kOMUNIkÁCIE
COMMUNICATIONS

OPERATION AND ECONOMICS

MECHANICAL ENGINEERING

ELECTRICAL ENGINEERING

MANAGEMENT SCIENCE AND INFORMATICS

SAFETY AND SECURITY ENGINEERING

Janusz Cwiklak - Marek Grzegorzewski - Kamil Krasuski

INFLUENCE OF THE IONOSPHERIC DELAY ON DESIGNATION OF AN AIRCRAFT POSITION

The article presents and describes research results concerning determination of an impact of the ionospheric correction upon the positioning accuracy of an aircraft. The main objective of examinations was to verify three ionospheric models (the Klobuchar model, SBAS model and IONEX model) in determining aircraft coordinates. In the framework of the conducted simulations, the authors determined the aircraft coordinates by means of the SPP code method in the GPS system. The article presents a comparison of the determined aircraft coordinates in the SPP code method in relation to an accurate solution in the RTK-OTF differential technique. Based on the obtained results, it was found that e use of the SBAS and IONEX model is exploited to improve the positioning accuracy of an aircraft in relation to the Klobuchar model, from 20% to 72%, in the geocentric XYZ frame. The obtained findings of a simulation indicate that the ionospheric correction in the SBAS and IONEX models may be used to improve the performance of aircraft coordinates in air navigation.

Keywords: *ionospheric delay, GPS, aircraft position, Klobuchar model, SBAS model, IONEX model*

1 Introduction

Use of the GPS satellite technology in aviation leads to designation of numerous systematic errors in satellite measurements. In general, the systematic GPS errors can be divided in geometry errors of the satellite-receiver system, errors associated with an accurate transfer of time and atmospheric errors [1]. Among the atmospheric errors it is possible to differentiate the tropospheric delay error and the ionospheric delay error.

The previous scientific investigations of determination of the ionospheric delay in aviation are mostly concerned with:

- impact of the ionospheric scintillation upon availability parameters and reliability of the GPS positioning within the approach procedures LPV-200 [2],
- determination of the slant value of the ionospheric delay STEC in kinematic measurements in aircraft positioning in aviation [3],
- determination of the slant value of the ionospheric delay STEC in the GPS single-frequency differential measurements in aircraft positioning in aviation [4],
- developing a methodology of determining the ionospheric delay STEC in single frequency code-phase GPS measurements within the NPA procedure and the APV indirect procedure [5],
- determination of the ionospheric delay in the GPS measurements within the GBAS augmentation system in air transport [6-11].

The problem of modelling the ionospheric delay in aviation is a huge challenge for scientists who are primarily focused on air navigation. In Poland, in which the air traffic is steadily increasing, there is a need to develop accurate models of the ionospheric delay as well as specifying an influence of this phenomenon on the positioning of an aircraft. Within the research conducted in Poland for modelling of the ionospheric delay, the Klobuchar and SBAS models were used for the EGNOS system [12-14]. In the above-mentioned works, value of the ionospheric delay in execution of flight tests in eastern Poland was examined. Based on the presented research findings in articles [12-14], it was found that monitoring the state of the ionosphere for air operations is indispensable and necessary.

The aim of this article was to specify an impact of the ionospheric delay in order to determine the aircraft coordinates. In the work the authors used the Klobuchar model, SBAS EGNOS model and the global ionospheric model GIM in the IONEX format for checking the impact of the ionosphere delay on the GPS kinematic data. In particular, the article presents results of an impact of the ionospheric delay on accuracy of the GPS positioning in aviation. The computations exploited real observations and navigation data from an on-board GPS receiver mounted in a Cessna 172. The solution to the problem of modelling the ionospheric delay in the GPS kinematic measurements, presented in the article, is extremely interesting and may be used on a larger scale in air tests.

2 The research method

In order to determine the aircraft coordinates, the SPP code method was used as a primary and universal technique of the GPS positioning in

Janusz Cwiklak, Marek Grzegorzewski, Kamil Krasuski*
Institute of Navigation, Military University of Aviation, Deblin, Poland
*E-mail of corresponding author: k.krasuski@law.mil.pl

aviation. The mathematical model of the SPP code method is shown below [15]:

$$C1 = \rho + c \cdot (dtr - dts) + Ion + \\ + Trop + TGD + Rel + Mp \,, \quad (1)$$

where:
$C1$ - code observations on the L1 frequency in GPS system (expressed in metres),
c - speed of light (expressed in m/s),
ρ - geometric distance between the satellite and a receiver on the L1 frequency in GPS system (expressed in metres),

$\rho = \sqrt{(X_{rec} - X_{sat})^2 + (Y_{rec} - Y_{sat})^2 + (Z_{rec} - Z_{sat})^2}$,
$(X_{rec}, Y_{rec}, Z_{rec})$ - coordinates of the aircraft in XYZ geocentric coordinates,
$(X_{sat}, Y_{sat}, Z_{sat})$ - satellite coordinates in the GPS system,
dtr - receiver clock bias in the GPS system (expressed in seconds),
dts - satellite clock bias in the GPS system (expressed in seconds),
Ion - ionosphere delay in the GPS system (expressed in metres),
$Trop$ - troposphere delay in the GPS system (expressed in metres),
TGD - Time Group Delay in the GPS system (expressed in metres),
Rel - relativistic effect in the GPS system (expressed in metres),
Mp - multipath effect and measurement noise in the GPS system (expressed in metres).

In Equation (1), the ionospheric delay is referred to as a standard in aircraft positioning by means of the Klobuchar model, which reduces the effect of the ionospheric impact by merely 50-60% [16]. Therefore, it is necessary to make a precise application of the ionospheric model to improve the performance of aircraft coordinates. The ionosphere delay in Klobuchar model is described as:

$$Ion = \begin{cases} \left[5 \cdot 10^{-9} + \sum \alpha_n \cdot \phi_m^n \cdot \left(1 - \frac{X_I^2}{2} + \frac{X_I^4}{24}\right)\right] \cdot F, & |X_I| \le 1.57 \\ 5 \cdot 10^{-9} \cdot F, & |X_I| \le 1.57 \end{cases} \quad (2)$$

where:
α_n - coefficients of the ionosphere delay from Broadcast ephemeris,
f_m^n - geomagnetic latitude of ionosphere pierce point,
X_I - phase of the ionospheric delay,
F - the mapping function.

One solution to this problem is to exploit the SBAS model for the EGNOS system, which relies on a regular grid (GRID), 5° by 5° of one mesh unit [17]. Values of the ionospheric corrections in the SBAS model are retransmitted in real time from EGNOS satellites to the onboard GPS receiver in the universal EMS format. Efficiency of the SBAS ionospheric model is considerably higher than in the Klobuchar model. The ionosphere delay in the SBAS model is described as:

$$Ion = MF \cdot \sum w_n \cdot VTEC_n \,, \quad (3)$$

where:
MF - mapping function,
w_n - weight, function of distance between the GRID coordinates and current ionosphere pierce point coordinates,
$VTEC_n$ - Vertical TEC at GRID SBAS coordinates.

In precise geodetic measurements, the IONEX format is exploited. It contains data of the global ionospheric model GIM in the form of ionosphere maps 2.5° by 5° of one mesh unit [18]. Efficiency of this model in geodetic solutions remains at a level of 80-90%. Thus, translating the ionospheric IONEX format into air navigation may prove crucial in improving the designation of aircraft coordinates. The ionosphere delay in the IONEX model is described as below:

$$Ion = MF \cdot \sum w_n \cdot VTEC_n \,, \quad (4)$$

where:

k - coefficient of time rate, $k = \frac{dt}{dT} = \frac{T_i - t}{T_i - T_o}$,
t - current time epoch,
(T_i, T_0) - reference time epoch of VTEC GRID maps,
$VTEC_n$ - Vertical TEC at GRID maps from IONEX file.

Impact of ionosphere on the aircraft positioning is visible in conducting navigation in the horizontal LNAV plane. Therefore, the ionospheric delay exerts an impact on determination of the aircraft horizontal coordinates. It has a direct influence during the aircraft take-off and landing at an airport. Thus, the ionospheric delay is a key navigation parameter in aircraft positioning.

3 The research test

In the framework of the research test, authors made a simulation of impact of the ionospheric delay on designation of an aircraft position. The computations exploited real observations and the GPS navigation data from an on-board GNSS receiver mounted in the Cessna 172 [19]. The Cessna 172 executed a test flight around the military airport EPDE in Deblin. In the framework of the conducted simulation, authors determined the aircraft position, using different models of the ionospheric delay, i.e. the Klobuchar model, SBAS EGNOS model and the global model GIM in the IONEX format. The research used:

- the RINEX observation file with the GPS code measurements,
- the GPS navigation file,
- the file from the EGNOS correction in the EMS format,
- the IONEX file with ionosphere maps.

In the first step, the aircraft position was determined, using the Klobuchar model from the GPS navigation message. In the second case, the aircraft position was determined, using the SBAS EGNOS ionospheric model within the format of EMS corrections. In the third case,

the aircraft position was determined, using the global ionosphere model GIM as a part of the universal IONEX format. The computations of the aircraft position were made in the RTKLIB v.2.4.3 programme in the RTKPOST library. In the computing strategy in the RTKPOST library, the following were used [20]:

- positioning method: single,
- elevation mask: 5°,
- source of ionospheric adjustments: on-board ephemeris Klobuchar model, SBAS model, IONEX format,
- source of tropospheric correction: Saastamoinen model,
- source of ephemeris data and satellite clocks data: on-board ephemeris data,
- coordinate frame: XYZ geocentric coordinates,
- mean error of pseudorange a priori: 1 m,
- type of observations: code on L1 frequency,
- weighting: in the function of the elevation angle: applied,
- maximum DOP coefficient: 30,
- observation interval: 1 s.

4 Results and discussion

The designated aircraft coordinates in the SPP code solution for different ionospheric models were compared to a precise reference position in order to determine the accuracy of the GPS positioning in aviation. The reference position of the aircraft was determined using the RTK-OTF differential technique for accurate GPS phase observations [21]. The reference position was designated in the Trimble Total Control v.2.7 programme. The coordinates of the aircraft reference position were also specified at a 1-second interval in the geocentric XYZ frame.

Within the conducted investigations, the authors analysed the impact of the ionosphere model on accuracy of the GPS positioning in aviation. Therefore, the accuracy of the GPS positioning was determined, taking into account an impact of an individual ionospheric model on designating the aircraft coordinates. The accuracy of the GPS positioning was specified as [22]:

- -for the X component as an absolute value of formula:

$$dX = \begin{cases} \left| X_{SPP}^{Klobuchar} - X_{RTK-OTF} \right| \\ \left| X_{SPP}^{SBAS} - X_{RTK-OTF} \right| \\ \left| X_{SPP}^{IONEX} - X_{RTK-OTF} \right| \end{cases}, \quad (5)$$

where:
$X_{SPP}^{Klobuchar}$ - X coordinate of aircraft based on the SPP code solution with the Klobuchar model, see Equation (1),
X_{SPP}^{SBAS} - X coordinate of aircraft based on the SPP code solution with SBAS model, see Equation (1),
X_{SPP}^{IONEX} - X coordinate of aircraft based on the SPP code solution with the IONEX model, see Equation (1),
$X_{RTK-OTF}$ - X coordinate of aircraft based on the RTK-OTF differential solution, reference coordinate of aircraft,
dX - accuracy of aircraft positioning along the X axis.

- for the Y component as an absolute value of formula:

$$dY = \begin{cases} \left| Y_{SPP}^{Klobuchar} - Y_{RTK-OTF} \right| \\ \left| Y_{SPP}^{SBAS} - Y_{RTK-OTF} \right| \\ \left| Y_{SPP}^{IONEX} - Y_{RTK-OTF} \right| \end{cases}, \quad (6)$$

where:
$Y_{SPP}^{Klobuchar}$ - Y coordinate of aircraft based on the SPP code solution with the Klobuchar model, see Equation (1),
Y_{SPP}^{SBAS} - Y coordinate of aircraft based on the SPP code solution with SBAS model, see Equation (1),
Y_{SPP}^{IONEX} - Y coordinate of aircraft based on the SPP code solution with the IONEX model, see Equation (1),
$Y_{RTK-OTF}$ - Y coordinate of aircraft based on the RTK-OTF differential solution, reference coordinate of aircraft,
dY - accuracy of aircraft positioning along the Y axis.

- for the Z component as an absolute value of formula:

$$dZ = \begin{cases} \left| Z_{SPP}^{Klobuchar} - Z_{RTK-OTF} \right| \\ \left| Z_{SPP}^{SBAS} - Z_{RTK-OTF} \right| \\ \left| Z_{SPP}^{IONEX} - Z_{RTK-OTF} \right| \end{cases}, \quad (7)$$

where:
$Z_{SPP}^{Klobuchar}$ - Z coordinate of aircraft based on the SPP code solution with the Klobuchar model, see Equation (1),
Z_{SPP}^{SBAS} - Z coordinate of aircraft based on the SPP code solution with SBAS model, see Equation (1),
Z_{SPP}^{IONEX} - Z coordinate of aircraft based on the SPP code solution with the IONEX model, see Equation (1),
$Z_{RTK-OTF}$ - Z coordinate of aircraft based on the RTK-OTF differential solution, reference coordinate of aircraft,
dZ - accuracy of aircraft positioning along the Z axis.

Figure 1 shows accuracy of the GPS positioning along the X axis for different ionospheric models, based on Equation (5). In the case of using the Klobuchar model, the GPS positioning accuracy ranges from 0.3 m to 9.1 m, with the value of arithmetic mean being equal to 4.6 m. In the case of using the SBAS model, the GPS positioning accuracy is between 0.1 m and 11.1 m, with the value of arithmetic mean being equal to 4.0 m. However, in the case of using the global model in the IONEX format, the GPS positioning accuracy ranges from 0.1 m to 5.8 m, with the value of arithmetic mean being equal to 1.3 m. Based on the obtained research results it can be observed that exploiting the SBAS model increased the positioning accuracy by approximately 13% in relation to using the Klobuchar model in the SPP code method. In addition, using the IONEX model increased the positioning accuracy by approximately 72% in relation to the exploitation of the Klobuchar model in the SPP code method.

Figure 2 shows accuracy of the GPS positioning along the Y axis for different ionospheric models, based on Equation (6). In the case of using the Klobuchar model, the GPS positioning accuracy ranges from 0.1 m to 2.6 m, with the value of arithmetic mean being equal to 0.5 m. In the case of using the SBAS model, the GPS positioning accuracy is between 0.1 m and 2.1 m, with the value of arithmetic mean being equal to 0.4 m. However, in the case of using the global model in the IONEX format, the GPS positioning accuracy ranges from 0.1 m to 2.1 m, with the value of arithmetic mean being equal to 0.4 m. Based on the

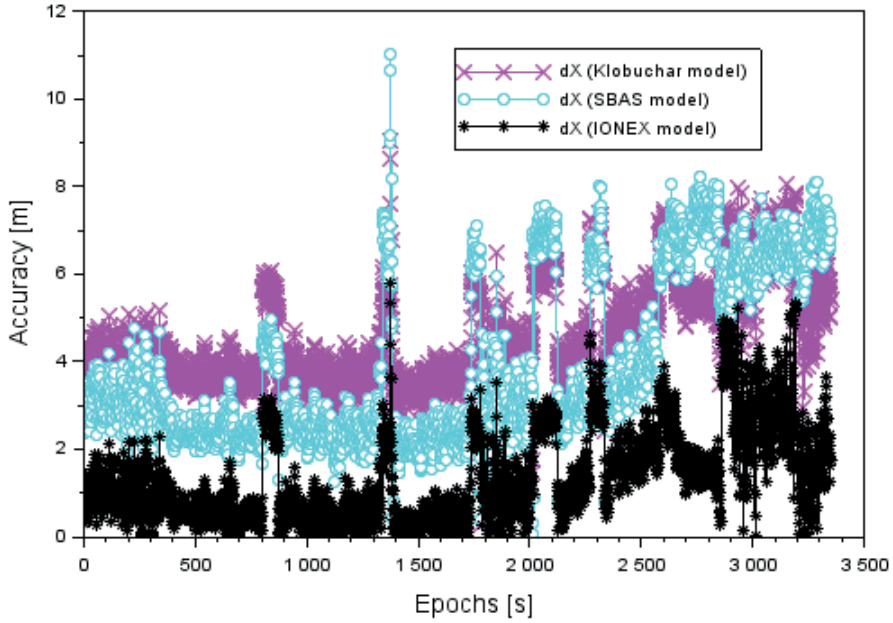


Figure 1 *Accuracy of aircraft positioning along the X axis*

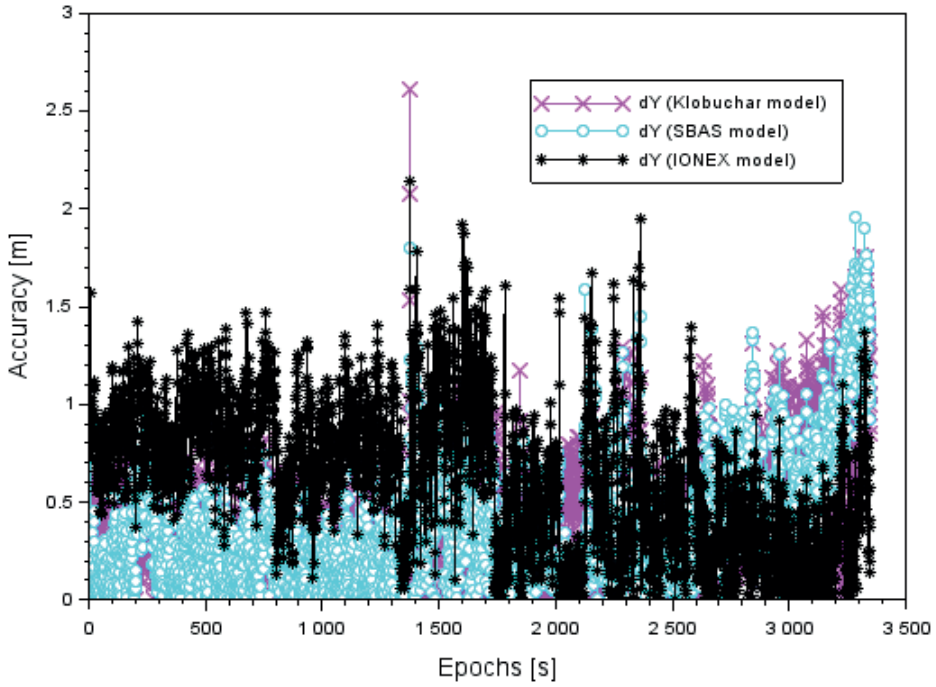


Figure 2 *Accuracy of aircraft positioning along the Y axis*

obtained research results it can be observed that exploiting the SBAS model increased the positioning accuracy by approximately 20% in relation to using the Klobuchar model in the SPP code method. In addition, using the IONEX model increased the positioning accuracy by approximately 20% in relation to the exploitation of the Klobuchar model in the SPP code method.

Figure 3 shows accuracy of the GPS positioning along the Z axis for different ionospheric models, based on Equation (7). In the case of using the Klobuchar model, the GPS positioning accuracy ranges from 0.3 m to 5.3 m, with the value of arithmetic mean being equal to 3.1 m. In the case of using SBAS model, the GPS positioning accuracy is between 0.1 m and 5.8 m, with the value of arithmetic mean being equal to 1.6 m. However, in the case of using the global model in the IONEX format, the GPS positioning accuracy ranges from 0.1 m to +5.7 m, with the value of arithmetic mean being equal to 2.2 m. Based on the obtained research results it can be observed that exploiting the SBAS model increased the positioning accuracy by approximately 47% in relation to using the Klobuchar model in the SPP code method. In addition, using the IONEX model increased

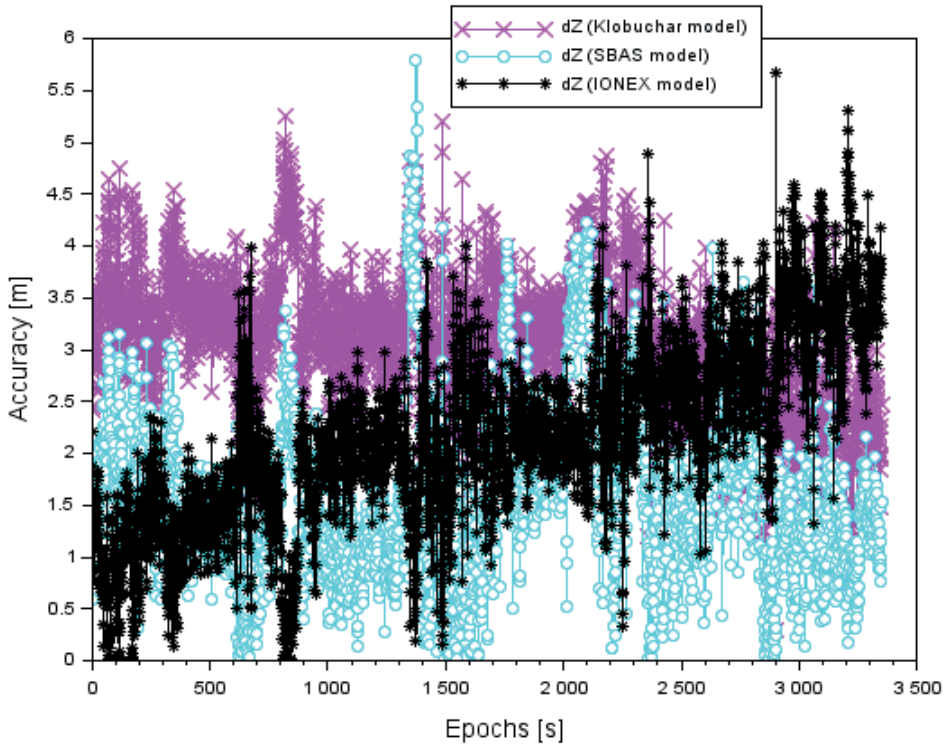


Figure 3 *Accuracy of aircraft positioning along the Z axis*

Table 1 *RMS error for the aircraft position in the XYZ geocentric coordinates*

	Klobuchar model	SBAS model	IONEX model
Coordinate X [m]	4.8	4.3	1.7
Coordinate Y (m)	0.5	0.5	0.7
Coordinate Z [m]	3.1	1.8	2.4

the positioning accuracy by approximately 29% in relation to the exploitation of the Klobuchar model in the SPP code method.

