdot \frac{V^2}{2} + m \cdot g \cdot C_r \cdot cos(\alpha) \tag{4}$$

where a is the vehicle acceleration, δ is the rotating mass factor, g is the acceleration due to gravity, α is the road grade, ρ_{air} is the air density, A_f is the vehicle frontal area, C_D is the vehicle aerodynamic drag coefficient, C_r is the vehicle rolling drag coefficient and *V* is the vehicle speed.

Alternatively, for the data recorded at the uniform time step, mechanical energy transmitted to the drive wheels may be calculated using the following equation:

$$E = \Delta t \cdot \sum_{i=1}^{N} \left(k_{p_i} \cdot F_{t_i} \cdot V_i\right) \tag{5}$$

where Δt is the time step.

As far as the usage of electric motor supporting conventional drive system is concerned, the regenerative braking energy must be taken into consideration. According to Eq. (4), the traction force is negative when the balance between acceleration component, road grade component, rolling resistance and air drag resistance is negative. This negative traction force in a conventional drive system is balanced by the brake system, while in hybrid or electric cars, it can be used for powering regenerative braking system. In the general case, however, the total energy that can potentially be delivered to the regenerative braking system can be calculated using the following equation:

$$E_{reg} = \Delta t \cdot \sum_{i=1}^{N} \left(k_{reg_i} \cdot F_{t_i} \cdot V_i\right) \tag{6}$$

where k_{reg} is the negative traction force factor:

$$k_{reg} = \begin{cases} -1 \ for\ idlling\ or\ braking \\ 0 \ \ for\ powered\ wheels \end{cases} \tag{7}$$

Regenerative braking specific energy (RBSE) for the covered distance can be calculated using the following equation:

$$RBSE = \frac{E_{reg}}{m \cdot L} \tag{8}$$

Power reception from the vehicle's drive system during braking is limited by specific restrictions and cannot be completed 100%. Typical restrictions come from the maximum power of the generator, its minimum operational speed or capacity of the battery. The part of energy that cannot be used for regenerative braking is transferred to the regular braking system. Distribution of this energy and efficiency of regenerative braking process are strongly influenced by the strategy applied by the system controller [17].

In electric cars, the electricity consumed by the drive system is measured at the input to the motor. Similarly, the electricity recovered from the braking process is measured at the generator output. The recorded result is the difference between the energy taken from the battery and that supplied to it from the braking process. Therefore, if we want to compare the results from the vehicle on-board system with the results of the traffic analyses, we should refer to the measuring points indicated above. Fig. 1 shows a diagram of the energy flow in the drive system of an electric vehicle with a regenerative braking system.

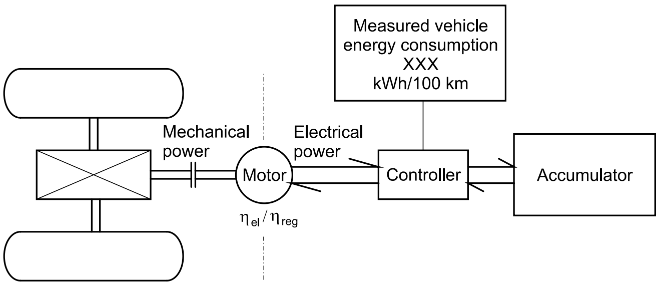


Fig. 1. Diagram of the energy flow in the drive system of an electric vehicle with a regenerative braking system

Calculation of the electric energy consumption for driving a vehicle, which corresponds to energy measured by the on-board system of the vehicle, can be performed using the following equation:

$$EEC = SEC \cdot m \cdot \frac{1}{\eta_{el}} - RBSE \cdot m \cdot \eta_{reg} \tag{9}$$

where η_{el} is the efficiency of electric drive system including electric motor and driveline and η_{reg} is the efficiency of regenerative braking system including electric generator and driveline.

The study did not consider the energy self-consumption of the vehicle, all tests were performed with the cooling or heating system turned off, and the energy consumption of other comfort systems (radio, displays, ventilation) was reduced to a minimum.

3. EFFICIENCY OF THE REGENERATIVE BRAKING PROCESS OF AN ELECTRIC VEHICLE

The analysis was carried out for three selected electric vehicles whose characteristics are presented in Tab. 1. The first two vehicles were designed for operation in typical urban conditions, and the third vehicle was equipped with a bigger battery, allowing its usage in extra-urban conditions also. The specified vehicle mass includes, in addition to the curb weight, the driver's weight. Due to the operating parameters, all vehicles can reach the speed of 120 km/h required in the test.

Tab. 1. Drive system parameters of the tested vehicles

No.	Vehicle	Mass [kg]	Power [kW]	Battery capacity [kWh]
1	Smart EQ	1175	60	17
2	Mazda MX-30	1725	107	35
3	Hyundai IoniQ 5	2100	224	72.6

To determine the efficiency of the electric drive and the regeneration process, a route taking into account typical urban conditions and bypass roads within the city has been selected in Gdańsk. The route is presented in Fig. 2 with the average speed (V) distribution. The road has been divided into 100-m-long sections, to which the average speed has been assigned. Fig. 3 shows the powered distance share (Lp/L) in indicated road sections. It has been assumed that the indicated road section can be covered by the vehicle in two states: powering and regenerative braking.

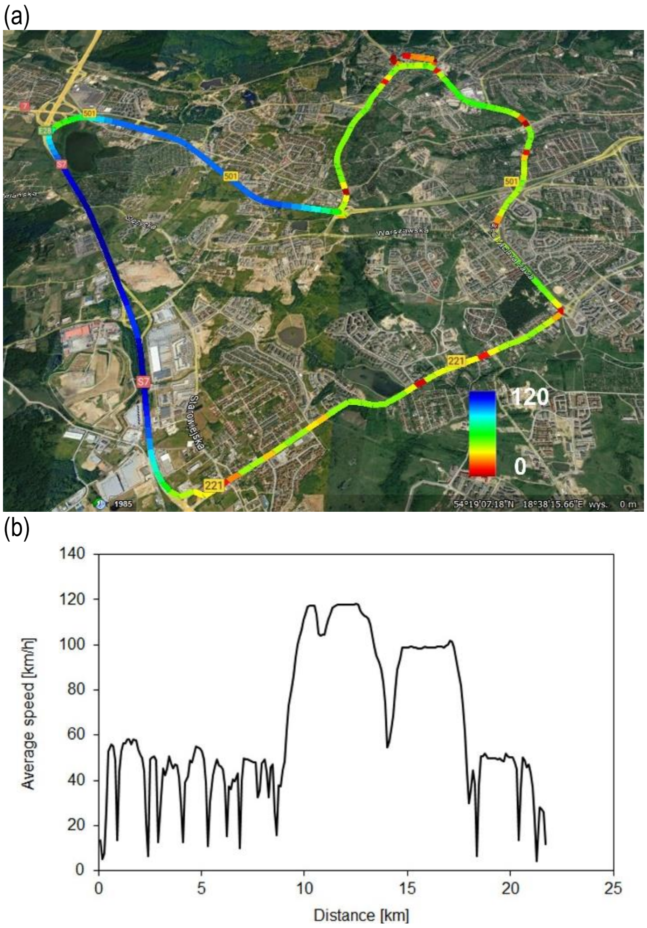


Fig. 2. The average speed (V) [km/h]: (a) distribution over the specified route and (b) distribution over the travelled distance

To determine the energy consumption and the energy that can potentially be used in regenerative braking process, the vehicle motion parameters (geographical coordinates, speed and altitude) have been determined on the basis of the recorded Global Positioning System (GPS) parameters. The traction force on the vehicle wheels has been calculated according to Eq. (4).

The road grade has been calculated using GPS coordinates. The altitude provided by the GPS receiver (Fig. 4) is not very accurate for terrestrial applications. The simple geometry of the system's satellite components means that the accuracy of the altitude measurement is about 10 times worse than the horizontal position measurement [12]. Typical vertical position error in signals for civilian use is about 5 m. However, the basic problem with the use of these receivers in road tests is the formation of reflections of the signal coming from satellite transmitters or its temporary disappearance. In urban conditions, this phenomenon is mainly caused by the proximity of buildings and the presence of tall buildings in the immediate vicinity of the road, as well as the presence of viaducts and tunnels.

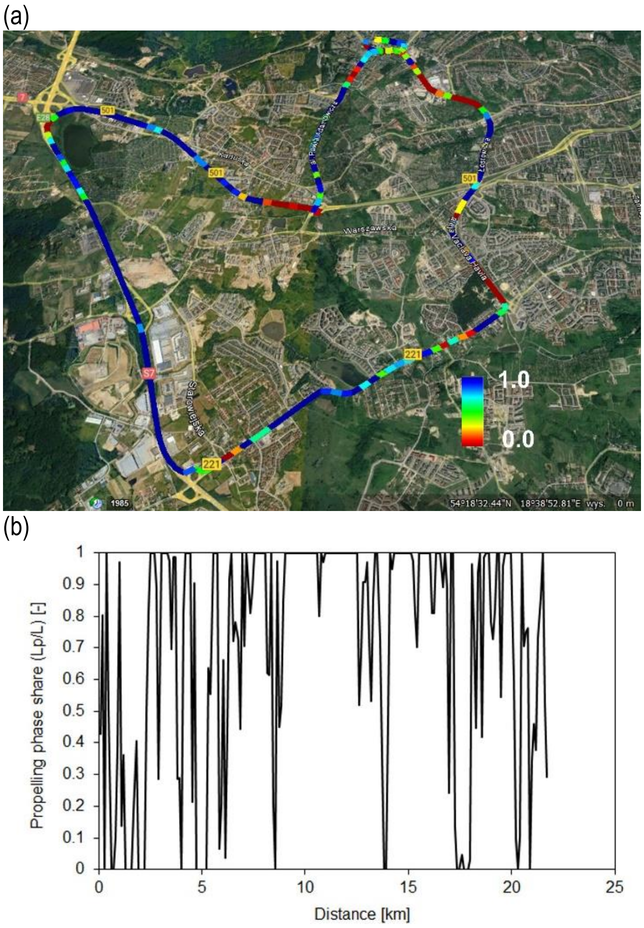


Fig. 3. The powered distance share (Lp/L) [-]: (a) distribution over the specified route and (b) distribution over the travelled distance