After specifying the accuracy, authors determined the RMS errors of the aircraft position for the obtained results of parameters (dX, dY, dZ) as shown below [23]:

$$\begin{cases} RMSdX = \sqrt{\dfrac{[dX^2]}{N}} \\ RMSdY = \sqrt{\dfrac{[dY^2]}{N}} \\ RMSdZ = \sqrt{\dfrac{[dZ^2]}{N}} \end{cases}, \tag{8}$$

where:
RMSdX - total RMS error along the X axis on the flight path,
RMSdY - total RMS error along the Y axis on the flight path,
RMSdZ - total RMS error along the Z axis on the flight path,
N - number of determinations of aircraft position.

Results of the RMS errors in the geocentric XYZ frame are shown in Table 1. For the application of the Klobuchar model, the RMS errors of the aircraft position range from 0.5 m to 4.8 m. For SBAS model, the RMS errors of the aircraft position range from 0.5 m to 4.3 m. For the IONEX model, the RMS errors of the aircraft position range from 0.7 m to 2.4 m. Based on Table 1, the RMS parameter is the smallest for e application of the IONEX model. In addition, the RMS parameter is the highest for the Klobuchar model.

In the next stage of the research, authors defined a shift in the aircraft position in the 2D plane and 3D space, as illustrated below [24-25]:

$$\begin{cases} 2D - error = \sqrt{dX^2 + dY^2} \\ 3D - error = \sqrt{dX^2 + dY^2 + dZ^2} \end{cases}, \tag{9}$$

where:
2*D* - *error* - shift in aircraft position in the horizontal 2D plane on the flight path,
3*D* - *error*- shift in aircraft position in 3D space on the flight path.

Figure 4 shows a shift in the aircraft position in the horizontal 2D plane. Therefore, the value of the parameter 2D-error was specified based on the obtained findings (dX, dY, dZ). For application of the Klobuchar model, the shift in the aircraft position in the horizontal 2D plane ranges from 0.7 m to 9.4 m. Then, for the application of the SBAS model, the shift in the aircraft position in the horizontal 2D plane ranges from 0.5 m to 11.2 m. In addition, for the application of the IONEX model, the shift in the aircraft position in the horizontal 2D plane is between 0.1 m and 6.2 m. Based on values of the parameter 2D-error it can be concluded that

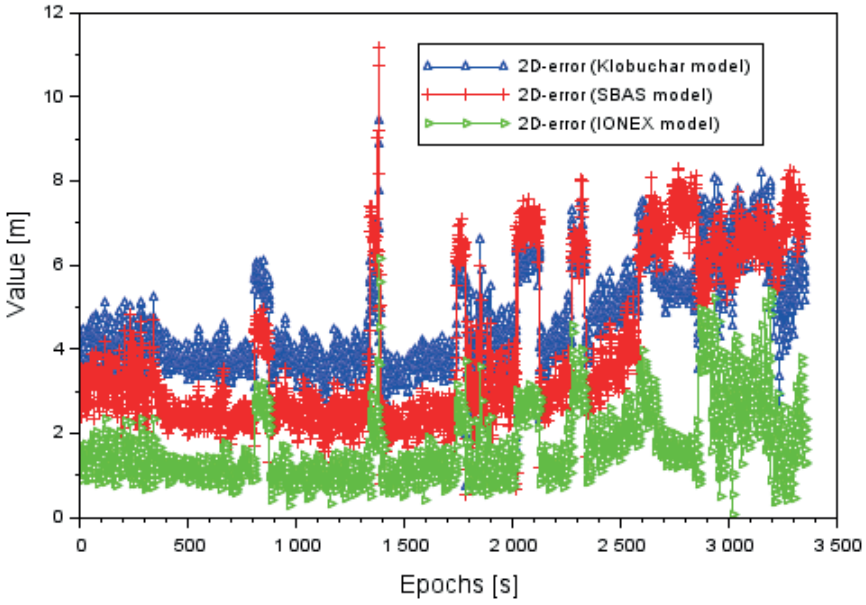


Figure 4 *Displacement of the aircraft position in the 2D horizontal plane*

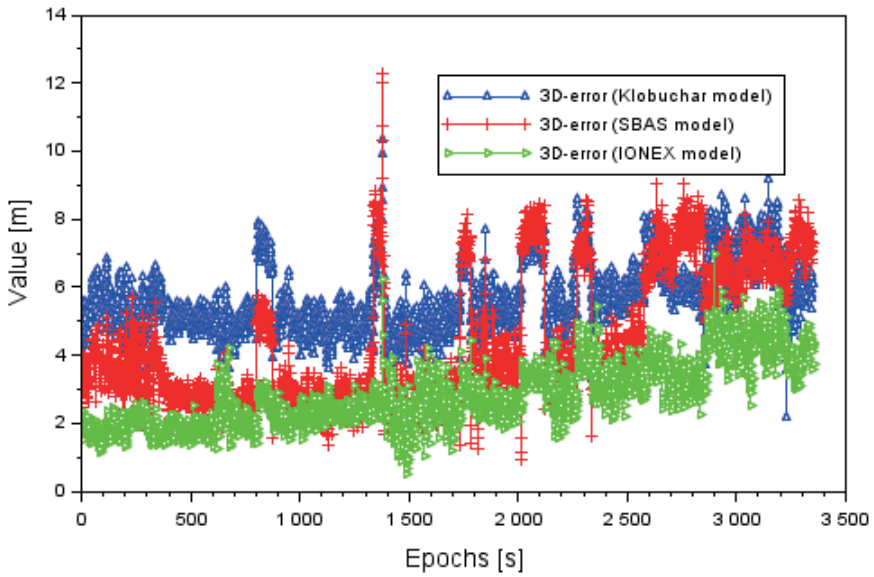


Figure 5 *Displacement of the aircraft position in the 3D space*

the smallest dispersion of results is for the application of the IONEX model and the largest for the SBAS model.

Figure 5 shows a shift in the aircraft position in the 3D space. Therefore, the value of the parameter 3D-error was specified based on the obtained findings (dX, dY, dZ). For the application of the Klobuchar model, the shift in the aircraft position in the 3D space ranges from 2.2 m to 10.4 m. Then, for application of the SBAS model, the shift in the aircraft position in the 3D space ranges from 0.9 m to 12.3 m. In addition for application of the IONEX model, the shift in the aircraft position in 3D plane is between 0.5 m and 7.0 m. Based on values of the parameter 3D-error it can be concluded that the smallest dispersion of results is for application of the IONEX model and the largest for the SBAS model.

5 Conclusions

The article describes results of research concerning specification of an impact of the ionospheric delay on determining the aircraft coordinates. The computations exploited the real observation and navigation GPS data from an on-board GNSS receiver. In the course of the conducted research, a number of computer simulations were made with regard to determination of the impact of the ionospheric correction in the SPP code solution. In order to determine the position of the aircraft, the SPP code method in the software RTKLIB v.2.4.3 in the RTKPOST library was used. In the SPP code method, three different ionosphere models were used, i.e. the Klobuchar model from the

on-board GPS message, the SBAS model from the EMS format in the EGNOS system and the global ionospheric model GIM in the IONEX format. The designated aircraft position using different ionosphere models in the SPP code method was compared to an accurate reference position determined in the RTK-OTF differential technique. Based on the conducted investigations, it was found that:

- application of the SBAS and IONEX models increases the accuracy of the SPP code positioning accordingly by 72% for coordinate X, 20% for coordinate Y and 47% for the coordinate Z in relation to the Klobuchar method,
- application of the SBAS and IONEX models causes the smallest shift in aircraft coordinates in relation to the accurate reference position,
- the SBAS and IONEX model can be used to determine the ionospheric delay in air navigation,
- the SBAS and IONEX models can be used to reduce the ionosphere delay and improve the aircraft position in air navigation.

References

[1] HOFMANN-WELLENHOF, B., LICHTENEGGER, H., WASLE, E. *GNSS - global navigation satellite systems: GPS, GLONASS, Galileo and more*. Wien, Austria: Springer Wien New York, 2008. ISBN 978-3-211-73012-6.

[2] SEO, J. *Overcoming ionospheric scintillation for worldwide GPS aviation*. PhD thesis. Stanford, CA: Stanford University, 2010.

[3] COLOMBO, O. L, HERNANDEZ-PAJARES, M., JANSSEN, V., GARCIA-RIGO, A. Comprehensive ionospheric delay modelling for single-frequency GNSS precise positioning: first steps. In: IGS Workshop : proceedings. 2016.

[4] WEBSTER, I. R. *Regional model for the prediction of ionospheric delay for single frequency users of the Global Positioning System*. Technical report no. 166. New Brunswick: Department of Geodesy and Geomatics Engineering, University of New Brunswick, 1993.

[5] OUZEAU, C, MACABIAU, C., ROTURIER, B., MABILLEAU, M., AZOULAI, L., LEVAN, J., BESSE, F. Ionospheric delay estimation in a single frequency mode for Civil Aviation. In: 21st International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2008 : proceedings. 2008. P. 1148-1158.

[6] VENKATA RATNAM, D. Estimation and analysis of user IPP delays using bilinear model for satellite-based augmented navigation systems. *Aviation* [online]. 2013, **17**(1), p. 65-69. ISSN 1648-7788, eISSN 1822-4180. Available from: https://doi.org/10.3846/16487788.2013.805864

[7] DATTA-BARUA, S., LEE, J., PULLEN, S., LUO, M., ENE, A., QIU, D-S., ZHANG, G., ENGE, P. Ionospheric threat parameterization for local area global-positioning-system-based aircraft landing systems. *Journal of Aircraft* [online]. 2010, **47**(4), p. 1141-1151. eISSN 1533-3868. Available from: https://doi.org/10.2514/1.46719

[8] HARRIS, I. L., MANNUCCI, A. J., IIJIMA, B. A., LINDQWISTER, U. J., MUNA, D., PI, X., WILSON, B. D. Ionospheric specification algorithms for precise GPS-based aircraft navigation. *Radio Science* [online]. 2001, **36**(2), p 287-298. eISSN 1944-799X. Available from: https://doi.org/10.1029/1999RS002428

[9] YOSHIHARA, T, FUJII, N., SAITO, A. A Study of the ionospheric effect on GBAS (Ground-Based Augmentation System) using the GPS nation-wide network date in Japan. In: National Technical Meeting of the Institute of Navigation : proceedings. 2004. p. 502-511.

[10] SAITO, S., SUNDA, S., LEE, J., PULLEN, S., SUPRIADI, S., YOSHIHARA, T., TERKILDSEN, M., LECAT, F. Ionospheric delay gradient model for GBAS in the Asia-Pacific Region. *GPS Solutions* [online]. 2017, **21**, p. 1937-1947. ISSN 1080-5370, eISSN 1521-1886. Available from: https://doi.org/10.1007/s10291-017-0662-1

[11] SAITO, S. Ionospheric effects on GBAS and mitigation techniques. In: Workshop on Ionospheric Data Collection, Analysis and Sharing this Support GNSS Implementation : proceedings. ENRI. 2011.

[12] KRASUSKI, K., WIERZBICKI, D. The impact of atmosphere delays in processing of aircraft's coordinates determination. *Journal of KONES* [online]. 2016, **23**(2), p. 209-214. ISSN 1231-4005. Available from: https://doi.org/10.5604/12314005.1213594

[13] GRUNWALD, G., CIECKO, A., BAKULA, M., KAZMIERCZAK, R. Examination of GPS/EGNOS integrity in north-eastern Poland. *EIT Radar, Sonar and Navigation* [online]. 2016, **10**(1), p. 114-121. ISSN 1751-8784, eISSN 1751-8792. Available from: https://doi.org/10.1049/iet-rsn.2015.0053

[14] GRZEGORZEWSKI, M., SWIATEK, A., OSZCZAK, S., CIECKO, A., CWIKLAK, J. Study of EGNOS safety of life service during the period of solar maximum activity. *Artificial Satellites* [online]. 2012, **47**, p. 137-145. eISSN 2083-6104. Available from: https://doi.org/10.2478/v10018-012-0019-5

[15] SANZ SUBIRANA, J., JUAN ZORNOZA J. M., HERNANDEZ-PAJARES M. Fundamentals and algorithms. Vol. 1. In: *GNSS data processing*. Noordwijk, Netherlands: ESTEC, ESA Communications, 2013. ISBN 978-92-9221-886-7, p. 139-144.

[16] XU, G., XU, J. *GPS theory, algorithms and applications* [online]. 3. Ed. Berlin Heidelberg: Springer Verlag, 2016. ISBN 978-3-662-50365-2. Available from: https://doi.org/10.1007/978-3-662-50367-6

[17] International Civil Aviation Organization Asia and Pacific Office. *SBAS safety assessment guidance related to anomalous ionospheric conditions*. Ed 1.0. 2016. Adopted by APANPIRG/27.

[18] SCHAER, S. *Mapping and predicting the Earth's ionosphere using Global Positioning System*. PhD thesis. Zurich, 1999.
[19] KRASUSKI, K. Aircraft positioning using SPP method in GPS system. *Aircraft Engineering and Aerospace Technology* [online]. 2018, **90**(8), p. 1213-1220. ISSN 0002-2667. Available from: https://doi.org/10.1108/AEAT-03-2017-0087
[20] RTKLIB ver. 2.4.3 Manual, RTKLIB: an open source program package for GNSS positioning - TAKASU T. [online]. 2013. Available from: http://www.rtklib.com/prog/manual_2.4.2.pdf
[21] CWIKLAK, J., JAFERNIK, H. The monitoring system for aircraft and vehicles of public order services based on GNSS. *Annual of Navigation*. 2010, **16**, p 15-24. ISSN 1640-8632.
[22] KRASUSKI, K., CWIKLAK, J., JAFERNIK, H., Aircraft positioning using PPP method in GLONASS system. *Aircraft Engineering and Aerospace Technology* [online]. 2018, **90**(9), p 1413-1420. ISSN 0002-2667. Available from: https://doi.org/10.1108/AEAT-06-2017-0147
[23] CHAI, T., DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* [online]. 2014, **7**, p. 1247-1250. ISSN 1991-959X, eISSN 1991-9603. Available from: https://doi.org/10.5194/gmd-7-1247-2014
[24] HENNING, W. *User guidelines for single base real time GNSS positioning***.** Version 2.1. National Geodetic Survey, 2011.
[25] JACOBSEN, K. S., DAHNN, M. Statistics of ionospheric disturbances and their correlation with GNSS positioning errors at high latitudes. *Journal of Space Weather and Space Climate* [online]. 2014, **4**, A27. eISSN 2115-7251. Available from: https://doi.org/10.1051/swsc/2014024

Annex - Nomenclature

The abbreviation	The full name
GPS	Global Positioning System
LPV	Localizer performance with vertical guidance
STEC	Slant TEC
NPA	Non-Precision Approach
APV	Approach with Vertical Guidance
GBAS	Ground Based Augmentation System
SBAS	Satellite Based Augmentation System
EGNOS	European Geostationary Navigation Overlay Service
GIM	Global Ionosphere Maps
IONEX	The IONospheric EXchange Format
SPP	Single Point Positioning
XYZ	Global geocentric coordinates
EMS	EGNOS Message
LNAV	Lateral Navigation
GNSS	Global Navigation Satellite System
EPDE	ICAO airport code
RINEX	Receiver Independent Exchange System
DOP	Dilution of Precision
RTK-OTF	Real Time Kinematic - On The Fly
RMS	Root Mean Square
VTEC	Vertical TEC

Kamil Krasuski - Stepan Savchuk

DETERMINATION OF THE PRECISE COORDINATES OF THE GPS REFERENCE STATION IN OF A GBAS SYSTEM IN THE AIR TRANSPORT

This paper presents results of research concerning determination of the GPS reference station coordinates located on the grounds of an EPDE airport in Deblin. The study uses a mathematical model of the PPP measurement technique in order to determine the coordinates of the reference station using the real GPS code-phase observations. The computations of the coordinates of the GPS reference station were carried out in numerical applications CSRS-PPP, APPS and GAPS. In this research was found that the accuracy of finding solutions to the XYZ geocentric coordinates of the reference station REF1 between solutions CSRS-PPP, APPS and GAPS ranges from 0.01 m to 0.13 m. In addition, the accuracy of determining the XYZ geocentric coordinates from the PPP method related to the GPS differential solution ranged from 0.01 m to 0.11 m.

Keywords: *GPS, reference station, PPP method, accuracy, XYZ geocentric coordinates*

1 Introduction

One of the elements of activating the GBAS augmentation system in air transport is the construction of a network of the RTK GPS permanent stations, located in the vicinity of both civil and military airports [1]. Ultimately, the GBAS augmentation system should consist of three components: the base station RTK GPS located at the airport, the mobile GPS receiver installed on board of an aircraft and transmission links of satellite data between the base station and the mobile receiver [2]. Such a configuration of components in the GBAS system ensures optimum utilization, during the precise positioning of an aircraft in the air transport. Within the GBAS augmentation system, it is possible to differentiate two basic methods of aircraft positioning: the DGPS differential technique and the RTK-OTF differential technique [3]. In the DGPS differential technique, the position of an aircraft is determined based on the GPS code observations, registered by the base station and the mobile receiver. On the other hand, in the RTK-OTF differential technique, the position of an aircraft is determined based on the GPS phase observations, registered by the base station and the mobile receiver [4].

The rules for the implementation and operation of the GBAS augmentation system in the air transport are clearly defined by the International Civil Aviation Organization ICAO [5]. In the framework of the ICAO guidelines and recommendations, the GBAS system finds its practical application in the procedure of a precision approach (PA) to landing. Among the types of a precision approach, it is possible to distinguish a precision approach, Category I, Category II and Category III. It is worth mentioning that the construction of the GBAS system infrastructure is costly and time consuming and it requires appropriate training for air traffic control personnel. In Poland, the GBAS system is currently being implemented for Cracow Balice Airport [6].

According to ICAO, the GPS measurements in the area of an airport should be realized with accuracy better than 0.1 m. In addition, in the carrier phase DGPS technique, typical accuracy of the GPS navigation system is about 0.01÷0.05 m, whereas, in static measurements in the post-processing mode, the GPS navigation system accuracy is around 0.01÷0.02 m [7]. The reference station coordinates in an airport must be estimated with the high level accuracy. In this way, the GPS reference station can be applied in conception of the GBAS technical infrastructure in the airport.

The aim of this article was to determine coordinates of the base station RTK GPS as one of the components of the GBAS augmentation system in the air transport. The precise coordinates of the RTK GPS reference station were determined using the PPP universal measurement technique for the GPS code-phase observations. The precise coordinates of the reference station REF1 were in calculations specified in the geocentric XYZ frame. A flight test was conducted for the reference station REF1, located within the military airport EPDE in Deblin.