Fig. 4. GPS receiver used to determine the position of the vehicle and its altitude. GPS, Global Positioning System

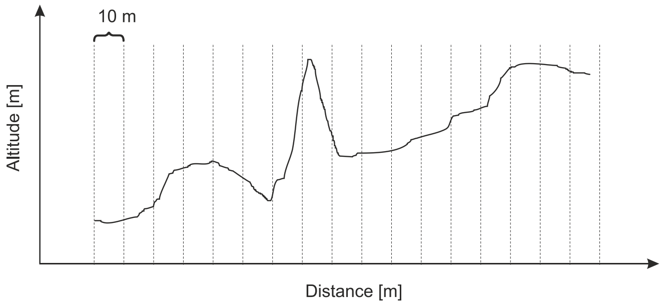


Fig. 5. Diagram of the division of the travel route in the horizontal direction into equal sections

Due to the above, the test has been carried out using a correction of the altitude signal based on phenomenological correction. The usage of the correction of the altitude signal depended on the exclusion of points giving greater road grade than allowed by the regulations. This method makes it possible to eliminate the influence of incorrect indications of the GPS regarding the measurement of altitude, as opposed to commonly used digital filtering methods, which only reduce this influence. The analysed route in the horizontal plane was divided into intervals of equal length of 10 m, and then the average value of the altitude in each interval indicated by the GPS was calculated (Fig. 5). The only exception is the last interval, the length of which results from the difference in the length of the entire route and values accumulated in the previous intervals.

Then, the first recorded altitude was connected with the average value in the middle of each section with straight lines. This way, a road profile was generated and mapped with straight sections with a constant horizontal component (except for the first and last sections). According to Polish regulations, the permissible absolute road grade for national roads should not exceed 3.7°. It was assumed that road grades smaller than 4° correspond to the actual course of the road profile and are marked with a bold line (Fig. 6), while the road grades which do not correspond to the given condition are marked with a dotted line. The condition of maintaining the permissible road grade can then be formulated as follows:

$$\alpha < \alpha_{perm} \tag{10}$$

where α is the road grade and α_{perm} is the permissible road grade, assumed to be 4°.

From the point where the incorrectly routed section begins, alternative road grade is starting, omitting the original altitude and the initial section is joined to the next one corresponding to the new road grade (Fig. 6).

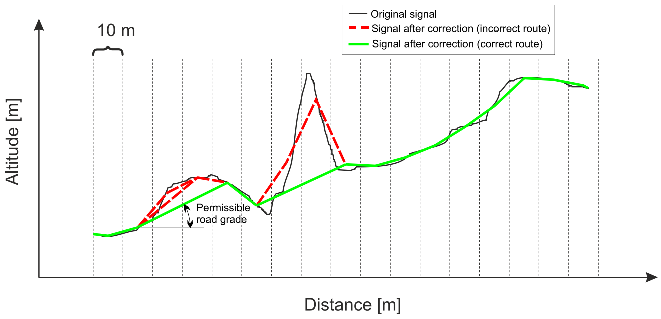

Original signal
Signal after correction (incorrect route)
Signal after correction (correct route)
10 m
Altitude [m]
Permissible road grade
Distance [m]

Fig. 6. Schematic drawing of phenomenological road profile correction: thin line – original route, dashed line – incorrect route, bold line – correct route meeting the condition (10)

In the event that the alternative road grade also fails to meet the condition of the permissible road grade (dotted line), the procedure is repeated. This procedure may also be performed in the opposite direction, if the first road grade, determined on the basis of the recorded original signal, does not meet the condition (10). In such a situation, the starting point for mapping the road profile can also be the first section fulfilling the condition (10), and the procedure of determining road grade will be carried out both in the forward and in the reverse directions. Fig. 7 shows an example of the road profile calculation with the use of phenomenological correction. Details of the mathematical basis of the method have been presented in Appendix A.

After calculating the road grade, speed and acceleration of the vehicle, it was possible to determine the power used to drive the vehicle and the power that can be used in the regenerative braking process. Fig. 8 presents the speed of the vehicle and the power calculated for vehicle No. 2.