2 Research method

The mathematical model of the GPS reference station coorrdinates determination is based on using the observation equations from the PPP measurement

Kamil Krasuski*, Stepan Savchuk
Institute of Navigation, Military University of Aviation, Deblin, Poland
*E-mail of corresponding author: k.krasuski@law.mil.pl

technique, as follows [8]:

$$\begin{cases} P_3 = \rho + c\cdot(dtr - dts) + Trop + \mathrm{R}el + M_{P3} \\ L_3 = \rho + c\cdot(dtr - dts) + Trop + Rel + B_3 + \delta_{wu} + M_{L3} \end{cases} \quad (1)$$

where:
$P_3 = \alpha_1 P1 + \alpha_2 P2$ - linear combination "Ionosphere-Free" for the GPS code measurements,
$L_3 = \alpha_1 L1 + \alpha_2 L2$ - linear combination "Ionosphere-Free" for the GPS phase measurements,
$(P1, P2)$ - GPS code measurements,
$(L1, L2)$ - GPS phase measurements,

$\alpha_1 = + \frac{f_1^2}{f_1^2 - f_2^2}$, linear coefficient,

$\alpha_2 = - \frac{f_2^2}{f_1^2 - f_2^2}$, linear coefficient,

(f_1, f_2) - carrier frequencies in the GPS system,
ρ - geometric distance between the GPS satellites and the receiver; it contains information about the parameters of the Earth's rotation, accurate coordinates of the satellite antenna and receiver, satellite antenna phase centre and receiver antenna phase centre, geodynamic and tidal effects, speed of the continental plate, etc.

$$\rho = \sqrt{(X - X_s)^2 + (Y - Y_s)^2 + (Z - Z_s)^2},$$

(X, Y, Z) - coordinates of the GPS reference station in the XYZ geocentric frame,
(X_S, Y_S, Z_S) - position of the GPS satellite in orbit,
c - speed of light,
dtr - receiver clock bias for the GPS observations,
dts - satellite clock bias for the GPS observations,
Trop- tropospheric delay for the GPS observations,
Trop = *SWD* + *SHD*,
SHD - Slant Hydrostatic Delay,
SWD - Slant Wet Delay,
$SHD = mf_H \cdot ZHD$,
$SWD = mf_W \cdot ZWD$,
(mf_H, mf_w) - mapping function for hydrostatic and wet delay,
ZHD - Zenith Hydrostatic Delay,
ZWD - Zenith Wet Delay,
R*el* - relativistic effects for the GPS observations,
δ_{wu} - phase wind up,
B_3 - real value of uncertainty phase,
M_{P3} - multipath effect for the GPS code measurements,
M_{L3} - multipath effect for GPS phase measurements.

The searched parameters in the PPP measurement technique are: coordinates of the receiver $[X, Y, Z]$ (3 parameters), receiver clock bias correction *dtr* (1 parameter), phase uncertainty principle B_3 (determined for each visible GPS satellite, from 1 to n, n-number of satellites), Zenith Wet Delay *ZWD* (1 parameter). The mentioned parameters are determined by the least squares method in the sequential process, as below [9]:

$$Sx = (A_{PPP}^T \cdot P_{PPP} \cdot A_{PPP} + C_X^{-1})\ A_{PPP}^T \cdot P_{PPP} \cdot l_{PPP}, \quad (2)$$

where:
Sx- vector of searched parameters,
A_{PPP} - matrix of plan,
P_{PPP} - matrix of weights,
C_x - variance-covariance matrix of determined parameters in the XYZ geocentric frame in the PPP measurement technique,
$C_x = (A_{PPP}^T \cdot P_{PPP} \cdot A_{PPP} + C_x^{-1})^{-1} + C_n$,
C_n - variance matrix of disturbances of the measurement process,
l_{PPP} - vector of constant terms.

In the stochastic process, in the PPP measurement technique, the determined parameters are modelled as [10]:
- coordinates of the receiver of the GPS reference station as a constant value in the stochastic model,
- receiver clock bias correction as a stochastic white noise model.
- *ZWD* parameter as a stochastic model of a random walk,
- phase uncertainty as a constant value in a stochastic model.

3 Research experiment

In the research experiment, the authors verified an application of the described research method in an accurate determination of coordinates of the reference station RTK GPS. In the analyzed example, coordinates of the reference station REF1 were determined, located at the military airfield in Deblin EPDE (see Figure 1). In test, the daily GPS observations data from REF1 GNSS station were applied, with time interval of 1 s. The experiment was realized on 1st June 2010.

The research experiment exploited the GPS navigation data recorded by the Topcon HiperPro receiver, mounted at the reference station REF1 in Deblin. In particular, in this research the P1/P2 and phase L1/L2 code measurements in the GPS system were used. In the calculations, there was the PPP precise measurement technique was used in order to determine coordinates of the reference station REF1 in the XYZ geocentric frame. Within the conducted studies, the numerical calculations were made in three independent geodetic programmes, using the PPP measurement technique in order to determine the position of the GPS receiver in a static mode. In the calculations, the available free geodetic programmes were used: CSRS-PPP, APPS and GAPS [12].

For the purposes of the conducted numerical calculations, the configuration of the CSRS-PPP programme was set, as below:
- GNSS system: GPS,
- linear combination: „Ionosphere-Free",
- type of positioning: PPP absolute method,
- positioning mode: static,
- computational mode: post-processing,

Figure 1 *Localization of the REF1 reference station [11]*

- type of GPS observation: dual-frequency, non-difference, code-phase observations,
- format of GPS observations: RINEX 2.xx,
- precise satellite ephemerides and satellite clocks: NRCan „Final"-type precise products,
- characteristics of the phase centre of the satellite/receiver antenna: based on the IGS ANTEX file,
- ultimate coordinate frame: ITRF2008 (reference epoch 2005.0),
- elevation mask: 10°,
- interval of calculations: 1 s,
- confidence level of the solution: 95%
- source of ephemeris and clock data: based on "SP3" and "CLK" formats
- weighting of measurement results: applied;
- gross error detection in the GPS measurements: applied,
- multipath effect: applied,
- relativistic effects: applied,
- Sagnac effect: applied
- correction of the GPS signal time from satellite to receiver: applied
- meteorological data for the tropospheric model: based on the GPT model
- model of troposphere: GPT (Global Pressure and Temperature) model,
- mapping function: GMF (Global Mapping Function),
- ionospheric delay: eliminated with the linear combination "Ionosphere-Free"
- accuracy of the GPS code observations a priori: 2 m,
- accuracy of the GPS phase observations a priori: 0.015 m,
- correction of receiver clock: determined,
- final recording of determined coordinates of the receiver: geocentric XYZ and ellipsoidal BLh coordinates.
- ZWD parameter: the hydrostatic delay is determined,
- uncertainty phase: determined for each tracked satellite,
- time of observation: GPS Time,
- hardware delay DCB: eliminated with the linear combination „Ionosphere-Free".

For the purposes of the conducted numerical calculations, the configuration of the APPS programme was set, as below:

- GNSS system: GPS,
- linear combination: „Ionosphere-Free",
- positioning mode: absolute method PPP,
- positioning mode: static,
- computational mode: post-processing,
- type of the GPS observations: dual-frequency, non-difference, code-phase observations,
- format of the GPS observations: RINEX 2.xx,
- precise satellite ephemerides and satellite clocks: JPL „Final"-type precise products,
- characteristics of the phase centre of the satellite/receiver antenna: based on the IGS ANTEX file,
- ultimate coordinate frame: ITRF2008 (reference epoch 2005.0),
- elevation mask: 7.5°,
- interval of calculations: 1 s,
- confidence level of solution: 95%
- final recording of determined coordinates of the receiver antenna: geocentric XYZ and ellipsoidal BLh coordinates,
- gross error detection with the GPS measurements: based on the programme TurboEdit,
- multipath effect: applied,
- relativistic effects: applied,
- Sagnac effect: applied
- the receiver clock bias: determined at an interval of 300 s,
- tropospheric delay: ZWD parameter (Zenith Wet Delay) and tropospheric gradients are determined at an interval of 300 s,
- uncertainty phase: determined for each tracked satellite,

- ionospheric delay: eliminated with the linear combination "Ionosphere-Free"
- time of observation: GPS Time,
- hardware delay DCB: eliminated with the linear combination „Ionosphere-Free".

For the purposes of the conducted numerical calculations, the configuration of the GAPS programme was set, as below:

- GNSS system: GPS system,
- File Type: RINEX: RINEX 2.xx,
- positioning mode: static,
- computational mode: post-processing,
- source of precise orbital data and GPS satellite clocks: IGS „Final" type precision ephemeris,
- intervals of precision orbital data and the GPS satellite clocks: 15 minutes,
- method of determining the satellite position and the GPS clocks from precision ephemeris: Lagrange polynomial model,
- used GPS observations: code observations (P1/P2) and phase observations (L1/L2)
- linear combination: Ionosphere-Free,
- weighting of the GPS observations: applied, in the function of the elevation angle,
- a priori standard deviation of code observations: 2 m,
- a priori standard deviation of phase observations: 1.5 cm,
- elevation cutoff angle (elevation mask): 10 degrees,
- interval of calculations: 1 s,
- maximum number of iterations in the stochastic process: 5,
- characteristics of the phase centre of the satellite/GPS receiver antenna: based on the IGS ANTEX file,
- initial values of aircraft coordinates: based on the RINEX file header,
- reference frame: global system ITRF2008,
- correction of receiver clock: eliminated,
- method of determination of the uncertainty phase: real Float value,
- final coordinates of the aircraft: geocentric coordinates (XYZ) and ellipsoidal coordinates (BLh)
- reference time: GPS Time,
- hardware delay DCB P1-P2: eliminated in the linear combination „Ionosphere-Free".
- ionospheric delay VTEC: first term of expansion of ionospheric delay, eliminated in the linear combination Ionosphere-Free,
- higher order ionospheric effects: not applied,
- satellite receiver orientation effect: applied,
- relativistic effect: applied
- movement of the pole correction, tidal effects, sea level and atmospheric pressure, movement of the continental plate, etc.: applied,
- Sagnac effect: applied
- model of troposphere: VMF1-gridded - VMFG_20100601,
- Zenith Hydrostatic Delay parameter: determined,
- value of the hydrostatic part of the tropospheric delay (ZHD) a priori: 2.332 m,
- mapping function of the tropospheric delay: Vienna mapping function.

4 Results and discussion

Findings of obtained results of the reference station REF1 coordinates, based on the CSRS-PPP, APPS and GAPS are shown in Tables 1, 2 and 3, respectively.

Table 1 shows results of the X coordinate for the reference station REF1. The scatter of results of the X coordinate, for the reference station REF1, in the solution CSRS-PPP, APPS and GAPS, equals approximately ±0.02÷0.03 m. Table 2 shows the results of the Y coordinate for the reference station REF1. The scatter of results for the Y coordinate, for the reference station REF1, in the solution CSRS-PPP, APPS and GAPS, also equals approximately ± 0.02÷0.13 m.

Table 3 shows results of the Z coordinate for the reference station REF1. The scatter of results for the Z coordinate, for the reference station REF1, in the solution CSRS-PPP, APPS and GAPS, also equals approximately ± 0.01÷0.06 m.

In Tables 4, 5 and 6 are presented the accuracy values in the form of standard deviations of determining the X, Y and Z coordinates of the reference station REF1, respectively. In Table 4 are shown results of the X coordinate mean values for the reference station REF1 in the CSRS-PPP, APPS and GAPS solution. The smallest value of the standard deviation for the X coordinate equals 0.01 m in the GAPS programme, whereas the largest one is equal to 0.03 m in the CSRS-PPP programme.

In Table 5 are presented results of the Y coordinate mean values for the reference station REF1 in the CSRS-PPP, APPS and GAPS solution. The smallest value of the standard deviation for the Y coordinate equals 0.02 m in the GAPS programme, whereas the largest one is equal to 0.07 m in the CSRS-PPP programme. In Table 6, there are results of mean values of the Z coordinate for the reference station REF1 in the CSRS-PPP, APPS and GAPS solution. The smallest value of the standard deviation for the Z coordinate equals 0.01 m in the GAPS programme, whereas the largest one is equal to 0.03 m in the CSRS-PPP programme.

Within the conducted experimental research, the authors verified the determined coordinates of the reference station REF1 in the PPP measurement technique. In the control test, the coordinate values of the reference station REF1, in the PPP measurement technique, were compared to the catalogue coordinates determined in the AUSPOS programme. In particular, the article compared the values of determined coordinates of the REF1 station in the geodetic XYZ frame. The catalogue coordinates of the REF1 reference station were determined in the AUSPOS programme. The programme AUSPOS ver. 2.2 is a free application tool to make computations of the GNSS receiver in a static mode, for the GPS phase observations [13]. The computational strategy for the determination of

Table 1 *Results of the X coordinate of the REF1 reference station*

Software	X coordinate [m]
CSRS-PPP	3687932.49
APPS	3687932.46
GAPS	3687932.47

Table 2 *Results of the Y coordinate of the REF1 reference station*

Software	Y coordinate [m]
CSRS-PPP	1480229.26
APPS	1480229.13
GAPS	1480229.24

Table 3 *Results of the Z coordinate of the REF1 reference station*

Software	Z coordinate [m]
CSRS-PPP	4972325.89
APPS	4972325.83
GAPS	4972325.88

Table 4 *Results of the X coordinate standard deviation of the REF1 reference station*

Software	Standard deviation of X coordinate [m]
CSRS-PPP	0.03
APPS	0.02
GAPS	0.01

Table 5 *Results of the Y coordinate standard deviation of the REF1 reference station*

Software	Standard deviation of Y coordinate [m]
CSRS-PPP	0.07
APPS	0.06
GAPS	0.02

Table 6 *Results of the Z coordinate standard deviation of the REF1 reference station*

Software	Standard deviation of Z coordinate [m]
CSRS-PPP	0.03
APPS	0.02
GAPS	0.01

the GNSS receiver coordinates in the AUSPOS programme is as follows:

- applied source software: Bernese GNSS Software Version 5.2,
- GNSS system: GPS system,
- preliminary GNSS data processing: elimination and repairing of cycle slips by means of triple difference phase technique for the phase observations L1/L2 in the GPS system; it is also possible to use various linear combinations to detect phase cycle slips,
- elevation mask: 7^0,
- interval of calculations: every 3 minutes, data cleaning every 30 seconds,
- weighting of the GPS observations: in the function of the elevation angle, $1/\sin^2(e)$, e- elevation angle,
- solution strategy: phase double-difference for the linear combination "Ionosphere-Free",
- characteristics of the phase centre of the satellite/receiver antenna: based on the ANTEX file,
- model of tropospheric delay: for the Zenith Hydrostatic Delay (ZHD), GMF model with DRY-GMF mapping function, tropospheric gradients G_N and G_E determined every 24 hours, Zenith Hydrostatic Delay determined every 2 hours with Wet-GMF mapping function,
- mapping function: GMF for tropospheric delay,
- ionospheric delay: the first term of expansion of ionospheric delay eliminated in the linear combination Ionosphere-Free, the 2nd and 3rd order ionospheric effects are used; moreover, it is possible to determine the global ionospheric maps to determine the uncertainty phase,
- tidal effects: on the recommendation of IERS 2010, effect of sea-level pressure is disregarded,
- atmospheric pressure: applied,
- source of orbital data: on the basis of IGS products,
- information about the parameters of the Earth's rotation: based on the IGS products,
- reference frame: ultimately ITRF2008,
- confidence level of solution: 95%,
- geoid model: EGM2008,
- method of determination of coordinates: method of least squares in the stochastic process,
- stochastic model of determining the GNSS receiver coordinates: modelling with boundary conditions imposed on the parameters, the change in horizontal deviation of horizontal coordinates equals 1 mm and for the vertical coordinate it is 2 mm.
- uncertainty phase solution: for vector's length of 180-6,000 km there is a strategy „Code-Based", for vector's length of 18-200 km there is a strategy „Phase-Based L5/L3", for vector's length of 18-2,000 km there is a strategy „Quasi-Ionosphere-Free (QIF)," for vector's length of 0-20 km there is a strategy „Direct L1/L2".
- final recording of determined coordinates of the receiver: geocentric XYZ and ellipsoidal BLh coordinates.

The catalogue coordinates of the reference station REF1 from the AUSPOS programme are as follows:

- for the X coordinate: 3687932.48 m with the standard deviation of 0.01 m,
- for the Y coordinate: 1480229.15 m with the standard deviation of 0.01 m,

Table 7 *The positioning accuracy of the REF1 reference station*

Parameter	CSRS-PPP vs AUSPOS	APPS vs AUSPOS	GAPS vs AUSPOS
dX [m]	0.01	-0.02	-0.01
dY [m]	0.11	-0.02	0.09
dZ [m]	0.02	-0.04	0.01

- for the Z coordinate: 4972325.87 m with the standard deviation of 0.02 m.

Owing to determination of the catalogue coordinates in the AUSPOS programme, it was possible to make a reliable evaluation of use of the PPP measurement technique in the static positioning. In particular, the verification process assesses the PPP positioning accuracy assessment in the GPS static mode. Based on that, it was possible to determine the difference in geocentric XYZ coordinates for the reference station REF1, as below [14]:

$$\begin{aligned} dX &= X_i - X_{AUSPOS} \\ dY &= Y_i - Y_{AUSPOS} \\ dZ &= Z_i - Z_{AUSPOS} \end{aligned}, \tag{3}$$

where:
(dX, dY, dZ) - parameters of accuracy positioning in the geocentric XYZ coordinates,
X_i - the X coordinate of the reference station REF1 from the solution of the CSRS-PPP, APPS and GAPS,
Y_i - the Y coordinate of the reference station REF1 from the solution of the CSRS-PPP, APPS and GAPS,
Z_i - the Z coordinate of the reference station REF1 from the solution of the CSRS-PPP, APPS and GAPS,
X_{AUSPOS} - the X catalogue coordinate of the reference station REF1 from the AUSPOS solution,
Y_{AUSPOS} - the Y catalogue of the reference station REF1, from the AUSPOS solution,
Z_{AUSPOS} - the Z catalogue coordinate of the reference station REF1 from the AUSPOS solution.

Table 7 shows results of the GPS accuracy positioning for the reference station REF1. Based on the obtained comparative results from Equation (3), it can be concluded that the highest positioning accuracy by means of the PPP measurement technique is visible along the X axis, being equal to approximately ±0.01÷0.02 m. On the other hand, the lowest accuracy of the GPS positioning is noticeable along the Y axis, particularly in the case of the CSRS-PPP and GAPS solution. Along the Z axis, the GPS positioning accuracy equals ± 0.01÷0.04 m.

5 Conclusions

The article presents results of determination of accurate GPS coordinates of the GPS reference station as one of the GBAS augmentation system elements in the air transport. The GPS reference station coordinates were determined using the PPP measurement technique for the GPS code-phase observations. In the research test, the authors determined the coordinates of the reference station REF1, localized on the grounds of the military airport EPDE in Deblin. Calculations of the reference station REF1 coordinates were performed in the CSRS-PPP, APPS and GAPS programmes. The paper presents results of determining the geocentric XYZ coordinates for the reference station REF1, in the PPP measurement technique. In addition, the article presents results of the geocentric XYZ coordinates standard deviations for the reference station REF1, in the PPP measurement technique. The article verifies the determined XYZ geocentric coordinates for the reference station REF1, in the PPP measurement technique against the catalogue coordinates, obtained in the AUSPOS programme. Based on the conducted examinations, it was found that the accuracy of the GPS positioning in the PPP measurement technique equals

- along the X axis up to ± 0.02m,
- along the Y axis up to ± 0.11m,
- along the Z axis up to ± 0.04my.

The research method presented in this paper can be fully exploited in the creation of the GBAS system for development of the air transport. Therefore, it is justified to use the PPP measurement technique for accurate determination of coordinates of the GPS reference stations, installed at airports. Determined coordinates of the GPS reference station in the PPP measurement technique seem to be reliable and accurate in relation to the catalogue coordinates, calculated using the dual-frequency differentiation of the phase observations. It must be stressed that the GBAS system in Poland is at the stage of construction, which means that its particular components must be properly verified so that their quality should be the best possible in the age of development of aviation. In the future, the authors plan to perform further research tests on use of the PPP measurement technique in determining the GPS reference stations located in the vicinity of civil airports in Mielec and Chelm, in south-eastern Poland.

Acknowledgement

The authots would like to thanks for CSRS-PPP, APPS, GAPS and AUSPOS on-line service for numerical computations of research. This paper was supported by Polish Air Force University for 2020 year.

References

[1] GRZEGORZEWSKI, M., CWIKLAK, J., JAFERNIK, H., FELLNER, A. GNSS for an Aviation. *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation*. 2008, **2**(4), p. 345-350. ISSN 2083-6473, eISSN 2083-6481.

[2] FELLNER, A. *Analysis of navigation systems and the concept of RTK DGPS permanent stations for the needs of aviation* (in Polish). Habilitation thesis. 1999. ISBN 83-912861-0-X, p. 11-14.