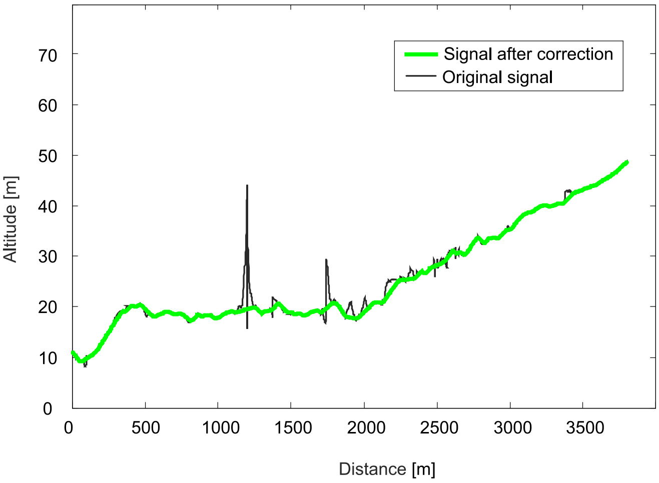


Fig. 7. Example of the road profile calculation with the use of phenomenological correction

Based on the relationships (1)–(8), SEC and RBSE have been determined. The calculations were performed for 100-m-long road sections. According to Eq. (1), high values of SEC correspond to a large amount of energy per distance travelled, which technically corresponds to intensive acceleration of the vehicle, climbing up a hill or driving at a very high speed. The results of this analysis have been presented in Figs. 9 and 10. Tab. 2 presents the parametric description of the route. This route was used in each analysed case.

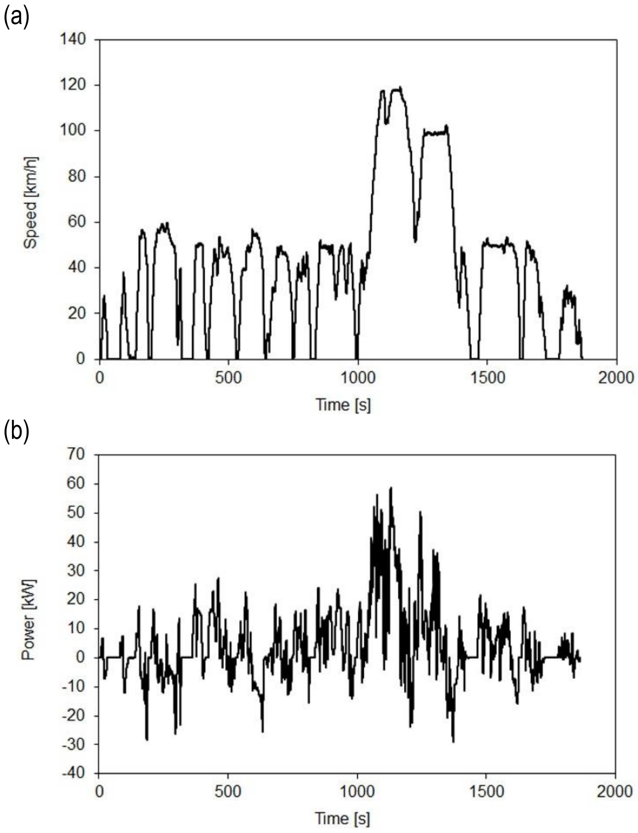


Fig. 8. The speed (a) of the vehicle and the power (b) calculated for vehicle No. 2

Tab. 2. Parametric description of the route used in tests (calculations processed for vehicle No. 2)

Distance	21.7 km
Average speed	41.9 km/h
SEC	9.11 J/(Mg*100 km)
RBSE	2.86 J/(Mg*100 km)
Idling time share	14.7%
Propelling phase share (Lp/L)	70.0%
Average power in propelling mode	10.3 kW
Average power in braking mode	5.1 kW

RBSE, Regenerative braking specific energy; SEC, specific energy consumption.