[3] GRZEGORZEWSKI, M., JARUSZEWSKI, W., FELLNER, A., OSZCZAK, S., WASILEWSKI, A., RZEPECKA, Z., KAPCIA, J., POPLAWSKI, T. Preliminary results of DGPS/DGLONASS aircraft positioning in flight approaches and landings. *Annual of Navigation*. 1999, **1**, p. 41-53. ISSN 1640-8632.

[4] KRASUSKI, K., CWIKLAK J., JAFERNIK, H. Verification of the precise position of the aircraft in air navigation based on the solution of the RTK-OTF technique. *Journal of KONES Powertrain and Transport* [online]. 2017, **24**(4), p. 117-124. ISSN 1231-4005, eISSN 2354-0133. Available from: https://doi.org/10.5604/01.3001.0010.3126

[5] International Civil Aviation Organization. ICAO standards and recommended practices (SARPS). Annex 10 volume I (Radio navigation aids) [online]. 2006. Available from: http://www.ulc.gov.pl/_download/prawo/prawo_miedzynarodowe/konwencje/zal10/Zal__10_Tom_I_popr__90.pdf

[6] Krakow Airport - Turystyka rp.pl [online]. Available from: https://www.rp.pl/Linie-lotnicze-i-lotniska/301259912-Krakow-Airport----najwiekszy-w-regionach-z-apetytem-na-wiecej.html, current on 2019

[7] International Civil Aviation Organization. *World Geodetic System - 1984 (WGS-84) manual*. 2. ed. 2002, DOC 9674 AN/946.

[8] SANZ SUBIRANA, J., JUAN ZORNOZA J. M., HERNANDEZ-PAJARES M. Fundamentals and algorithms. Vol. 1. In: *GNSS data processing*. Noordwijk, Netherlands: ESTEC, ESA Communications, 2013. ISBN 978-92-9221-886-7, p. 139-144.

[9] LEANDRO, R., SANTOS, M., LANGLEY, R. Analyzing GNSS data in precise point positioning software. *GPS Solutions* [online]. 2011, **15**, (1), p. 1-13. ISSN 1080-5370, eISSN 1521-1886. Available from: https://doi.org/10.1007/s10291-010-0173-9

[10] HADAS, T. GNSS-WARP software for real-time precise point positioning. *Artificial Satellites* [online]. 2015, **50**(2), p. 59-76. ISSN 2083-6104. Available from: https://doi.org/10.1515/arsa-2015-0005

[11] The localization of REF1 reference station – Google Maps [online]. https://www.google.pl/maps/place/51%C2%B033'20.0%22N+21%C2%B052'08.8%22E/@51.5562621,21.8661988,1503m/data=!3m1!1e3!4m5!3m4!1s0x0:0x0!8m2!3d51.555555!4d21.869117, current on 2019

[12] MALINOWSKI, M., KWIECIEN, J. A comparative study of precise point positioning (PPP) accuracy using online services. *Reports on Geodesy and Geoinformatics* [online]. 2016, **102**(1), p. 15-31. ISSN 2391-8152. Available from: http://dx.doi.org/10.1515/rgg-2016-0025

[13] CELLMER, S., RAPINSKI, J. Tests of selected automatic positioning systems in post-processing mode. *Technical Sciences*. 2011, **14**(1), p. 45-56. ISSN 2083-4527, eISSN 1505-4675.

[14] BAKULA, M. Static code DGPS positioning based on three reference stations. *Geodesy and Cartography*. 2005, **54**(2), p. 81-92. ISSN 1648-3502.

Annex - Nomenclature

The abbreviation	The full name
EPDE	ICAO airport code
GPS	Global Positioning System
PPP	Precise Point Positioning
CSRS-PPP	Canadian Spatial Reference System - Precise Point Positioning
GAPS	GPS Analysis and Positioning Software
APPS	Automatic Precise Positioning Service
GBAS	Ground Based Augmentation System
RTK	Real Time Kinematic
DGPS	Differential GPS
OTF	On The Fly
ICAO	International Civil Aviation Organization
PA	Precision Approach
RINEX	Receiver Independent Exchange System
ANTEX	Antenna Exchange Format
NRCan	Natural Resources Canada
ITRF	International Terrestrial Reference Frame
DCB	Differential Code Biases
VTEC	Vertical TEC
VMF	Vienna Mapping Functions
AUSPOS	The Australian Surveying and Land Information Group‘s Online GPS Processing Service
IERS	International Earth Rotation and Reference Systems Service
IGS	International GNSS Service
XYZ	Global geocentric coordinates
BLh	Latitude, Longitude, ellipsoidal height
EGM	The Earth Gravitational Model
GNSS	Global Navigation Satellite System
JPL	Jet Propulsion Laboratory

Tomasz Matyja

ASSESSMENT OF THE PALLET UNIT LOAD STABILITY BY SIMULATION METHODS

The paper presents a simulation model of a loading unit dynamics, with a layered structure, wrapped with a stretch film. Stretch film wrapping is the most commonly used and cheap method of securing the load unit. The model, proposed by the author, allows assessing the stability of the unit during the transport operations, when the loading unit is subjected to inertial forces. Deformations of a unit and prediction of its disintegration in extreme cases can be evaluated based on results of a simulation. The value of the necessary containment force as well as the number of wraps with the pre-stretched film, can also be estimated. In effect, simulations can reduce the amount of film used and the number of stability tests performed experimentally.

Keywords: *load unit, stability, stretch film*

1 Introduction

In the supply chain, a unit load is understood as a single mass that can be easily stored, mechanically stacked and moved between the two locations [1]. A load unit is usually made up of a series of smaller loads that are properly integrated and secured against disintegration. The pallet is the most commonly used platform for forming load units, which facilitates the transportation and storage of broadly understood material goods. Wrapping the unit load with the stretch film is the most frequently used method of protecting it against disintegration. The main advantage of the stretch film is its very low weight compared to the weight of the load and the fact that it is possible to apply the film both manually and mechanically using wrappers. The force bonding the whole load, called in the literature the containment force, is created by suitable pre-stretching the stretch film [2]. Of course, the more film layers the greater the force will be. However, the use of excessive numbers of layers unnecessarily increases packaging costs, contributes to redundant waste production and negatively affects the environment. Due to huge amounts of the stretch film used worldwide, it is important to apply the film as optimally as possible.

Pallet load units are classified into three classes due to the shapes of packages from which the unit is formed and their dimensional adjustment to the size of the pallet surface. Class A pallet units are made of identical packages, stacked in columns or bridges, which perfectly fill the surface of the pallet. The load on the A-class pallet creates a rectangular block, which is easy to effectively protect with stretch film. For Class B units, the individual load layers may differ in dimensions. The stretch film protection is still effective, but the film is more exposed to puncture. Class C pallet units are formed from packages of various shapes and are the least susceptible to being protected by foil [3].

During the transportation, even under the normal operating conditions, inertial forces act on the loading unit, what may cause displacement of individual packages, deformation of the unit treated as a solid and, in extreme cases, its disintegration. This problem occurs in all the modes of transport [4-6]. Stability, or also called rigidity, of the load unit is defined in literature and in transport safety standards as the integrity of the packages included in the unit and keeping unchanged shape [7]. Accidents involving trucks are very often caused by inadequate load securing [8-9]. Suitably rigid and properly secured against shifting in the vehicle's cargo space load unit improved human safety and protects products against damage. This problem has been noticed in the European Union, which has issued Directive 2014/47 [10].

The load unit stability is checked by experimental methods. Simple static tilt tests on an inclined plane or dynamic impact tests are carried out. Standard EUMOS 40509 recommends conducting dynamic tests [11]. Because real life truck braking tests are too expensive, they are carried out on special mobile platforms accelerated and braked. Typical laboratory stand is equipped with a fast camera for documenting the load's behavior and sensors measuring forces between packaging and film. The results of such tests depend not only on the stretch film performance and wrapping methods, but also on the weight and dimensions of the packages, friction coefficients and on packages arrangement pattern on the pallet [12]. Rationalizing film consumption, based on dynamic tests, requires at least several trials to be performed and is therefore time consuming.

The properties of the stretch film can vary widely and depend largely on the materials used and the production

Tomasz Matyja
Faculty of Transport and Aviation Engineering, Silesian University of Technology, Katowice, Poland
E-mail of corresponding author: tomasz.matyja@polsl.pl

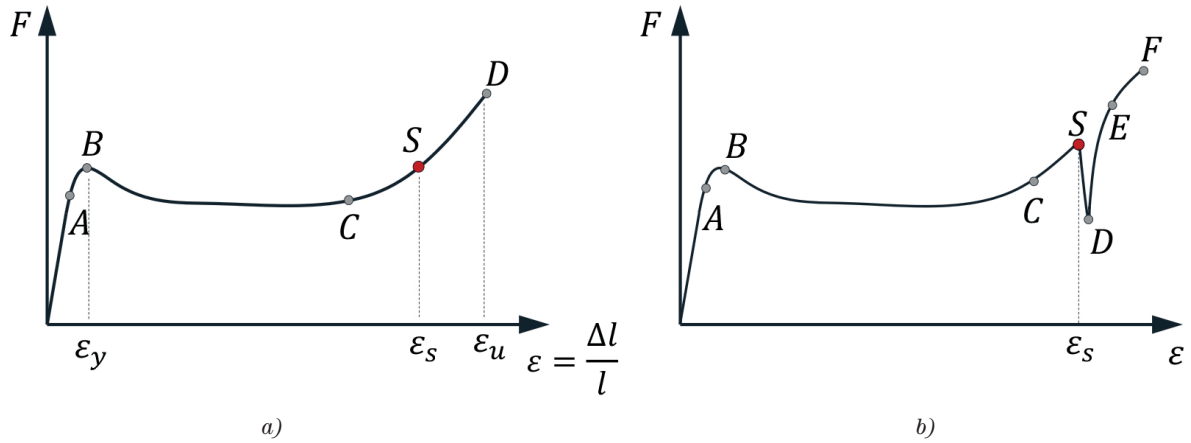


Figure 1 *The stretch film tensile tests: a) determining the sweet spot (sweet spot stretch level* $\varepsilon_s \approx 0.75\varepsilon_u, \varepsilon_u$ *- ultimate stretch), b) interruption of the tensile test to allow the stress relaxation*

technology. Any change of the film distributor usually requires new braking tests. For this reason, it is important to search for effective calculation methods to estimate the necessary number of film layers to ensure the load stability.

The simple simulation model of a loading unit secured with stretch film is presented below. This model is adapted to class A pallet with a layered structure. Performing the simulation requires collecting several experimental data. Among others, measurements of the stretch films strength properties are necessary. However, such tests are much cheaper and faster than the acceleration tests.

2 Basic properties of the stretch film

The stretch film is the most effective and efficient packaging material. The specific properties of the stretch film, such as: significant elongation of 400% and a certain type of shape memory - elastic recovery, result from its layered structure. Usually, the foil sheet is made of a few to several layers. Polyethylene and vinyl-based plastics, such as PVC, LDPE, EVA, LLDPE, are used for the film production. The final properties of the film depend on the selection of the above-mentioned materials and the production process and can be significantly variable.

A number of tests are performed to determine the properties of the film: classical ultimate tensile test; retention test, which measures the maximum load holding force and reduction in force over time; puncture test, which measures the maximum force required to puncture the film; cling test, which measures ability of a film to stick to itself at a selected pre-stretch level; quality test a selected stretch level, which checks homogeneity of the film along its length based on the measurement of the current stretch at a constant tensile force [13].

The stretch film tensile tests indicate that the film behaves like a hyper elastic body. On a typical graph obtained from a tensile test can be distinguished: linear elastic region, yield stress, necking and strain hardening region (Figure 1.a). The test is carried out in accordance with the direction of winding the film on the roll (MD - Machine Direction). The results obtained for the perpendicular direction (CD - Cross Machine Direction or Transverse Direction) are usually different.

Pre-stretching of film in the load wrapping process should be planned so as to reach the strain hardening zone, on the rising part of the tensile test graph. In technical jargon this point is referred to as a “sweet spot” [14]. Usually this point corresponds to a strain between 200% and 350%. By pre-stretching the film, an adequate value for the containment force is obtained. An additional positive effect of the pre-stretching is less film consumption. Based on the value of the tensile force at the sweet spot and the slope of the ascending part of the tensile graph, the approximate relationship between the containment force and the current stretch of the film can be determined as:

$$F_C \approx F_S + \frac{dF}{d(\Delta l)}\Big|_{\Delta l S}(\Delta l - \Delta l_S), \tag{1}$$

where $F_S, \Delta l_S$ - the initial value of the containment force and the elongation corresponding to the sweet spot, respectively. The derivative value (curve slope) can be considered as the spring stiffness.

Stresses in the stretch film relax very quickly and decrease by up to 50%. Therefore, some authors recommend interrupting the tensile test after reaching the sweet spot for one hour and then continue until the film breaks [2]. Thanks to this procedure, both the initial value of the containment force and the slope of the graph are closer to reality (Figure 1.b).

3 Equations of motion

The problem of investigation of the dynamics of packages forming the load unit is a complex issue of the multi-body contact (MBS) with friction [15]. The solution of this problem is complicated by the interaction between the

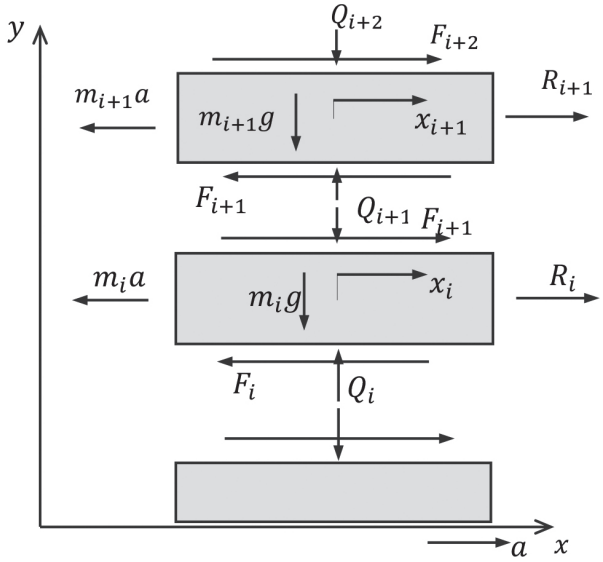


Figure 2 *Free body diagram of the load layers*

practically massless stretch film and packages. One way to consider the stretch film in the load unit model is to replace it with a cloud of equidistant points connected by springs of known stiffness [16]. At each step of simulation, both the positions of the packages relative to each other and the packages relative to the point cloud should be controlled. This type of simulation requires considerable computing power and is time consuming.

The paper proposes a simplified model of a class A pallet load unit, which has a layered structure. The assumption of a layered structure simplifies the model and means that the layers can only move horizontally with respect to each other.

Based on the free body diagram of the load layers (Figure 2), equations of motion of the individual layers can be written in the form:

$$\begin{aligned} &m_i\ddot{x}_i = -m_i a - F_i + F_{i+1} + R_i, i = 1,\ldots,N \\ &F_{N+1} = 0, \end{aligned} \tag{2}$$

where:

m_i - mass of the layer,

a - acceleration of the global system (vehicle cargo space),

F_i - friction force between layer i and *i-1*,

R_i - the resultant reaction (restoring) force of the stretch film acting on the layer (appears from the containment force and film tensions),

x_i - displacement of the layer i relative to the global system.

The proposed layered model of the loading unit takes into account only the horizontal displacement of the layers. This is sufficient when the purpose of the simulation is to choose the appropriate value of the containment force. If the containment force is insufficient or the load is not wrapped in the film, one or more adjacent layers may lose their stability due to significant displacements. Therefore, the simulation stops automatically when it detects a loss of stack stability when the center of gravity of the top or several subsequent top layers extend beyond their base. This particular case is shown in the Figure 7.

The issue of the stacked block stability is well known from elementary physics [17]. The condition of continuing the simulation can be written as the following product of logical expressions:

$$\begin{aligned} &\prod_{k=0}^{N-1}\left[\left|\frac{\sum_{i=N-k}^{N} m_i x_i}{\sum_{i=N-k}^{N} m_i} - x_{N-k-1}\right| < \frac{L}{2}\right] = 1, \\ &x_0 = 0 \end{aligned} \tag{3}$$

where L - is the length of the load layers.

4 Friction force modeling

The slip-stick friction was considered in the model. The friction force is defined by Equation [18]:

$$F_i(\Delta\dot{x}_i, F_i^{input}) = \begin{cases} F_i^{stick}(F_i^{input}) & |\Delta\dot{x}_i| < V_{break} \\ F_i^{slip}(\Delta\dot{x}_i) & |\Delta\dot{x}_i| \geq V_{break} \end{cases}, \tag{4}$$

where:

$\Delta\dot{x}_i = \dot{x}_i - \dot{x}_{i-1}$ - it is the speed of layer "i" with respect to the layer located below, V_{break} - is the velocity of the layers breaking.

When the relative velocity of the layers is lower than the assumed very low speed of the break $|\Delta\dot{x}_i| < V_{break}$ then stick occurs and the force of static friction should be determined from the condition of balance of the upper layer relative to the lower one:

$$F_i^{stick}(F_i^{input}) = \begin{cases} -F_i^{break}\, sign(F_i^{input}) & |F_i^{input}| > F_i^{break} \\ F_i^{input} & |F_i^{input}| \leq F_i^{break} \end{cases}, \tag{5}$$

where:

$F_i^{break} = \mu_{si} Q_i$ is the maximum static friction value,

μ_{si}, Q_i - are the static friction coefficient and normal force between the surfaces of the layer "i" and "i - 1", respectively.

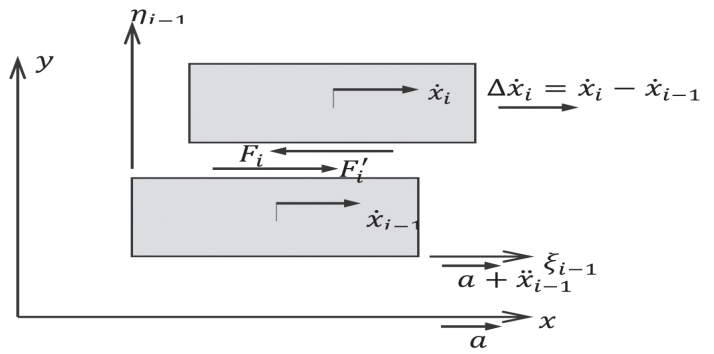


Figure 3 *The local non-inertial coordinate system for determining the stick friction force*

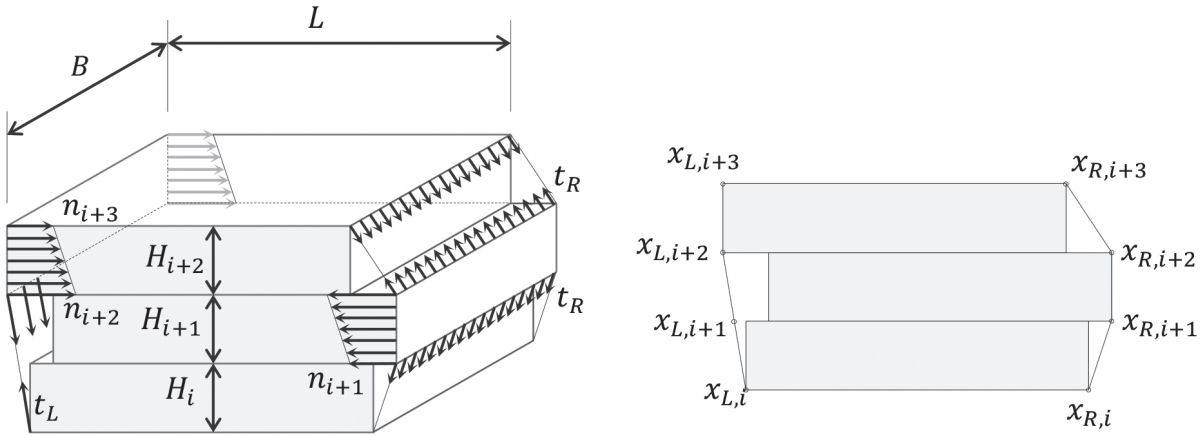


Figure 4 *Assumed tensions distributions in the stretch film as a result of mutual displacement of the load layers and control points on the edge of the foil*

The force F_i^{input} can be determined from the equilibrium equation of a layer "i", which is temporarily at rest, relative to the layer "i - 1", i.e. in a non-inertial system moving with acceleration $a + \ddot{x}_{i-1}$ (Figure 3):

$$F_i^{input} = -m_i(a + \ddot{x}_{i-1}) + F_{i+1} + R_i, i = N, \ldots, 1, \quad (6)$$

To determine the force F_i^{input} one must know the friction force F_{i+1} from the upper layer. The task is statically determined. However, the process of calculating the value of the friction force must take place from top to bottom.

In the case of slip $|\Delta\dot{x}_i| \geq V_{break}$, the friction force is determined from Equation [19]:

$$F_i^{slip}(\Delta\dot{x}_i) = \sqrt{2e}(F_i^{break} - F_i^{coulomb}) \cdot \exp\left[-\left(\frac{\Delta\dot{x}_i}{V_{stribeck}}\right)^2\right]\frac{\Delta\dot{x}_i}{V_{stribeck}} + F_i^{coulomb}\tanh\left(\frac{\Delta\dot{x}_i}{V_{coulomb}}\right) + f\Delta\dot{x}_i, \quad (7)$$

where:
$F_i^{coulomb} = \mu_{ki} Q_i$ - kinetical friction force,
μ_{ki} - kinetical friction coefficient between layers "i" and "i - 1",
$V_{stribeck}$ - speed threshold of the Stribeck phenomenon,
$V_{coulomb}$ - Coulomb speed threshold,
f - viscous friction coefficient.

The hyperbolic tangent in Equation (7) guarantees the continuity of the friction force when the relative velocity passes through zero. In the model it was assumed that [19]:

$$V_{stribeck} = \sqrt{2}\, V_{break} \qquad V_{coulomb} = 0.1 \cdot V_{break}, \quad (8)$$

5 Restoring force modeling

A two-dimensional problem was considered. Mobile control points have been introduced on both sides of the film, which are always at the same height regardless of the vertical stretching of the film. These points correspond to boundaries of the layers. In Figure 4 the edges of the deformed foil are visible in the form of polylines. The control points on the left and right of the palette are also shown.

To determine the position of the film, relative to the displaced layers of cargo, a simple recursive algorithm for subsequent stretch film tension was proposed. Algorithm does not require solving the contact problem. The foil, represented by the control points, is stretched successively to the most protruding vertices of the packages. The adopted algorithm requires additional simplifying assumptions. Friction between the foil and the load was omitted. Due to that, it can be assumed that the film is uniformly tensioned in the vertical direction. This tension results from the

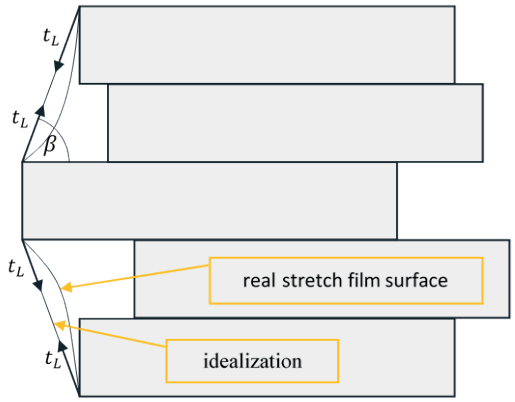


Figure 5 *Additional forces acting on layers due to the vertical stretching of the film*

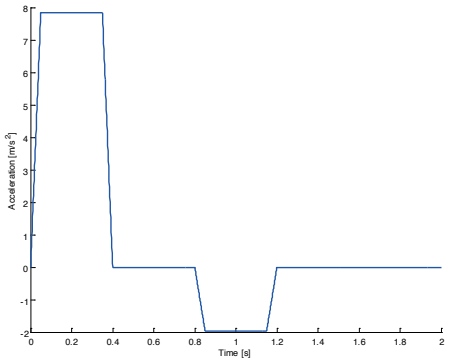


Figure 6 *Acceleration graph during the emergency braking test*

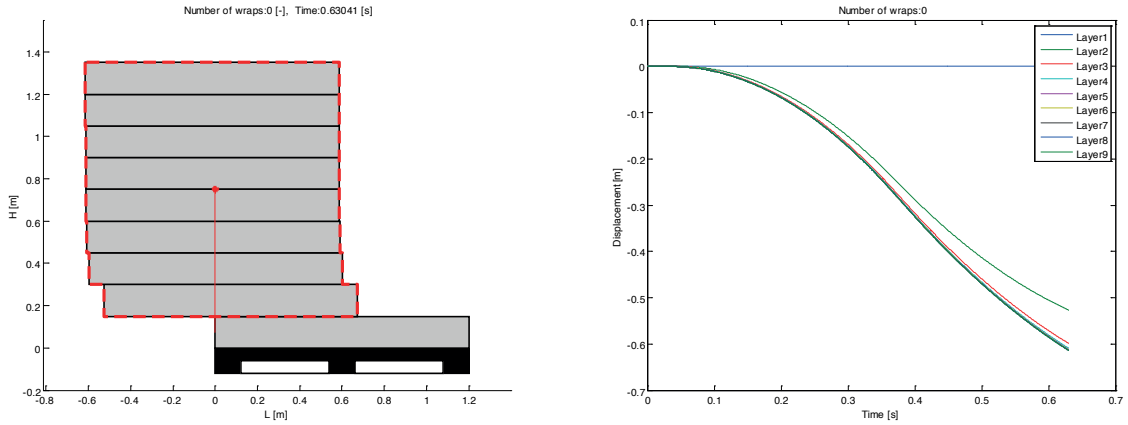


Figure 7 *Displacements of unsecured layers*

current length of the polyline, whose vertices are the control points. The curvature of the foil between the control points was also omitted (Figure 5). The operation of the algorithm explains the following pseudo code:

```
StretchFoilOverPoints(StartPoint, EndPoint)
{   FindPoint = {Find the node of the rectangle representing
    the layer of cargo located furthest on the left (right)
    side relative to the current position of the left (right)
    polyline representing the foil edge};
If exist FindPoint
{   Move the left (right) polyline nodes to stretch the foil
    on the nodes: (StartPoint, FindPoint, EndPoint);
    StretchFoilOverPoints(StartPoint, FindPoint);
    StretchFoilOverPoints(FindPoint, EndPoint);
    }
}
```

When the load is undeformed, the film pressure acts evenly on all sides. After the load layers have moved, the pressure distribution changes. The foil now presses on the most shifted layers of charge. When the layer moves away from the edge of the film, the pressure quickly decreases to zero. However, when the layer presses against the edge of the film, an additional pressure increase occurs as a result of stretching the film in the horizontal direction.

The horizontal film tension force per unit of length at selected control points can be expressed by formula:

$$n_i = n_0 + k_H(x_{Ri} - x_{Li} - L), \tag{9}$$

where:
n_0 - tensile force obtained initially in the wrapping process (containment force),
k_H - horizontal stiffness of the wrapped film,
$x_{Li}(t), x_{Ri}(t)$ - coordinates of the foil control points on the left-hand and right-hand side.

Similarly, the vertical force of film stretching per unit length on the left- (right-hand) side of the loading unit is

$$t_{L(R)} = t_0 + k_v\left(\sum_{i=1}^{N} l_{L(R)i} - H\right), \tag{10}$$

where:
t_0 - pre-tension of the film due to wrapping,
k_v - vertical film stiffness,
$l_{L(R)i}$ - distance between the control points on the left (right) side.

The individual layers are affected by the resultant force due to the film tensions. Part of the force from the horizontal tension of the film can be expressed by formula:

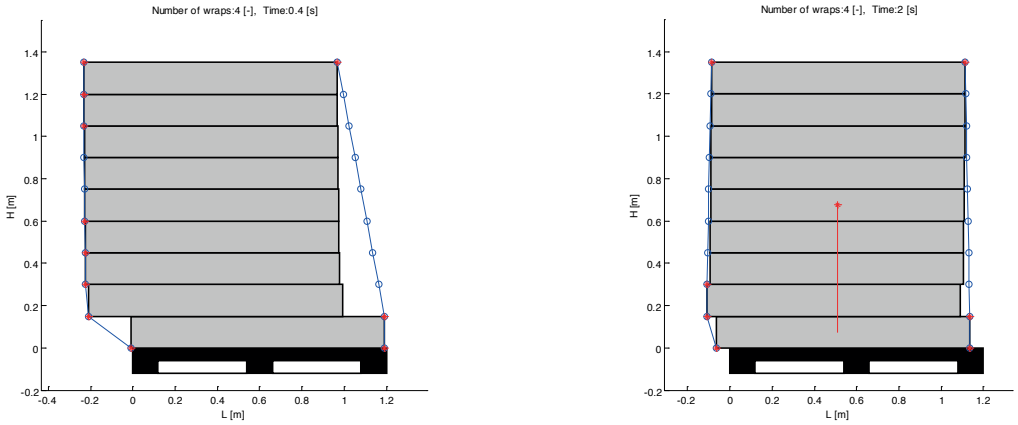


Figure 8 *Deformation the of loading unit (first variant of wrapping)*

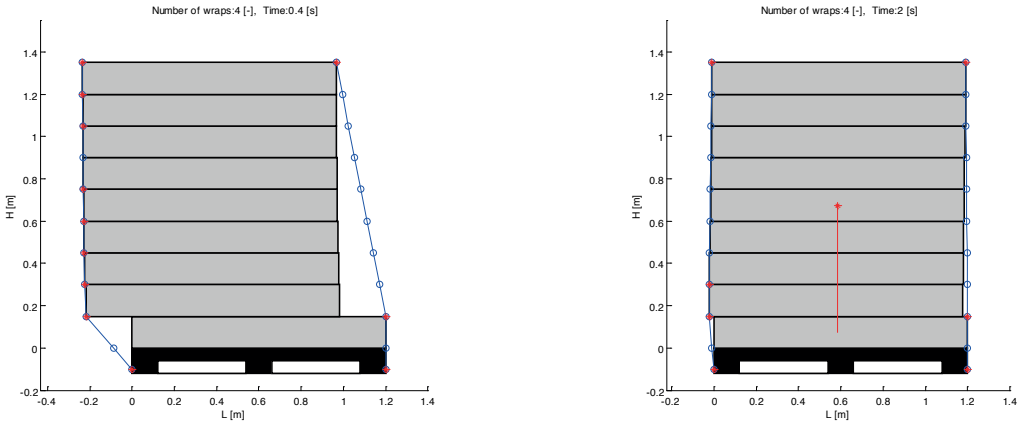


Figure 9 *Deformation the of loading unit (second variant of wrapping)*

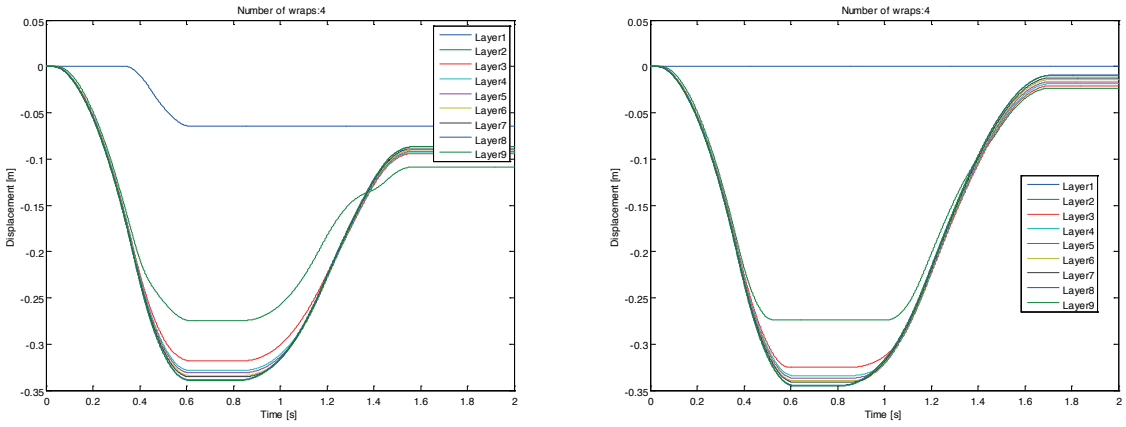


Figure 10 *Comparison of the displacements of layers (4 wraps, variant 1 left and variant 2 right)*

$$R_i = 2n_0H_i\left\{\exp\left[\alpha\left(\frac{x_i - \min\limits_{i=1:N}(x_i)}{\Delta}\right)^2\right] - \exp\left[\alpha\left(\frac{\max(x_i) - x_i}{\Delta}\right)\right]\right\} + \begin{cases} (n_p + n_{p+1} - n_0)H_p & x_{Lp} = x_{Lp+1} = x_p = \min\limits_{i=1:N}(x_i) \\ -(n_q + n_{q+1} - n_0)H_q & x_{Rq} = x_{Rq+1} = x_q = \max\limits_{i=1:N}(x_i) \\ 0 & i \neq p \wedge i \neq q \end{cases} \tag{11}$$

where: H_i - layer height,

α - dimensionless containment force vanishing factor (e.g. $\alpha = \ln 0.01$),

Δ - assumed distance of the containment force disappearance.

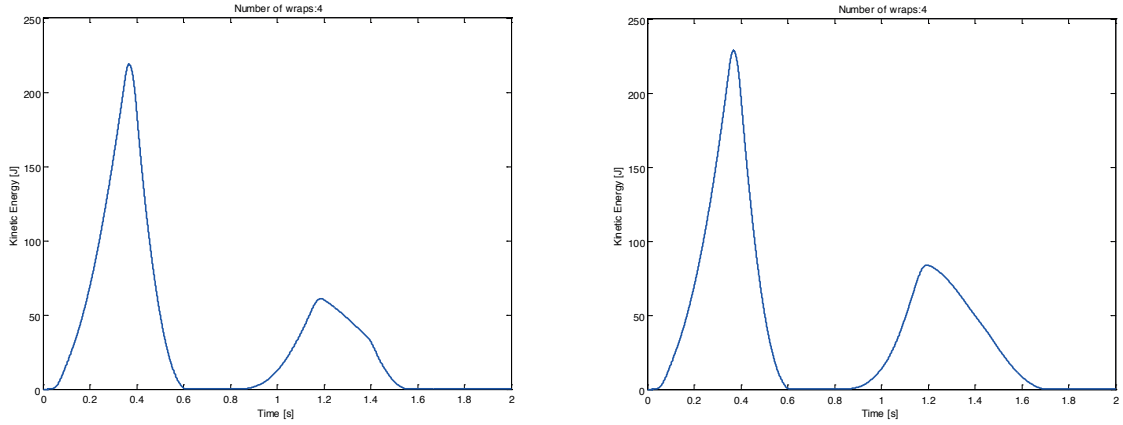


Figure 11 *Comparison of the total kinetic energy (4 wraps, variant 1 left and variant 2 right)*

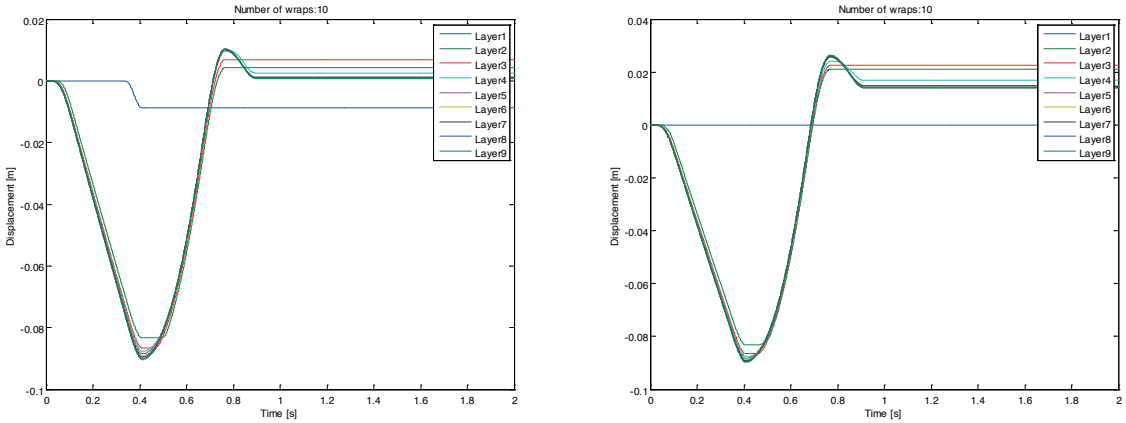


Figure 12 *Comparison of the displacements of layers (10 wraps, variant 1 left and variant 2 right)*

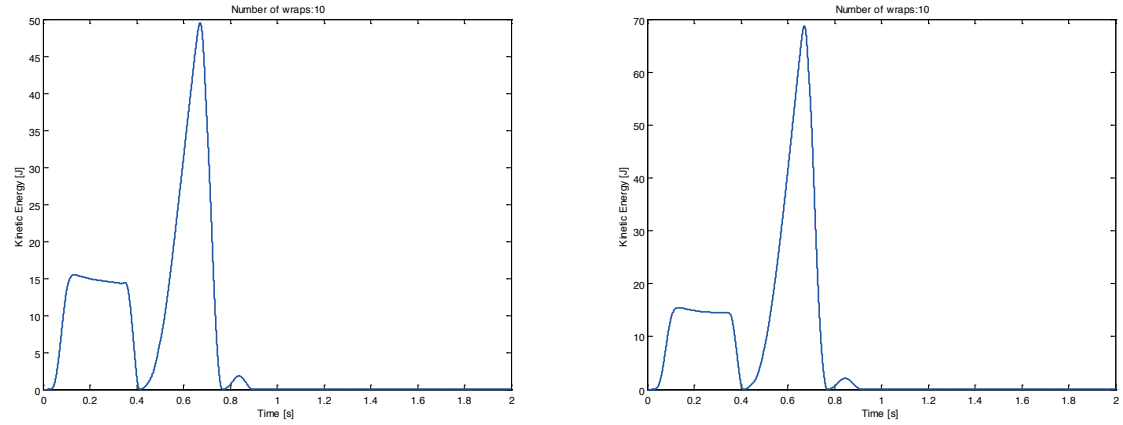


Figure 13 *Comparison of the total kinetic energy (10 wraps, variant 1 left and variant 2 right)*

Horizontal projections of forces caused by the vertical tension of film will additionally modify values of the restoring forces. However, the vertical projections will change the value of normal forces and consequently also the friction forces (Figure 4).

The containment force and required stiffnesses of the stretch film can be estimated from the film tensile tests, pre-stretching level and number of wraps. The containment force can be also measured [20].