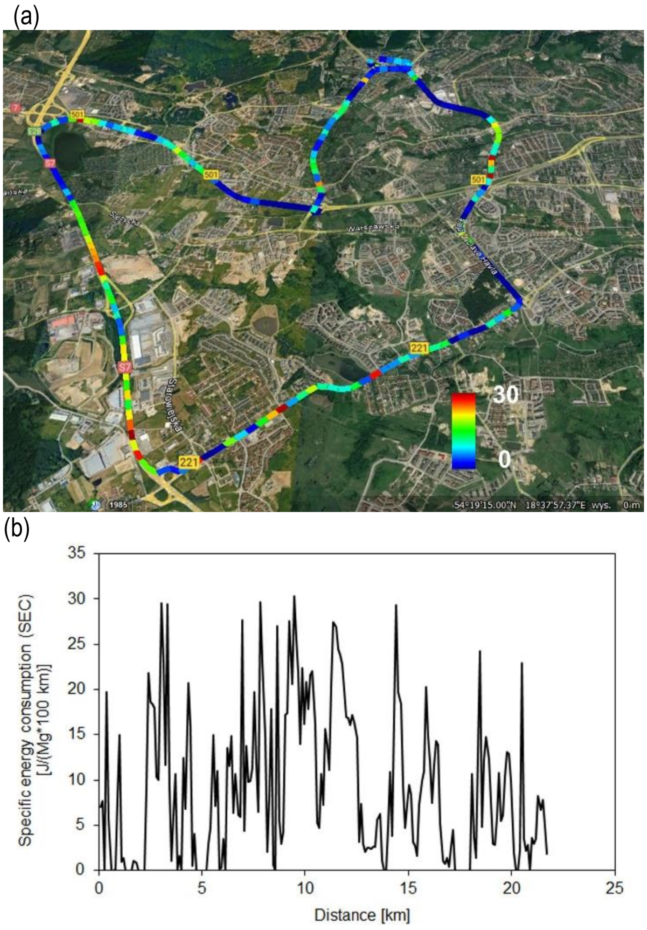

(a)
501
S7
221
30
0
54°19'15.00"N 18°37'57.37"E
(b)
Specific energy consumption (SEC) [J/(Mg*100 km)]
Distance [km]
0
5
10
15
20
25
30
35

Fig. 9. The SEC [J/(Mg*100 km)]: (a) distribution over the specified route and (b) distribution over the travelled distance. SEC, specific energy consumption

In the first stage of determining the efficiency of regenerative braking system, it was necessary to determine the efficiency of the electric drive system. For this purpose, the energy consumption of vehicle No. 2 with the regenerative braking system turned off has been measured. The obtained result is presented in Tab. 3. Due to the lack of technical possibilities to turn off the regenerative braking system in the other two vehicles, in further analysis, it was assumed that the drive system efficiency is the same in each of the tested vehicles.

DOI 10.2478/ama-2023-0035

Tab. 3. Energy consumption of vehicle No. 2 with the regenerative braking system turned off

Vehicle	Vehicle energy consumption [kWh/100 km]	Drive system efficiency [%]
Mazda MX-30	18.1	87

For the assumed route (Fig. 2), tests were conducted with selected cars. The electric energy consumption measurements in these tests, carried out with the use of the on-board systems, are presented in Tab. 4. The regenerative braking system was turned on in all the vehicles used in this test. In the tests, the control algorithms of the regenerative braking system were not influenced in any way. The strategies applied to control them were not the subject of the research and the original strategies implemented by vehicle manufacturers were used. Regenerative braking system efficiency was calculated using Eq. (9), where the SEC along with the RBSE was calculated for the registered route of the vehicle. The calculated regenerative braking system efficiency refers to the energy potentially available in a braking process, not to the total energy consumed by the drive system during operation.

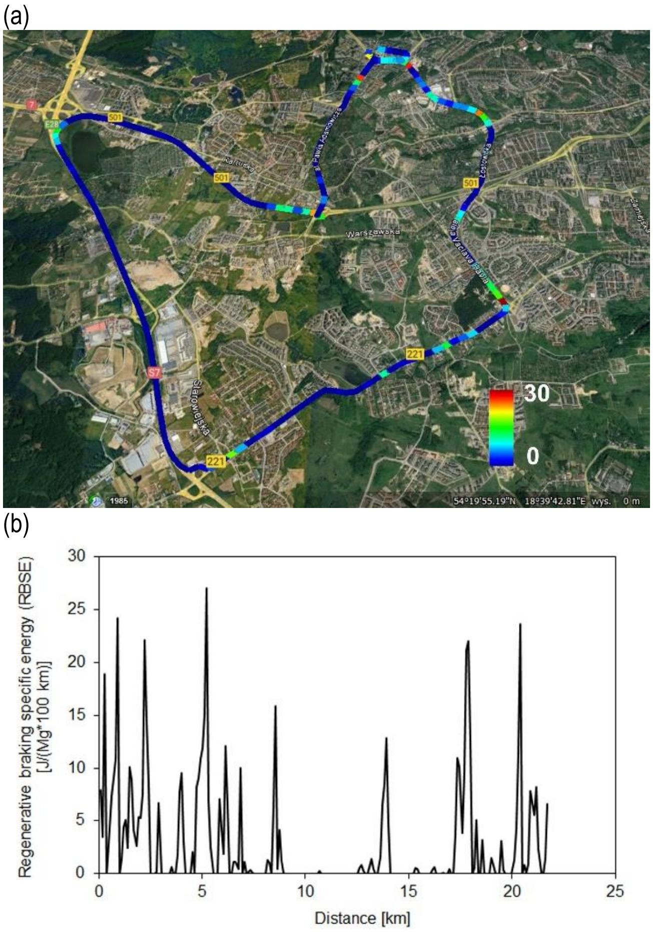


Fig. 10. The RBSE [J/(Mg*100 km)]: (a) distribution over the specified route and (b) distribution over the travelled distance. RBSE, regenerative braking specific energy

Tab. 4. Electric energy consumption of the tested vehicles with the regenerative braking system turned on

Vehicle	Vehicle energy consumption [kWh/100 km]	Drive system efficiency [%]	Regenerative braking system efficiency [%]
Smart EQ	13.6	87	21
Mazda MX-30	16.0	87	43
Hyundai IoniQ 5	15.5	87	77

4. CONCLUSIONS

The presented method of determining the efficiency of regenerative braking process requires relatively simple measuring devices, i.e. a GPS and an on-board energy consumption recording system. The procedure has been adapted for using the on-board energy consumption recording system of electric vehicles, which essentially simplifies the measurement procedure. Thus, the energy supplied from an external power source for charging the battery is not measured directly. Instead, the result of the electricity consumption balance recorded by the vehicle controller is taken into account.

Technically, the input and output energy can hardly be measured with sensors mounted on wheels, which enable taking into account the dual-direction energy flow considering the regenerative brake [25]. In this work, the applied calculation model was not verified due to the lack of technical possibility of interfering with the drive systems of the tested vehicles.

The obtained results enabled the formulation of the following conclusions:

- The adopted method of identifying vehicle operating conditions and its usage in the form of distribution of SEC (Fig. 9) and RBSE parameters (Fig. 10) over the specified route, for defined 100 m long sections, significantly facilitates the calculations of regenerative braking system efficiency.
- The applied original method consisting in using the on-board energy consumption recording system for measuring the electric energy consumption enabled a non-intrusive measurement of the parameters required to determine the regenerative braking system efficiency.
- A new phenomenological correction method was used to eliminate the influence of the incorrect indications of the GPS on the measured altitude.
- Road tests conducted for three electric vehicles enabled the measurement and comparison of the braking system efficiency (9) in real operating conditions.
- The obtained results show that the recuperation system efficiency achieves relatively low values for vehicle No. 1, merely 21%, which corresponds to the lowest electric motor power to mass relation (51 W/kg), as well as the lowest battery capacity to mass relation (14 Wh/kg). Vehicle No. 2 achieved 43% efficiency, which corresponds to the electric motor power to mass relation (62 W/kg) and the battery capacity to mass relation (20 Wh/kg). Vehicle No. 3 achieved the best efficiency amounting to 77%, which corresponds to the electric motor power to mass relation (107 W/kg) and the battery capacity to mass relation (35 Wh/kg). The factor that had a considerable impact on the results consisted of the wide range of altitude change, which has a particularly strong influence during long-term braking combined with going downhill, taking place at the end of the route section with a registered speed of 100 km/h.

NOMENCLATURE

a – vehicle acceleration,
A_f – vehicle frontal area,
C_D – vehicle aerodynamic drag coefficient,
C_r – vehicle rolling drag coefficient,
E – mechanical energy delivered by drive system to the wheels,
F_t – traction force,
g – acceleration due to gravity,
k_p – positive traction force factor,
k_{reg} – negative traction force factor,
L – distance covered by the car,
m – gross vehicle mass,
RBSE – regenerative braking specific energy,
SEC – specific energy consumption,
V – vehicle speed,
α – road grade,
α_{perm} – permissible road grade,
δ – rotating mass factor,
Δt – time step,
η_{el} – efficiency of electric drive system including electric motor and driveline,
η_{reg} – efficiency of regenerative braking system including electric generator and driveline,
ρ_{air} – air density,
CTCRDC – China Typical City Regenerative Driving Cycle
GPS – Global Positioning System,
RDE – Real Driving Emissions.