6 Examples of simulation results

The mathematical model described above has been implemented in Matlab environment. In most cases, the differential equations were solved numerically using a standard ode45 solver and an automatically chosen variable time step. However, it has been observed that when the containment force is much higher than necessary and it significantly limits the movement of the load layers,

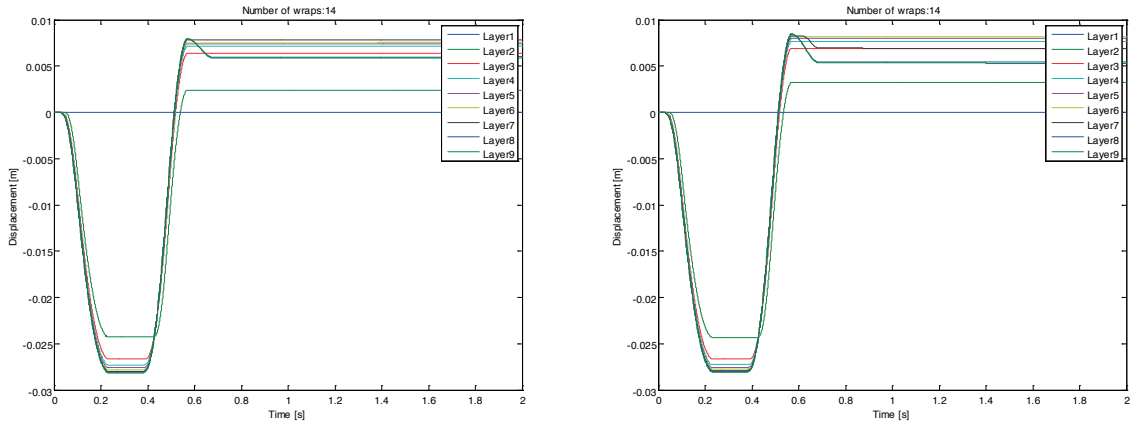


Figure 14 *Comparison of the displacements (14 wraps, variant 1 left and variant 2 right)*

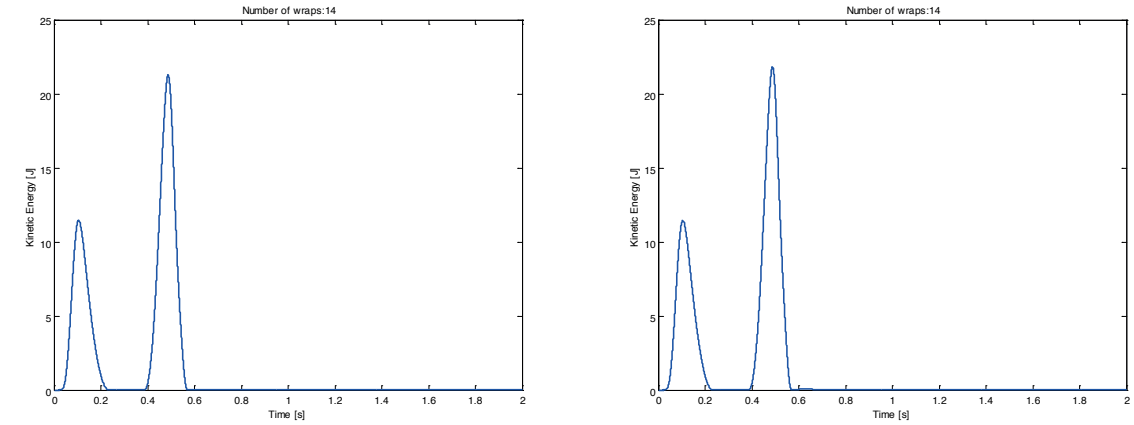


Figure 15 *Comparison of the total kinetic energy (14 wraps, variant 1 left and variant 2 right)*

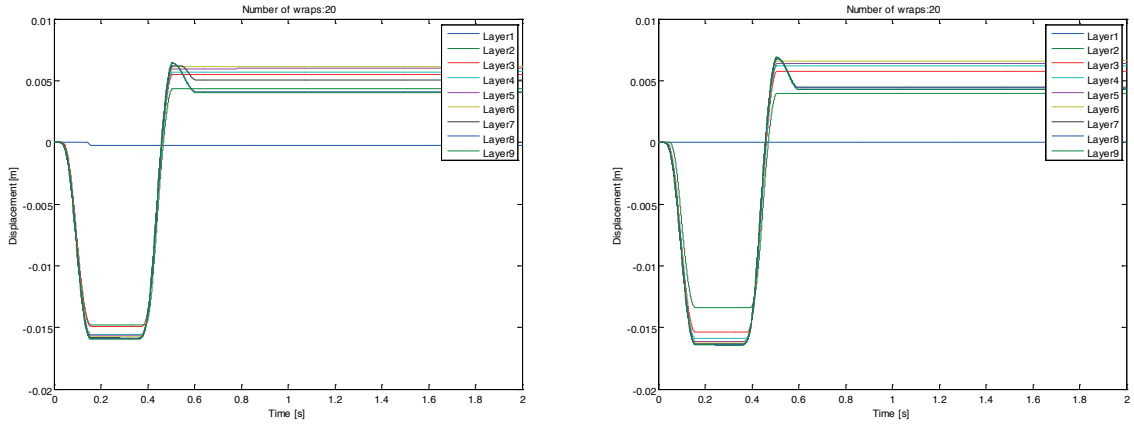


Figure 16 *Comparison of the displacements (20 wraps, variant 1 left and variant 2 right)*

the time of numerical calculations increases rapidly. In this case, solvers dedicated to stiff problems are much more effective.

The test was carried out in accordance with EUMOS guidelines. These types of tests are usually carried out in real conditions using special mobile platforms [21]. Emergency braking of a vehicle transporting a load unit with retardation of $0.8g$ during the first $400ms$ and next acceleration of $0.2g$ was simulated (Figure 6).

A medium heavy unit made of 9 identical layers was considered. The length and height of the layer were $1.2m$ and $0.15m$, respectively. The layer weight $50kg$. The coefficient of kinetic friction between layers was equal 0.3, while between layer and pallet 0.4. To model the Stribeck effect, it was assumed that the coefficient of increase in friction force is equal 120% and breakaway speed is equal $10^{-5} m/s$. The coefficient of viscous friction was assumed to be equal to $20Ns/m$. The containment

force per wrap and per unit length was equal $85 N/m$. Assumed distance of the containment force disappearance was $0.02 m$.

Figure 7 shows the visualization and values of the layers displacements if the cargo is unsecured with stretch film. The simulation stops automatically when it detects a loss of stack stability when the center of gravity of the top or several subsequent top layers extend beyond their base.

Wrapping the load four times with foil reduces the layer displacement. However, the loading unit is significantly deformed in the braking phase. Two variants were analyzed: the first when only the load is wrapped and the second when the pallet is also partially wrapped. Figures 8 and 9 visualize the deformation of the unit at the end of the emergency braking and at the end of the simulation test in both variants of wrapping. Comparison of displacement values (Figure 10) and total kinetic energies (Figure 11) does not allow to clearly determine which of the wrapping methods is better. Wrapping layers of packaging together with the pallet means that the final displacement of the load on the pallet is smaller and the bottom layer does not move. However, there is no doubt that the number of layers of stretch film is too small.

A simulation was then carried out in the case of wrapping the load with stretch film ten times. The results of both wrapping options are comparable. However, in the first variant the final displacement of the layers is slightly smaller (Figure 12). This is done at the expense of slightly shifting the first layer relative to the pallet. In addition, the peak of kinetic energy in the second phase of motion is smaller in the first variant, which is visible in the energy diagrams (Figure 13). Increasing the number of wraps to fourteen means that the results are practically identical in both variants (Figures 14 and 15).

Further increasing the number of wraps to twenty makes no sense, because it does not significantly improve the results (Figure 16). It can be considered that a load unit wrapped fourteen times is sufficiently stable.

7 Conclusions

The model of dynamics of the load unit, secured with the stretch film, presented in the paper is greatly simplified due to assumed layered structure and the method of modeling the impact of film on packages. If the load on the pallet is stacked in columns or on bricks, then obtained results can be treated only as a rough approximation of reality. The model guarantees more accurate results when the layer is one package or panel.

The smaller the range of movements performed by the layers, the higher the stability (also called rigidity) of the load unit. Of course, rigidity in the context of solid mechanics will never be achieved. At the same time, the smaller the displacements, the more accurate the method of determining the force of the package's interaction with the stretch film. Because the main purpose of the simulation was to estimate the value of the containment force (or the number of film layers if initial pre-stretching is known) guaranteeing the stability of the loading unit, the proposed model seems to be sufficiently effective. The goal is not to accurately determine the displacement of individual layers if the stability is not achieved.

Wrapping the load together with the pallet is a common practice. Studies have shown that this effectively blocks the movement of the bottom layer, but it is not strictly necessary to ensure the stability of the loading unit. More important is how many layers of film were used. This note applies when there is increased friction between the bottom layer of the load and the pallet.

It was also noted that there is a certain limit quantity of the stretch film wraps beyond which the results will no longer improve. This is the basis for optimizing the stretch film consumption.

The numerical investigation suggests that a good measure of the load unit stability assessment can be the maximum kinetic energy generated as a result of movement of individual layers. The lower kinetic energy, the better the load protection and stability.

References

[1] WHITE, M. S, HAMNER, P. Pallets move the world: the case for developing system-based designs for unit load. *Forest Product Journal.* 2005, **55**, p. 8-16. ISSN 0015-7473.

[2] BISHA, J. *Correlation of the elastic properties of stretch film in unit load containment.* Ph.D. dissertation. Virginia Tech, Blacksburg, VA, USA: Virginia Tech, Wood Science and Forest Products, 2012.

[3] DUNNO, K. D., WYNS J., COOK, J. Evaluation of containment force variability between different grades of stretch film. *International Journal of Advanced Packaging Technology* [online]. 2017, **5**(1), p. 267-274. eISSN 2349-6665. Available from: https://doi.org/10.23953/cloud.ijapt.318

[4] CIESLA, M., OPASIAK, T. *Load securing in cargo transport units.* Katowice: PAN, 2014.

[5] FIGLUS, T., KUCZYNSKI, L. Selection of a semi-trailer for the haulage of long oversize loads, taking into account an analysis of operational damage. In: 2018 XI International Science-Technical Conference Automotive Safety: proceedings. IEEE. ISBN 978-1-5386-4578-9.

[6] SKRUCANY, T., SEMANOVA, S., KENDRA, M., FIGLUS, T., VRABEL, J. Measuring mechanical resistances of a heavy good vehicle by coastdown test. *Advances in Science and Technology-Research Journal* [online]. 2018, **12**(2), p. 214-221. ISSN 2299-8624. Available from: https://doi.org/10.12913/22998624/91889

[7] SINGH, J., SAHA, K., SEWELL, T.: Evaluation of stability of unit loads for tilt and shock events during distribution. *Journal of Applied Packaging Research* [online]. 2017, **9**(3), p. 73-93. ISSN 1557-7224. Available from: https://scholarworks.rit.edu/japr/vol9/iss3/5

[8] SKRUCANY, T., GNAP, J. The effect of the crosswinds on the stability of the moving vehicles. *Applied Mechanics and Materials* [online]. 2014, **617**, p. 296-301. ISSN 1662-7482. Available from: https://doi.org/10.4028/www.scientific.net/AMM.617.296

[9] FEDORKO, G., MOLNAR, V., HONUS, S., NERADILOVA, H., KAMPF, R. The application of simulation model of a milk run to identify the occurrence of failures. *International Journal of Simulation Modelling* [online]. 2018, **17**(3), p. 444-457. ISSN 1726-4529, eISSN 1996-8566. Available from: http://doi.org/10.2507/IJSIMM17(3)440

[10] Directive 2014/47/EU of the European Parliament and of the Council on the technical roadside inspection of the roadworthiness of commercial vehicles circulating in the Union [online] [Viewed 2016-12-16]. Available from: https://publications.europa.eu

[11] EUMOS 40509-2012: Test method for load unit rigidity [online] [Viewed 2019-12-16]. Available from: https://eumos.eu/quality-standards

[12] FINNEMORE, D. Stretch wrap film and pallet load stability. *Powder and Bulk Engineering* [online]. 2017, **5**, [accessed 2019-12-16]. ISSN 0897-6627. Available at https://www.powderbulk.com/enews/2017/editorial/story_pdf/pbe_04_19_17rihf.pdf

[13] ASTM D4649-03. Standard guide for selection and use of stretch wrap films. West Conshohocken, PA, USA: American Society for Testing and Materials (ASTM) International, 2016.

[14] BP-Film-Testing [online] [Viewed 2019-12-16]. Available from: https://www.bestpackaging.com/wp-content/uploads/2018/02/ BP-Film-Testing.pdf

[15] FEDORKO, G., IVANCO, V., MOLNAR, V., HUSAKOVA, V. Simulation of interaction of a pipe conveyor belt with moulding rolls. *Procedia Engineering* [online]. 2012, **48**, p. 129-134. ISSN 1877-7058. Available from: http://doi.org/10.1016/j.proeng.2012.09.495

[16] DANJOU, S., OSTERGAARD, N. Application of computational MBD for simulation of wrap packaging performance. *Journal of Applied Packaging Research* [online]. 2018, **10**(3), p. 25-36. ISSN 1557-7224. Available from: https://scholarworks.rit.edu/japr/vol10/iss3/3

[17] HALL, J. F. Fun with stack blocks. *American Association of Physics* [online]. 2005, **73**(12), p. 1107-1116. ISSN 0002-9505, eISSN 1943-2909. Available from: https://doi.org/10.1119/1.2074007

[18] PENNESTRI, E., ROSSI, V., SALVINI, P., VALENTINI, P. P. Review and comparison of dry friction force models. *Nonlinear Dynamics* [online]. 2016, **83**(4), p. 1785-1801. Available from: https://doi.org/10.1007/s11071-015-2485-3

[19] MATLAB documentation. Friction in contact between moving bodies [online] [Viewed 2019-12-16]. Available from: https://www.mathworks.com/ref/translationalfriction

[20] The transportable test pallet [online] [Viewed 2019-12-16]. Available from: https://www.bestpackaging.com.au/

[21] Pallet stability testing - ESTL [online] [Viewed 2019-12-16]. Available from: https://s3-eu-west-1.amazonaws.com/upload

Galimkair Mutanov - Sayabek Ziyadin - Askhat Serikbekuly

APPLICATION OF SYSTEM-DYNAMIC MODELING TO IMPROVE DISTRIBUTION LOGISTICS PROCESSES IN THE SUPPLY CHAIN

This article presents an approach of improve logistics processes with the system dynamics modeling of two different processes scenarios. System dynamics is used, not only as a causal loop diagram, but calculated measures of end-user satisfaction indicators were provided by experts, as well.

Literature review of fundamental definitions and recent ideas in subject of supply chain management (SCM), system dynamics (SD), business processes model and notation (BPMN) and distribution logistics for dozens of major sources, has been made.

In the work, methods of expert interviews, content analysis of recent publications, system-dynamic modeling and business process modeling, were used. Developed BPMN-models of distribution logistics business processes are given in Appendices #1 and #2.

Keywords: *distribution logistics, system dynamics, BPMN, supply chain, modelling*

1 Introduction

Logistics is the process of planning, managing and implementing efficient and powerful flows, and storing goods and related information from the point of departure to the point of consumption, the purpose of which is to satisfy customer needs [1].

I. Gros [2] defines logistics as the planning, management and execution of goods flows, starting from development and purchase through production and distribution in accordance with the end customer, so that all the market requirements are met with minimal costs and capital expenditures.

According to Straka [3], logistics is a system in which there is an effect on elements in order to coordinate material, information and financial flows, as a result of which the customer's requirements are accordingly met and the corresponding economic effect is ensured.

As can be seen from the above definitions, one of the key objectives of logistics is to ensure the satisfaction of end consumers.

Importance of ensuring the customer satisfaction is also confirmed in the SCOR model (Supply Chain Operations Reference-model), proposed by the international organization - The Supply Chain Council (SCC), one of the five main attributes of which is response.

A response is the speed at which tasks are performed. The speed at which the supply chain delivers products to the customer [4].

Kazakhstan's wholesale companies, which are elements of the supply chain of consumer goods, especially small and medium ones, cannot always ensure full compliance with the above principles and goals of logistics, due to insufficient resources to improve the quality of service, dynamically changing needs of their customers and other complex factors. The task of supply chain management, according to Ivanov et al. [5], in this case is to provide a quick response to market changes and customer needs, while reducing the costs.

In order to successfully apply the supply chain management strategy in their activities, companies must first carefully analyze their current activities and identify the main weaknesses, eliminating which can ensure a higher speed and quality of service delivery and increase the level of the customer satisfaction. One of the effective ways to carry out such an analysis can be business processes modeling in the company. Modeling processes by describing actions and measuring the results of their processes allows organizations to constantly analyze them, thereby contributing to their improvement [6].

Business process management, according to [7] helps organizations to achieve their strategic goals through analysis, modeling, implementation, execution, control and continuous improvement of processes, which leads to increased productivity, increased innovation and quality.

In the process of modeling existing and proposed business processes in companies, the Business Process Model and Notation (BPMN 2.0) [8] was used - the most popular notation of business process modeling today.

Sterman [9] also emphasizes that modeling is an iterative, ongoing process for testing and validating

Galimkair Mutanov[1], Sayabek Ziyadin[2,*], Askhat Serikbekuly[3]
[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Korkyt Ata Kyzylorda State University, Kazakhstan
[3]UP Consulting, Almaty, Kazakhstan
*E-mail of corresponding author: sayabekz@gmail.com

models. Experiments in the virtual world determine the design and execution of experiments in the real world; in turn, experience in the real world leads to changes and improvements in the virtual world and in mental models [9].

According to Tolujew [10], the main purpose of modeling is to obtain numerical estimates of performance indicators in the system under consideration for such modes that have not yet been observed in reality. Continuous models based on differential equations as models of system dynamics are the most often used to reproduce production and logistics processes [11].

To determine the most effective scenario, which allows ensuring end-user satisfaction in the distribution chain of consumer goods, the system-dynamic modeling approach was applied. In the role of the end consumer for the company in the distribution chain, this article considers an independentdelivery point. Two process models in the BPMN 2.0 notation, which are considered in the work, act as scenarios of the system-dynamic model.

2 Literature review

2.1 Supply chain management

The term "supply chain management" (SCM) was first proposed by Oliver and Weber [12]. Oliver defined supply chain management as the process of planning, implementing and controlling the operations of the chain with the purpose to satisfy customer requirements as efficiently as possible. The SCM spans all the movements and storage of raw materials, work-in-progress inventory, and finished goods from point of origin to point of consumption [13].

Based on Chen et al. [14], Christopher [15], Harrison [16], managing a new type of logistics system called the supply chain in terms of minimizing total costs, increasing profits, improving customer service and reducing the impact of uncertainty on the system is called the chain management supplies [17].

Van Leeuwen and Tijssen [18] say that the supply chain management concept combines the tasks of various disciplines: logistics (minimizing costs in the supply chain), operational management (efficient inventory and production management), marketing (focus on creating values for the client) and market relationships (interaction with partners in the supply chain).

Akhmetkalieva [17] calls the concept of the supply chain management one of the most effective ways of managing commercial structures, which can significantly improve business optimization and help synchronize all the ypes of activities performed during production and distribution; optimization of added value for the end user; eliminate inefficient business activities. Currently, one of the main goals of the supply chain management concept is to optimize the entire system, which will achieve higher quality of service at lower overall costs, and not only maximize the productivity and profitability of individual business units of the chain [17].

Ivanov et al. [19] expressing their opinion based on the works of Sheffi and Rice [20], Tomlin [21], Wilson [22], Klibi et al. [23], Hahn and Kuhn [24], Simchi-Levi and Wei [25], Baghalian et al. [26], called the main problems in supply chain management uncertainty and dynamics, which require an understanding of the effectiveness of the supply chain from a dynamic point of view as part of efficiency and flexibility / sustainability.

The modern concept of supply chain management, in addition to maximizing the level of service and minimizing costs, is complemented by the need to ensure flexibility and sustainability of the supply chainin the face of increasing environmental problems and limited resources [5].

2.2 Distribution logistics

Distribution logistics is an integral and one of the most important components of the entire supply chain. There are various definitions that describe the essence of distribution logistics.

Ihde [27] suggests that distribution logistics deals with issues related to temporal and spatial related transfers of real goods and among the systems of labour division.

Distribution logistics is responsible for all warehouse and transport movements of goods for consumers and related information, management and control of activities [28].

Distribution logistics, as stated by DHL Logbook [29], includes all types of activities related to the provision of finished goods for the client, which can be sent to the client immediately after their production, to a warehouse for further processing or, if necessary, to other regional distribution warehouses.

Distribution logistics should provide the most appropriate way to analyze, select and implement all the activities, as well as strategic and other decisions related to the provision of products to the client in such a way as to achieve error-free market activity [3].

A detailed definition of distribution logistics is described by Kappauf et al. [30], who describe it as a business process for selling goods, including delivery, transportation to the customer and subsequent billing. It links the production logistics of the company (and/or its external procurement department in the case of goods produced from outside) with the requirements (orders) of customers. The main objective of distribution logistics is the efficient provision of goods for customers according to established criteria, such as quantity, time and price (p. 171).

In general, distribution logistics plays an important role both at the strategic level, determining the sales policy, the distribution network structure of the company, regions of distribution of the product, the location of system elements, methods of order processing and at the tactical level, determining the most efficient use of resources in building warehouses and buying and the use of necessary equipment for storage and transportation of goods, types of vehicles, etc. [31].

Based on the presented definitions and opinions of various authors regarding distribution logistics, its importance is obvious not only in the logistics processes of the company, but also in general for the company's activities, its successful development and gaining an attractive position in the market.

2.3 System dynamics

System dynamics was founded back in the late 1950s at the Massachusetts Institute of Technology under the direction of J Forrester [32]. In 1961, Forrester published the first classic book, which set out the concept of system dynamics, which is a direction in the study of complex systems, exploring their behavior over time and depending on the structure of system elements and the interaction between them, including causal relationships, feedback loops, reaction delays, environmental influences and others [33].

In part, system dynamics is a method for developing and testing formal mathematical models and computer modeling of complex nonlinear dynamic systems; it has much in common with other modeling methods [9].

Sterman [9] describes the key principles that must be considered when building models for the successful application of system dynamics:

A model should be created to solve a specific, existing rather than abstract problem;

1. The model needs to be integrated into the project at the very beginning of its implementation - the existing problem should be identified before the start of the project, based on past experience of the company;
2. The model must be constantly tested;
3. The main attention should be paid to the most significant, key elements of the model and not to focus on the details [9].

Importance of considering the project as a holistic system within the framework of the system dynamics methodology, without focusing on details and operations, was discussed by Rodrigues and Bowers [34].

According to Bendoly E. [35], the system dynamics methodology emphasizes the importance of considering the interaction and development of system elements, as well as providing a structured analytical image of processes.

Mutanov et al. [36] call the system dynamics methodology, which allows modeling complex systems at a high level of abstraction, without taking into account small details, the basis of an intelligent information system.

The methodology of the system-dynamic modeling is widely used in various fields of activity. Examples of the application of system dynamics can serve as the work of Sterman [9] - in building a car leasing strategy, Mutanov et al. [36] - in making investment decisions, Uehara et al. [37] - in the study and modeling of ecological and economic systems, Ivanov et al. [19] in the modeling and analysis of the supply chain, Tama et al. [38] - to optimize profits in the network of supplies of vegetable products and many others.