REFERENCES

1. Abdurazzokov U, Sattivaldiev B, Khikmatov R, Ziyaeva S. Method for assessing the energy efficiency of a vehicle taking into account the load under operating conditions. In: E3S Web of Conferences 264 (8). 2021.
2. Alves J, Baptista PC, Gonçalves GA, Duarte GO. Indirect methodologies to estimate energy use in vehicles: Application to battery electric vehicles. Energy Convers Manag. 2016;124:116–29.
3. Andrych-Zalewska M, Chłopek Z, Merkisz J, Pielecha J. Analysis of the operation states of internal combustion engine in the Real Driving Emissions test. Arch Transp. 2022;61(1): 71–88.
4. Berzi L, Delogu M, Pierini M. Development of driving cycles for electric vehicles in the context of the city of Florence. Transp Res Part D Transp Environ. 2016;47:299–322.
5. Brady J, O'Mahony M. Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas. Appl Energy. 2016;177:165–78.
6. Cieslik W, Zawartowski J, Fuc P. The Impact of the Drive Mode of a Hybrid Drive System on the Share of Electric Mode in the RDC Test. In: SAE Technical Papers. 2020. p. 1–8.
7. Damiani L, Repetto M, Prato AP. Improvement of powertrain efficiency through energy breakdown analysis. Appl Energy. 2014;121:252–63.
8. Fiori C, Ahn K, Rakha HA. Power-based electric vehicle energy consumption model: Model development and validation. Appl Energy. 2016;168:257–68.
9. Fiori C, Arcidiacono V, Fontaras G, Makridis M, Mattas K, Marzano V, et al. The effect of electrified mobility on the relationship between traffic conditions and energy consumption. Transp Res Part D Transp Environ. 2019; 275–90.
10. Galvin R. Energy consumption effects of speed and acceleration in electric vehicles: Laboratory case studies and implications for drivers and policymakers. Transp Res Part D Transp Environ. 2017;53:234–48.
11. Huang J, Qin D, Peng Z. Effect of energy-regenerative braking on electric vehicle battery thermal management and control method based on simulation investigation. Energy Convers Manag. 2015;105:1157–65.
12. Krasuski K, Ćwiklak J. Accuracy Analysis of Aircraft Position at Departure Phase Using DGPS Method. Acta Mech Autom. 2020;14(1):36–43.
13. Kropiwnicki J, Kneba Z, Ziółkowski M. Test for assessing the energy efficiency of vehicles with internal combustion engines. Int J Automot Technol. 2013;14(3).
14. Kropiwnicki J. A unified approach to the analysis of electric energy and fuel consumption of cars in city traffic. Energy. 2019;182.
15. Li L, Li X, Wang X, Song J, He K, Li C. Analysis of downshift's improvement to energy efficiency of an electric vehicle during regenerative braking. Appl Energy. 2016;176:125–37.
16. Li L, You S, Yang C, Yan B, Song J, Chen Z. Driving- behavior-aware stochastic model predictive control for plug-in hybrid electric buses. Appl Energy. 2016;162:868–79.
17. Liu W, Qi H, Liu X, Wang Y. Evaluation of regenerative braking based on single-pedal control for electric vehicles. Front Mech Eng. 2020;15(1):166–79.
18. Mamala J, Graba M, Bieniek A, Prażnowski K, Augustynowicz A, Śmieja M. Study of energy consumption of a hybrid vehicle in real-world conditions. Maint Reliab. 2021;23(4):636–45.
19. Mamala J, Śmieja Michałand Prażnowski K. Analysis of the total unit energy consumption of a car with a hybrid drive system in real operating conditions. Energies. 2021;14(3966).
20. Orofino, Luifi; Cilimingras, Luis; Morello E. Ecodrive: Driver Behaviour Evaluation System To Reduce CO2 Emissions. FISITA World Automot Congr. 2010;1(E3-1: Intelligent Transportation Systems):261–76.
21. Pielecha I, Cieślik W, Fluder K. Analysis of energy management strategies for hybrid electric vehicles in urban driving conditions. Combust Engines. 2018;173:14–8.
22. Pielecha I, Cieślik W, Szałek A. The use of electric drive in urban driving conditions using a hydrogen powered vehicle – Toyota Mirai. Combust Engines. 2018;172:51–8.
23. Pielecha I. Control algorithms for a Range Extender vehicle with an combustion engine. Combust Engines. 2020;183(4): 3–10.
24. Qian KF, Liu XT. Hybrid optimization strategy for lithium-ion battery's State of Charge/Health using joint of dual Kalman filter and Modified Sine-cosine Algorithm. J Energy Storage. 2021;44.
25. Qiu C, Wang G. New evaluation methodology of regenerative braking contribution to energy efficiency improvement of electric vehicles. Energy Convers Manag. 2016;119:389–98.
26. Racewicz S, Kazimierczuk P. Light Two-Wheeled Electric Vehicle Energy Balance Investigation Using Chassis Dynamometer. Acta Mech Autom. 2020;14(4):175–9.
27. Rakha HA, Ahn K, Moran K, Saerens B, den Bulck E Van. Virginia Tech Comprehensive Power-Based Fuel Consumption Model: Model development and testing. Transp Res Part D Transp Environ. 2011;16:492–503.
28. Rydh CJ, Sandén BA. Energy analysis of batteries in photovoltaic systems. Part I: Performance and energy requirements. Energy Convers Manag. 2005;46:1957–79.
29. Smith R, Shahidinejad S, Blair D, Bibeau EL. Characterization of urban commuter driving profiles to optimize battery size in light-duty plug-in electric vehicles. Transp Res Part D Transp Environ. 2011;16:218–24.
30. Triantafyllopoulos G, Kontses A, Tsokolis D, Ntziachristos L, Samaras Z. Potential of energy efficiency technologies in reducing vehicle consumption under type approval and real world conditions. Energy. 2017;140:365–73.
31. Zhang R, Yao E. Electric vehicles' energy consumption estimation with real driving condition data. Transp Res Part D Transp Environ. 2015;41:177–87.