Distribution logistics usually crosses borders between organizations and an understanding of incentives and organizational barriers is a prerequisite for the successful application of any "scientific" solution [39], which also confirms the validity of applying the system dynamics methodology in modeling distribution logistics processes.

2.4 BPMN 2.0

Business Process Model and Notation (BPMN) - a system of conventions (notation) and their descriptions in XML for modeling business processes, which provides companies the opportunity to understand their internal business procedures in graphical notation and makes it possible to transfer these procedures in a standard way [8].

In addition, graphical notation will facilitate understanding of the interaction of productivity and business transactions between organizations. This ensures that enterprises will understand themselves and their business participants and will allow organizations to quickly adapt to new internal conditions and B2B business conditions [8].

The BPMN 2.0 is the latest, most relevant version of notation to date, which is widely used in modeling business processes.

Effectiveness of using this method in process modeling is confirmed by Bisogno et al. [40], where they say that the analysis of business processes using the BPMN notation can help to identify and correct the corresponding operating problems.

The BPMN is the leading language for noting business process modeling and allows business analysts and managers to increase the efficiency and quality of processes and performs the main function of modeling business processes - presenting processes so that the current process can be analyzed and improved in the future [41].

The advantage of the BPMN is that it is accessible to a wide range of users in the company - from analysts who create process diagrams and developers responsible for implementing business process execution technology, to managers and ordinary users who manage these business processes and monitor their implementation. Thus, the BPMN aims to bridge the gap between business process models and their implementation [42].

3 Methods and materials

The research object is a small and medium-sized production and wholesale trading companies of Kazakhstan, distributing goods to independent retail stores (delivery point). These retail stores, as the final consumer for the company in the chain of distribution of fast moving consumer goods (FMCG), are considered in this article. The subject of the study is the business process of receiving and processing an order from a delivery point by small

Table 1 *Input data for scenario No. 1 of the system-dynamic model*

Process step name	Time of execution, working hours	Time of waiting before previous process step, working hours	Probability, %
Visiting a delivery point by a sales agent	0.00	0.00	100
Reception and agreement of a DP's order by a sales agent	0.17	0.03	88
Possible adjustments by the sales agent to the order	0.08	0.00	43
Transfer of an order by a sales agent to an accountant / operator	0.08	0.00	100
Reception of order by operator from a sales agent	0.08	4.00	100
The operator checks the availability of goods in the warehouse and overdue receivables at the DP	0.02	0.08	100
Notification by the operator of the DP if there is a debt or shortage of goods in the warehouse	0.08	0.00	3
Registration by the operator of the order and preparation of accompanying documents (invoice)	0.03	0.00	100
The operator transfers the order to the warehouse to form	0.00	0.00	100
Collection and order confirmation	0.08	0.00	100
Correction, if necessary, accompanied by the coordination of adjustments with the DP	0.08	0.00	3
Determination of transport for delivery of the order, printout of route and loading lists	0.02	0.00	100
Transfer of the order to the forwarder, accompanied by the transfer of the necessary documents and information	0.01	0.00	100
Forwarder receiving goods and documents from the warehouse	0.01	0.00	100
Delivery of goods to the DP	0.08	0.00	100
Transferring an order to an DP	0.08	0.03	96
Acceptance of payment and receipt of signed documents (invoice) depending on the financial conditions of cooperation with the point and payment	0.02	0.00	67
Transfer of payment and signed documents (invoice) to the operator / accountant to complete the process.	0.02	0.00	67
Total duration per one order (in case of all events)	0.95	4.143	
Probability-adjusted duration per one order	0.71	4.139	
Total average duration of the process, taking into account expectation and probability*		4.85	

*average duration of the process, taking into account expectation and probability for the client - delivery point

and medium-sized manufacturing and wholesale trading companies of Kazakhstan.

In the process of modeling the existing and proposed business process diagrams, the BPMN 2.0 was used - a system of conventions and their descriptions in XML for modeling business processes [8].

Models in the BPMN 2.0 are used as the data source with two alternative scenarios for the system dynamic model.

In the process of modeling the distribution chain between companies and retail outlets, its main elements and the interaction between them, the system-dynamic modeling method was used [33].

The Ventana Systems' simulation software "VenSim" was used for the system dynamics modelling and Microsoft Visio 2016 for modelling processes in the BPMN 2.0 standard.

The following indicators were used as input data for calculating the order processing time for each of the scenarios of the system-dynamic model:

- duration of the process - the time that each steptakes;
- waiting time - the time between the steps of the process;
- probability - the probability of a specified process event occurring.

All the indicators taken as the basis of the calculations were obtained as a result of an expert interview and are the average values of the data of experts from their real experience.

Results of this expert interview are also described in the work of Mutanov and Serikbekuly [43].

In the framework of this work, the following basic terms were adopted:

- Delivery point (DP) - a retail outlet selling goods to consumers;

Table 2 *Input data for scenario No. 2 of the system-dynamic model*

Process step name	Time of execution, working hours	Time of waiting before the previous process step, working hours	Probability, %
Visiting a delivery point by a sales agent	0.00	0.00	0
Reception and agreement of a DP's order by a sales agent	0.00	0.00	0
Possible adjustments by the sales agent to the order	0.00	0.00	0
Transfer of an order by a sales agent to an accountant/operator	0.00	0.00	0
Reception of order byoperator from a sales agent	0.00	0.00	0
The operator checks the availability of goods in the warehouse and overdue receivables at the DP	0.00	0.00	0
Notification by the operator of the DP if there is a debt or shortage of goods in the warehouse	0.00	0.00	0
Registration by the operator of the order and preparation of accompanying documents (invoice)	0.00	0.00	0
The operator transfers the order to the warehouse to form	0.00	0.00	100
Collection and order confirmation	0.08	0.00	100
Correction, if necessary, accompanied by the coordination of adjustments with the DP	0.00	0.00	0
Determination of transport for delivery of the order, printout of route and loading lists	0.02	0.00	100
Transfer of the order to the forwarder, accompanied by transfer of the necessary documents and information	0.01	0.00	100
Forwarder receiving goods and documents from the warehouse	0.01	0.00	100
Delivery of goods to the DP	0.08	0.00	100
Transferring an order to an DP	0.08	0.03	96
Acceptance of payment and receipt of signed documents (invoice) depending on the financial conditions of cooperation with the point and payment	0.02	0.00	67
Transfer of payment and signed documents (invoice) to the operator/ accountant to complete the process.	0.02	0.00	67
Total duration per one order (in case of all events)	0.32	0.030	
Probability-adjusted duration per one order	0.30	0.029	
Total average duration of the process, taking into account expectation and probability		0.33	

- Sales agent (SA) - a representative of a trading company that periodically visits and receives orders from retail outlets (delivery points).

The numerical indicators, used in the model for calculating scenario No. 1 (the existing process), are shown in Table 1.

Details of the process steps:

- duration of the process of visiting a delivery point and duration of waiting are equal to zero, since it represents the moment the employee arrives at the point and has a 100% probability, due to the fact that it is the initial link in the entire process of receiving an order;
- reception and agreement of the order takes an average of 10 minutes and has a waiting time of about 2 minutes. The expectation may be due to presence of buyers, sales agents of other companies or delivery of goods from other companies at the time of the visit;
- making adjustments can take 10 minutes and does not have a waiting time, as it occurs immediately after receiving an order from a point in 50% of cases;
- the transfer of the order to operator takes 10 minutes, as it includes the transfer of information on all the commodity items and quantities for each, does not have a waiting time and is a mandatory process (100% probability);
- reception of order by the operator takes 5 minutes, has a 100% probability and the average waiting time is 4 hours, given that this step of the process is carried out in parallel for many orders;
- checking the availability of goods and debt takes an average of 1 minute and is carried out using the accounting system, has a waiting time of 5 minutes and a probability of 100%;
- notification of the DP in the absence of goods or the

Table 3 *Matching lead time and customer (DP's) satisfaction*

Order processing time,working hours	Level of satisfaction
1	5
3	4
5	3
8	2
13	1

presence of debt may take 5 minutes, waiting time = 0, probability 3% - the process is quite rare;

- placing an order and preparing documents takes 2 minutes, does not have a waiting time and is always carried out;
- the process of transferring an order to the warehouse has a duration of zero, since this process is carried out for all orders for the day at the same time and not for each separately;
- collection and confirmation of the order in the warehouse takes about 5 minutes;
- adjustment of the order when collecting it in the warehouse can take 5 minutes and is carried out quite rarely - in about 3% of cases;
- determination of transport for order delivery takes a short period of time - about 1 minute;
- the transfer of the order to the forwarder and its acceptance by the forwarder are also very short-term steps with a duration of approximately 30 seconds each, since it consists of loading the order into transport;
- delivery of the order to the DP has an average duration of 5 minutes, which is calculated taking into account the average number of freight forwarder orders per day, equal to 30 and the length of the working day;
- transferring an order to a point of sale takes 5 minutes, while waiting, as in the case of the process of receiving an order by a sales agent, lasting 2 minutes on average may be due to the presence of buyers, sales agents or forwarders of other companies at the point at the time of order transfer. The probability of this process is 95%, since provided that the DP has a type of payment in cash, but does not pay for the order upon delivery, the forwarder does not leave the order and returns the goods to the warehouse;
- acceptance of payment and documents from the DP can take about a minute and is carried out in 50% of cases, which depends on the type of financial conditions for the DP (cash/cashless/consignment);
- transfer of payment and documents by the forwarder to the cash desk also takes 1 minute and has a 50% probability, which is confirmed by the conditions described in the previous paragraph.

The numerical indicators, used for scenario No. 2 (the proposed process scheme), are shown in Table 2.

The key difference between scenarios is that the first eight steps of the process have zero indicators (execution time, waiting time and probability). This difference is provided by the using of a different online-services by trade companies. An important condition for the successful application of the second scenario is availability and use of a warehouse reservation system in a trading company. It will allow the delivery point to see the actual availability of goods in the warehouse of the supplier company.

It also excludes the step of adjusting the order and coordinating it with the point of sale when collecting at the warehouse, since the point cannot initially place an order in which there will be commodity items that are not in stock.

Otherwise, the steps of the process, including the duration of their fulfillment and expectations, as well as the probability, remain unchanged. The main indicator of the level of customer satisfaction (delivery points) is a processing time of the order - the period from the moment the order is received until delivery to the point.

The final satisfaction score is a score from 1 to 5, where 5 is the best indicator and 1 is the worst indicator. The assessment depends on the processing time of the order and this compliance is shown in Table 3

This correlation of the level of satisfaction with the processing time of the order was compiled on the basis of experimental data obtained by the author from experts in the wholesale industry. The basis of the correspondence between the values of the duration of order processing and customer satisfaction was the Fibonacci sequence.

4 Results and discussion

Based on the data obtained by Mutanov and Serikbekuly [43], on the structure of the process of receiving and processing orders in small and medium-sized wholesale trading companies of Kazakhstan, obtained as a result of a field study by expert interviews and as a continuation of this study and more and a detailed description and study of this process, modeling of the process of receiving and processing orders between participants in the studied supply chain was carried out.

The authors identified a standard scheme used by most of the studied production and wholesale trading companies in Kazakhstan. The process diagram in the BPMN 2.0 notation is presented in Appendix 1.

As can be seen in Appendix 1, the process of receiving and processing an order consists of a large number of steps and is accompanied by many approvals and the transfer of data and documents at its various steps.

The process includes 5 participants, the main roles and functions of which are described in Table 4.

Table 4 *Functions and roles of participants in the process of receiving and processing an order*

Participant role	Main role function	Workinglocation
Delivery point	gives the initial order, confirms it, expects, accepts, pays for this order, signs the supporting documents	Field
Sales agent	visits the point, accepts the order, communicates with the point on all issues (debts, balances, marketing campaigns, fulfillment of plans, legal relations), transfers the order to operator, makes adjustments to the order	Field
Operator	accepts an order from the merchant, checks the availability of goods and debts, prepares accompanying documentation, transfers the order to warehouse, monitors and accepts payment from the merchant/forwarder	Backoffice, office
Warehouse Manager	acceptance, assembly and loading of an order	Backoffice, office
Forwarder	receives goods and documents, delivers the order to the point, receives payment and supporting documents, transfers payment to the cashier	Field

In the majority of the small and medium-sized trading companies in Kazakhstan today, as shown in Appendix 1, the process consists of such main steps as:

1. Visiting a delivery pointby a sales agent
2. Reception and agreement of a DP's order by a sales agent
3. Possible adjustments to the order by the sales manager
4. Transfer of an order by a sales agent to an accountant / operator
5. Reception of order by operator from a sales agent
6. The operator checks availability of goods in the warehouse and overdue receivables at the DP
7. Notification by the operator of the DP if there is a debt or shortage of goods in the warehouse
8. Registration by the operator of the order and preparation of accompanying documents (invoice)
9. The operator transfers the order to the warehouse for picking goods for delivery
10. Collection and confirmation of the order with possible adjustment if necessary, accompanied by the coordination of adjustments with the point of sale (subject to the availability of goods in the accounting system, but its actual absence in the warehouse)
11. Determination of transport for delivery of the order, printout of route and loading lists
12. Transfer of the order to the forwarder, accompanied by the transfer of the necessary documents and information
13. Forwarder receiving goods and documents from the warehouse
14. Delivery of goods to the DP
15. Transferring an order to an DP
16. Acceptance of payment and receipt of signed documents (invoice) depending on the financial conditions of cooperation with the point and payment
17. Transfer of payment and signed documents (invoice) to the operator/accountant to complete the process.

As can be seen, the process contains many tasks, complicated by the need for approvals at various steps, which significantly increases its duration and also contains great risks of making mistakes and inaccuracies during the transfer of order information and documents at some steps.

The authors propose a more effective process of interaction with the DP, presented in Appendix 2.

The main improvement of the proposed interaction scheme, in comparison to the existing one at the moment, is the exclusion of a sales representative from the process, which automatically reduces a significant number of process steps previously performed by this participant. The exclusion of these steps from the process is ensured by use of a company information system that automatically performs these actions. Examples of such systems can be class systems: CRM (Customer Relationship Management), SFA (Sales Force Automation), OMS (Order Management System).

This change brings benefits such as:

1. Reduced costs of maintaining sales representatives
2. Reduced order lead time
3. Eliminated need to coordinate the order at the acceptance step, since the point makes the order on its own.

The next difference is the exclusion from the process of confirming the availability of goods in the warehouse, since this information is available at the initial checkout of the delivery pointand it is impossible to order the missing goods. This improvement is suitable for those companies in the information system of which there is a reservation system for orders - in this case, other agents and points will not be able to book the same product and the delivery pointwill receive its order. Accordingly, the steps of adjustment and confirmation of the order by the DP are also excluded from the process.

The operator's task of placing an order and creating related documents is not available in the information system, since upon receipt of an order from a point of sale these actions can be performed by the program.

Changes and improvements after applying the proposed business process scheme occur not only in the composition of the participants and in the number of steps, but also in the methods of transferring and storing information:

1. Order data now comes into the system directly from the source, without going through several participants, which eliminates the risk of loss or distortion of this information.

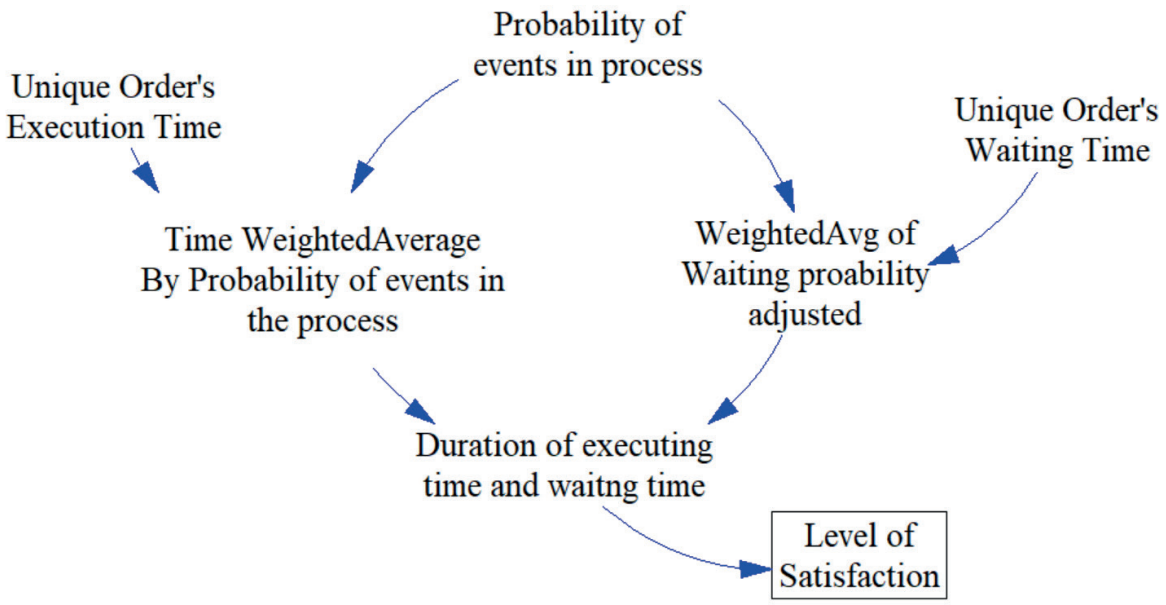


Figure 1 *System dynamic model for the two business process scenarios of order management in the distribution chain*

Table 5 *Results of comparison of the two scenarios*

	Time and speed	Level of satisfaction
Scenario #1	4.85	3
Scenario #2	0.33	5

2. The format for providing the forwarder with a routing sheet is changing. Now there is no need to compile and print it. The route sheet, which is created by the system, is available for the forwarder in the mobile application. This not only reduces the time it takes to compile, print and transmit, but also ensures maximum efficiency when creating the most optimal routes for the system, as well as flexible and quick adjustment of the route sheet by the system during the delivery process when the route actually changes.
3. The need for transmitting information to the forwarder about the availability of debt and the financial conditions of the DP- all this information is also available to him in the mobile application in real time.

As one can see from the comparison results, implementation of the proposed business process scheme for receiving and processing orders in a trading company greatly simplifies one of its most basic and significant business processes, improves its quality, reliability and efficiency, reduces the labor costs, and the most importantly the implementation of the process reduces the length of the cycle, because this indicator can be considered one of the key, since the speed and timeliness of fulfillment of obligations to the client, especially for small and medium enterprises' trade is of a great importance for the formation of positive and long-term relationship with the client and his loyalty to the company.

The importance of reducing such an indicator as the duration of the cycle is also confirmed by the fact that the companies that are considered are directly involved in distribution of the FMCG that are of paramount importance to the end consumers, respectively, ensuring the timely availability of these goods in the required volume is also of social significance.

In the VenSim program, a system-dynamic model was calculated (Figure 1) according to the two described scenarios.

Based on results of the modeling processes and comparing their scenarios in dynamics, the results of each scenario were summarized in Table 5.

As one can see from results shown in Table 5, the level of satisfaction and time spent on one execution of one order are improving. Particularly noticeable is improvement in the speed of order completion from 4.85 to 0.33 working hours.

5 Conclusion

Results of the study show an improvement in the speed indicators, and accordingly, satisfaction with the quality of logistics services due to implementation of the proposed scheme (Appendix 2).

The practical significance of the work is in the fact that the simulation results in the form of business process scenarios can be applied in creating and optimization of processes in real trading companies or distribution chains using information systems for managing the class orders:

- CRM (Customer Relationship Management),
- SFA (Sales Force Automation),

- OMS (Order Management System).

One example is ezapp.kz, popular among Kazakhstani trading companies, an online service designed to automate the process of accepting and processing orders.

Note that application of the proposed process scenario and use of the proposed tools would not be effective for absolutely all wholesale companies in Kazakhstan. There are such companies, processes, which have certain features and specifics and require a more thorough and individual approach to improving the organization of the process of receiving and processing orders.

To build a complete and high-quality supply chain model, one needs very accurate and detailed data for each individual supply chain process and each individual company. Receiving and processing an order is one of the important processes that were already carefully analyzed. Continuing the modeling of other logistics processes using the proposed methodology, it is possible to build a single complete model of the complete supply chain.

Further areas of application and development of this research are the trade "digital twin" as a combination of process and dynamic models into a single synchronized one. This will require obtaining actual data from information systems for order management of companies in the entire supply chain in real time. Given development of IoT, BPMS and RPA technologies - this approach seems possible at the moment.

Additional benefits of implementing online supply chain management are:

- improving the quality of customer service,
- the provision of additional services due to ability to monitor the movement of goods,
- reduction of paper operations,
- reduction of labor costs at each automated section of logistics processes.