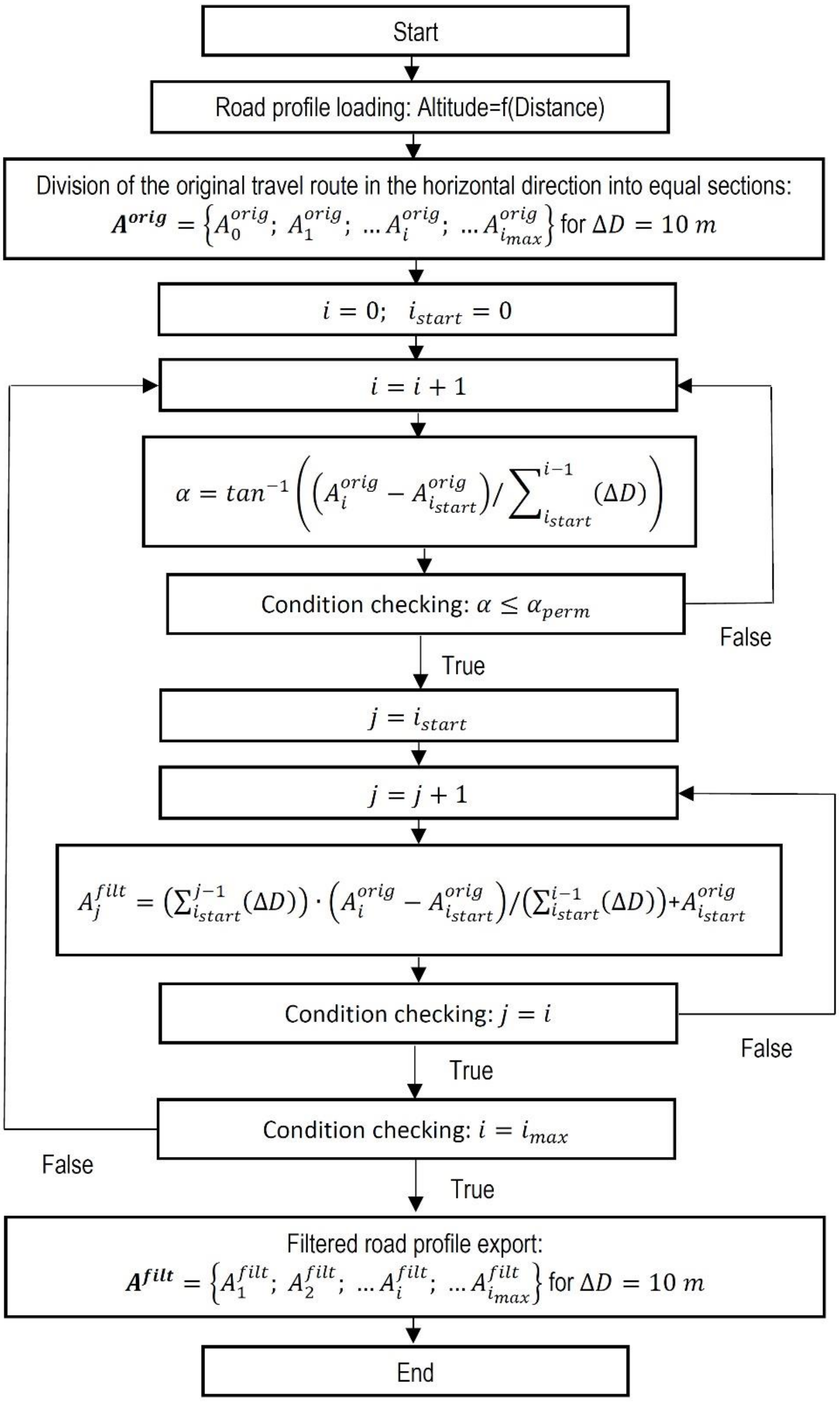
Jacek Kropiwnicki: https://orcid.org/0000-0001-7412-7424

Tomasz Gawłas: https://orcid.org/0009-0005-1699-6193

DOI 10.2478/ama-2023-0035

Appendix A

Mathematical basis of the phenomenological correction of the road profile


Start
Road profile loading: Altitude=f(Distance)
Division of the original travel route in the horizontal direction into equal sections: $\boldsymbol{A^{orig}} = \left\{A_0^{orig};\ A_1^{orig};\ \dots A_i^{orig};\ \dots A_{i_{max}}^{orig}\right\}$ for $\Delta D = 10\ m$
$i = 0;\quad i_{start} = 0$
$i = i + 1$
$\alpha = tan^{-1}\left(\left(A_i^{orig} - A_{i_{start}}^{orig}\right) / \sum_{i_{start}}^{i-1}(\Delta D)\right)$
Condition checking: $\alpha \le \alpha_{perm}$
False
True
$j = i_{start}$
$j = j + 1$
$A_j^{filt} = (\sum_{i_{start}}^{j-1}(\Delta D)) \cdot \left(A_i^{orig} - A_{i_{start}}^{orig}\right) / (\sum_{i_{start}}^{i-1}(\Delta D)) + A_{i_{start}}^{orig}$
Condition checking: $j = i$
False
True
Condition checking: $i = i_{max}$
False
True
Filtered road profile export: $\boldsymbol{A^{filt}} = \left\{A_1^{filt};\ A_2^{filt};\ \dots A_i^{filt};\ \dots A_{i_{max}}^{filt}\right\}$ for $\Delta D = 10\ m$
End