References

[1] Council of Logistics Management. 1961

[2] GROS, I. Logistika ano ci ne? / Logistics yes or not? (in Czech). *Logistika*, *MesicnikHospodarskychNovin*.1995, **3**, p. 58. ISSN 1213-7693.

[3] STRAKA, M. Logistika distribuce. Ako efektivne dostat vyrobok na trh / Logistics distribution. How to effectively get the product on the market (in Slovak). Ruzomberok: EPOS, 2013. ISBN 9788056200155.

[4] Supply Chain Council [online]. 2019. [Viewed 2019-11-15]. Available from: http://www.apics.org/

[5] IVANOV, D., TSIPOULANIDIS, A., SCHONBERGERB, J. Operations and supply chain strategy. In*: Global supply chain and operations management* [online]. Springer, Cham: Springer Texts in Business and Economics. 2017. ISBN 978-3-319-24215-6, eISBN978-3-319-24217-0. Available from: https://doi.org/10.1007/978-3-319-24217-0_4

[6] PEREIRA, M. F., JACOBSEN, A. L., BIANCHI, I. S., MORITZ, G. O. Modelo de producao de material didatico: ouso da notacao BPMN emcurso a distancia / Model for production of teaching material: the use of BPMN notation in distance learning courses (in Portuguese). *RAI Revista de Administracao e Inovacao*. 2011, **8**(4), p. 45-66. ISSN 1809-2039. Available from: https://doi.org/10.5773/rai.v8i4.898

[7] MINONNE, C., TURNER, G. Business process management - are you ready for the future? *Knowledge and Process Management* [online]. 2012, **19**(3), p. 111-120. eISSN1099-1441. Available from:https://doi.org/10.1002/kpm.1388

[8] OMG. Business Process Modeling and Notation (BPMN). Version 2.0 [online]. 2019. Available from: http://www.bpmn.org/

[9] STERMAN, J. *Business dynamics*, *systems thinking and modeling for a complex world*. Boston: Irwin/McGraw-Hill, 2000.

[10] TOLUJEW, J. Graphic method for modeling material flows in supply chains. *The Journal of Economic Research and Business Administration*. 2018, **126**(4), p. 170-178. ISSN 1563-0358, eISSN2617-7161.

[11] REGGELIN, T., TOLUJEW, J. A mesoscopic approach to modelingand simulation of logistics processes. In Winter Simulation Conference: proceedings. IEEE, 2011. ISBN 978-1-4577-2108-3, p. 1513-1523

[12] OLIVER, R. K., WEBBER, M. D. Supply-chain management: logistics catches up with strategy. Outlook, Booz, Allen & Hamilton, Inc, 1982. Reprinted in: CHRISTOPHER, M. G. *Logistics: the strategic issue*. London: Chapman and Hall, 1992, p. 63-75.

[13] POUNDER, P., BOVELL, G., PILGRIM - WORELL, S. A Review of supply chain management and its main external influential factors. *Supply Chain Forum* [online]. 2013, **14**(3), p. 42-50. ISSN 1625-8312, eISSN1624-6039. Available from: https://doi.org/10.1080/16258312.2013.11517320

[14] CHEN, F., FEDERGRUEN, A., ZHENG, Y. S. Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Management Science* [online]. 2001, **47**(5), p. 693-708. ISSN 1526-550. Available from: https://doi.org/10.1287/mnsc.47.5.693.10484

[15] CHRISTOPHER, M. The agile supply chain: competing in volatile markets. *Industrial Marketing Management* [online]. 2000, **29**(1), p. 37-44. ISSN 0019-8501. Available from: http://dx.doi.org/10.1016/S0019-8501(99)00110-8

[16] HARRISON, A. Logistics Management and Strategy: Competing through the Supply Chain. 4 ed. Financial Times/ Prentice Hall, 2010, p. 384. ISBN 978-0273730224.

[17] AKHMETKALIEVA, S. K. The concept of supply chain management as a method of optimizing the doing business. *The Journal of Economic Research and Business Administration* [online]. 2018, **123**(1), p. 4-11. ISSN 1563-0358, eISSN2617-7161. Available from: https://be.kaznu.kz/index.php/math/article/view/1938

[18] VAN LEEUWEN T., TIJSSEN, R. Interdisciplinary dynamics of modern science: analysis of crossdisciplinary citation flows. *Research Evaluation* [online]. 2000, **9**(3), p. 183-187. ISSN 0958-2029, eISSN1471-5449. Available from: https://doi.org/10.3152/147154400781777241

[19] IVANOV, D., SOKOLOV, B., DOLGUI, A. The Ripple effect in supply chains: trade-off "efficiency - flexibility-resilience" in disruption management. *International Journal of Production Research* [online]. 2014, **52**(7), p. 2154-2172. ISSN 0020-7543, eISSN 1366-588X. Available from: https://doi.org/10.1080/00207543.2013.858836

[20] SHEFFI, Y., RICE, J. B. A supply chain view of the resilient enterprise. *MIT Sloan Management Review* [online]. 2005, **47**(1), p. 41-48. ISSN 1532-9194. Available from: https://sloanreview.mit.edu/article/a-supply-chain-view-of-the-resilient-enterprise/

[21] TOMLIN, B. On the Value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science* [online]. 2006, **52**, p. 639-657. Available from: https://be.kaznu.kz/index.php/math/article/view/2039

[22] WILSON, M. C. The impact of transportation disruptions on supply chain performance. *Transportation Research Part E: Logistics and Transportation Review* [online]. 2007, **43**, p. 295-320. ISSN 1366-5545. Available from: https://doi.org/10.1016/j.tre.2005.09.008

[23] KLIBI, W., MARTEL, A., GUITOUNI, A. The design of robust value-creating supply chain networks: A critical review. *European Journal of Operational Research* [online]. 2010, **203**(2), p. 283-293. ISSN 0377-2217. Available from: https://doi.org/10.1016/j.ejor.2009.06.011

[24] HAHN, G.J., KUHN, H. Value-based performance and risk management in supply chains: A robust optimization approach. *International Journal of Production Economics* [online]. 2012, **139**(1), p. 135-144. ISSN 0925-5273, eISSN1873-7579. Available from: https://doi.org/10.1016/j.ijpe.2011.04.002

[25] SIMCHI-LEVI, D., WEI, Y. Understanding the performance of the long chain and sparse designs in process flexibility. *Operations Research* [online]. 2012, **60**, p. 1125-1141. ISSN 0030-364X, eISSN 1526-5463. Available from: https://doi.org/10.2307/23323684

[26] BAGHALIAN, A., REZAPOUR, S., FARAHANI, R. Z. Robust supply chain network design with service level against disruptions and demand uncertainties: a real-life case. *European Journal of Operational Research* [online]. 2013, **227**(1), p. 199-215. ISSN 0377-2217. Available from: https://doi.org/10.1016/j.ejor.2012.12.017

[27] IHDE, G. B. Distributions - Logistik / Distribution logistics (in German). Stuttgart, New York: Gustav Fischer Verlag, 1978. ISBN 3437400517.

[28] SCHULTE, CH. *Logistik / Logictic* (in German). Munchen: Vahlen, 1991. ISBN 9783800614547.

[29] DHL Logbook [online]. 2013. Available from: http://www.dhldiscoverlogistics.com/cms/en/course/processes/distribution_logistics/definition.jsp

[30] KAPPAUF, J., LAUTERBACH, B., KOCH, M. *Logistic core operations with SAP* [online]. Verlag Berlin Heidelberg: Springer, 2011. ISBN 978-3-642-18203-7, eISBN978-3-642-18204-4. Available from: https://doi.org/10.1007/978-3-642-18204-4

[31] STRAKA, M. The position of distribution logistics in the logistic system of an enterprise. *Acta Logistica- International Scientific Journal about Logistics* [online]. 2017, **4**, p. 23-26. ISSN 1339-5629. Available from: https://doi.org/10.22306/al.v4i2.5

[32] FORRESTER, J.W. Industrial dynamics - a major breakthrough for decision makers. *Harvard Business Review* [online]. 1958, **36**(4), p. 37-66. ISSN 0017-8012. Available from: https://doi.org/10.1225/58404

[33] FORRESTER, J.W. *Industrial Dynamics*. Cambridge: MIT Press, 1961. ISBN 9780262060035.

[34] RODRIGUES, A., BOWERS, J System dynamics in project management: a comparative analysis with traditional methods. *System Dynamics Review* [online]. 1996, **12**(2), p. 121-139. eISSN1099-1727. Available from: https://doi.org/10.1002/(SICI)1099-1727(199622)12:2<121::AID-SDR99>3.0.CO;2-X

[35] BENDOLY, E. Systems dynamics understanding in projects: information sharing, psychological safety, and performance effects. *Production and Operations Management* [online]. 2014, **23**(8), p. 1352-1369. eISSN1937-5956. Available from: https://doi.org/10.1111/poms.12024

[36] MUTANOV, G., MILOSZ, M., SAXENBAYEVA, Z. H., KOZHANOVA, A. Investments decision making on the basis of system dynamics. In: *Modern approaches for intelligent information and database systems*. SIEMINSKI, A., KOZIERKIEWICZ, A., NUNEZ, M., HA, Q.-T. (eds.) [online]. Springer International Publishing AG, part of Springer Nature, 2018, p. 293-303. Available from: https://doi.org/10.1007/978-3-319-76081-0_25

[37] UEHARA, T., NAGASE, Y., WAKELAND, W. Integrating economics and system dynamics approaches for modelling an ecological-economic system. *Systems Research and Behavioral Science* [online]. 2016, **33**, p. 515-531. eISSN 1099-1743. Available from: https://doi.org/10.1002/sres.2373

[38] TAMA, I.P., AKBAR, Z., EUNIKE, A. Implementation of system dynamic simulation method to optimize profit in supply chain network of vegetable product. *IOP Conference Series: Materials Science and Engineering*. 2017, **337**(1), 012014. ISSN 1757-8981, eISSN1757-899X. Available from: https://doi.org/10.1088/1757-899X/337/1/012014

[39] BRANDIMARTE, P., ZOTTERI, G. *Introduction to distribution logistics*. John Wiley & Sons, 2007, ISBN 978-0-471-750-44-4.

[40] BISOGNO, S., CALABRESE, A., GASTALDI, M., GHIRON, N. L. Combining modelling and simulation approaches: how to measure performance of business processes. *Business Process Management Journal* [online]. 2016, **22**(1), p. 56-74. ISSN 1463-7154. Available from: https://doi.org/10.1108/BPMJ-02-2015-0021

[41] ALMEIDA, L.C., SALLES, S.A.F., CARVALHO, R.L., MORAIS, A.S.C., SILVA, S.V. S. BPMN e ferramentas da qualidade para melhoria de processos: um estudo de caso. GEPROS / BPMN and quality tools for process improvement: a case study (in Portuguese). *GEPROS. Gestao da Producao, Operacoes e Sistemas* [online]. 2019, **14**(4), p. 156-175. eISSN1984-2430. Available from: https://doi.org/10.15675/gepros.v14i4.2308

[42] VLASOV, A.I., GONOSHILOV, D. S. Simulation of manufacturing systems using BPMN visual tools. *Journal of Physics Conference Series* [online]. 2019, 1353:012043. ISSN 1742-6588, eISSN1742-6596. Available from: https://doi.org/10.1088/1742-6596/1353/1/012043

[43] MUTANOV, G. M., SERIKBEKULY, A. (2019). Fast reengineering for distribution logistic processes in small companies in Kazakhstan based on graphical and system dynamic model. 34th International Business Information Management Association Conference, IBIMA 2019 : proceedings. 2019. ISBN 978-0-9998551-3-3.

Appendix 1

The scheme of receiving and processing an order used by trading companies is available online at:
http://journals.uniza.sk/data/MutanovEtAl/Appendix1.vsdx

Appendix 2

The proposed scheme of the business process of receiving and processing orders is available online at:
http://journals.uniza.sk/data/MutanovEtAl/Appendix2.vsdx

Andrej Novak - Alena Novak Sedlackova - Anna Stelmach - Doris Novak

SAFETY IMPLICATIONS OF GNSS SIGNAL INTERFERENCE AT ZILINA AIRPORT

The paper deals with the issue of GNSS interference and its subsequent impact on airport approach procedures. It discusses the problem of GNSS signal interference and interference identification on a practical example of a small regional airport in Zilina, located close to the highway, through research aimed at its identification, a proposal for the location of a monitoring station and subsequent practical verification. The paper seeks to analyse and provide recommendations for enhancing safety and reliability in GNSS approaches. Given the need to develop the air transport, it is important to ensure the safety and continuity of service provision at small regional airports. The GNSS approach at airports with insufficiently equipped navigation infrastructure seems to be one of the most suitable. Introduction of the GNSS interference monitoring in the final instrument approach phase would increase the safety and reliability of the flight.

Keywords: *aircraft, GNSS, signal interference, safety*

1 Introduction

Satellite navigation systems have existed since the early 1960s when development of the first navigation satellites began. This resulted in the development of the Transit system in the US and development of the Cyclone navigation system and its civilian version of Cikada in the Soviet Union. Initially, those systems provided positioning in 2D space. The positional error of the first satellite navigation systems was approximately 500 m during the signal reception and had a relatively low time determination accuracy. Those navigation systems worked on the Doppler principle. In the US, after experience with the Doppler systems in the early 1970s, they decided to build a 3D navigation system with a precise time standard. This project has launched a new era of satellite navigation systems that are also suitable for use in civil aviation. The new navigation systems in the 1980s were the GPS NAVSTAR developed in the USA and the GLONASS system developed in the Soviet Union. Those two global satellite navigation systems set the foundation for the development and production of new satellite navigation systems Galileo (in the EU), BeiDou (in China) and Gagan (in India).

Development of the satellite systems and their implementation in aviation in the 1990s were also helped by the massive development of computer technology, the enhancement and implementation of new procedures in airspace management, but also problems with the navigation performance of existing terrestrial navigation systems. One of the first significant milestones in use of the satellite navigation systems was to switch off the deliberate misleading of the GPS L1 signal on May 1, 2000. The second significant milestone in development and use of the satellite navigation systems was introduction of enhanced satellite navigation systems, either based on a terrestrial augmentation called GBAS (Ground Based Augmentation System) or based on a satellite augmentation called SBAS (Satellite Based Augmentation System). Both variants of the satellite navigation system augmentation offer various aviation users a variety of options and benefits. Highlights include increased 3D positioning accuracy, continuity, integrity and safety.

The third milestone in development is integration of the new satellite navigation systems, Galileo and BeiDou, into air navigation and their application to aviation. This step will make it possible to switch from one L1C/A GPS signal to eight signals from four satellite navigation systems [1]. The US GPS system on L1C/A has so far provided the required Performance Based Navigation (PBN) in air navigation, aircraft positioning for Automatic Dependent Surveillance (ADS-B) and navigation support for safety systems (TAWS - Terrain Avoidance Warning Systems). In the future it is expected that in air or other types of navigation Dual Frequency Multi Constellation Services (DFMC) - combining two frequencies from different satellite navigation systems of American GPS, Russian GLONASS, European Galileo and Chinese BeiDou will be used to help civil aviation to improve navigation performance, continuity and integrity of the precision air navigation system. Therefore, DFMC GNSS in aviation means using a dual frequency signal from up to four GNSS sources. This will enable improvement of the Airborne Based Augmentation Systems (ABAS),

Andrej Novak[1,*], Alena Novak Sedlackova[1], Anna Stelmach[2], Doris Novak[3]
[1]University of Zilina, Slovakia
[2]Warsaw University of Technology, Poland
[3]University of Zagreb, Croatia
*E-mail of corresponding author: andrej.novak@fpedas.uniza.sk

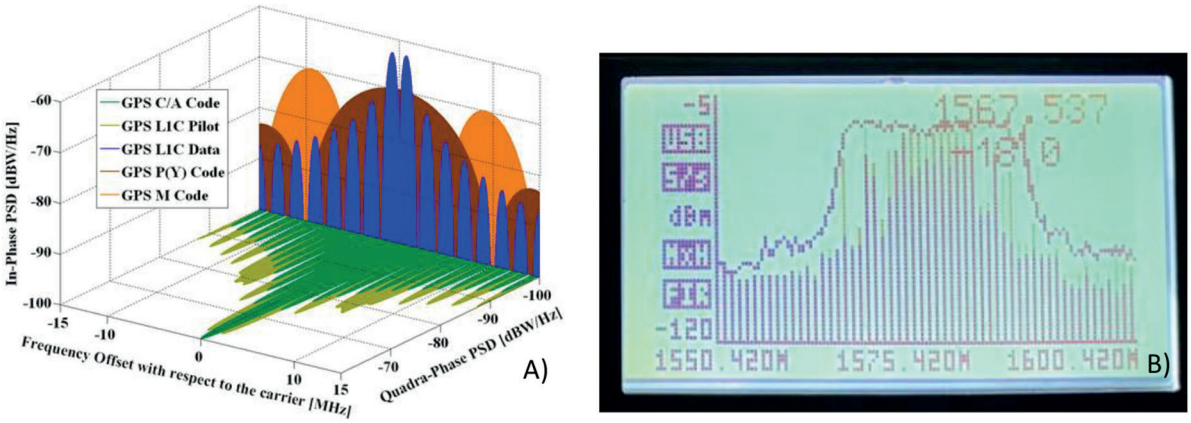


Figure 1 *Spectrum GPS L1 (C/A, L1C, P(Y) code, M-code) and spectrum of jammer with power -18dBm*

the Satellite Based Augmentation System (SBAS) and the Ground Based Augmentation System (GBAS). Introduction of the DFMC for use in civil aviation is expected in years 2025 to 2028. The idea of introducing the DFMC is not new and its use has been common in geodesy and cartography for more than 10 years, with the first devices with this functionality being marketed by Leica and Trimble.

The problem of the GNSS systems dependence on only one data source, the GPS L1C/A, is very restrictive in terms of safety and regularity of air transport. From the economic efficiency point of view, such a monopoly in the provision of navigation information at first glance may seem advantageous, but it has limits in terms of the need to maintain the large-scale ground and airborne infrastructure based on the conventional VOR/DME or DME/DME navigation. This GNSS monopoly built on GPS L1C/A has only a little resistance to accidental and deliberate frequency interference. In Europe, but also elsewhere in the world, there were accidental local interference of satellite navigation systems, which were losing their navigation performance, continuity of service and, last but not the least, location and time accuracy.

2 GNSS signal interference identification

Issue of the GNSS signal interference identification has been elaborated by several authors of publications. This issue can be divided into two partial problems. The first is the interference of the GNSS signal, which is undesirable, not only in the field of the air transport in real-time positioning, but in the intelligent transport systems, as well. The second partial problem is identification and location of GNSS signal interference. The non-public GNSS signals that are provided for military and special applications are usually encrypted. Their immunity to the interference signal is relatively greater due to use of the two or more carrier frequencies for the transmission of navigation data. However, in publicly available services, referred to as "open service" (OS), the signal is transmitted on one frequency, so the resilience of this system decreases significantly. Given that the satellite signal has a very low signal strength on the earth's surface, typically -120 to -130 dBW (the receiver is capable of processing up to -158.5 dBW) (Figure 1), the easiest attack on the GNSS signal is to generate interference or jamming. In the case of using meaconing or spoofing, this type of attack is much more demanding and therefore this type of interference requires relatively good technical equipment, provision of which is relatively expensive. Meaconing, sending a false signal with the aim of taking the target to another area, requires sophisticated know-how, as well as tens of thousands of euros of technical equipment (antenna system, GNSS signal generator, radar or LIDAR or camera system for detecting the location of interfered target).

The GNSS signal interference can be divided into two categories. The first category consists of interference that is not intentional and arises from various electrical devices operating at harmonic frequencies close to the GNSS systems. Be it various radio relay links, TV transmitters, radars, but also damaged base stations of mobile operators (BTS). The second group consists of interference or jamming caused by deliberate interference in the frequency spectrum of the GNSS signals. These are either intentionally coordinated attacks, or people using personal radio jammers, also known as PPD (Personal Privacy Device). A typical jamming case is the GNSS signal jamming at Newark Airport in New Jersey in 2009, where the GBAS was newly installed. [2-3]. This system showed short-term system integrity failures due to the GNSS signal interruptions. After two months of investigation by the FAA (United States' Federal Aviation Administration), it has been found that the interference was caused by a passing vehicle with a driver using a freely available PPD. There was no damage to property and lives during the incident, but this case showed how easily it is possible to disrupt such a sophisticated GBAS as long as it is based on receiving one GNSS (GPS L1) signal, see Figure 1 [4-5].

The interference itself can be defined based on the following characteristics:

- Depending on the type (sine wave interference, carrier wave - single tone, or interference by AM, FM, PM

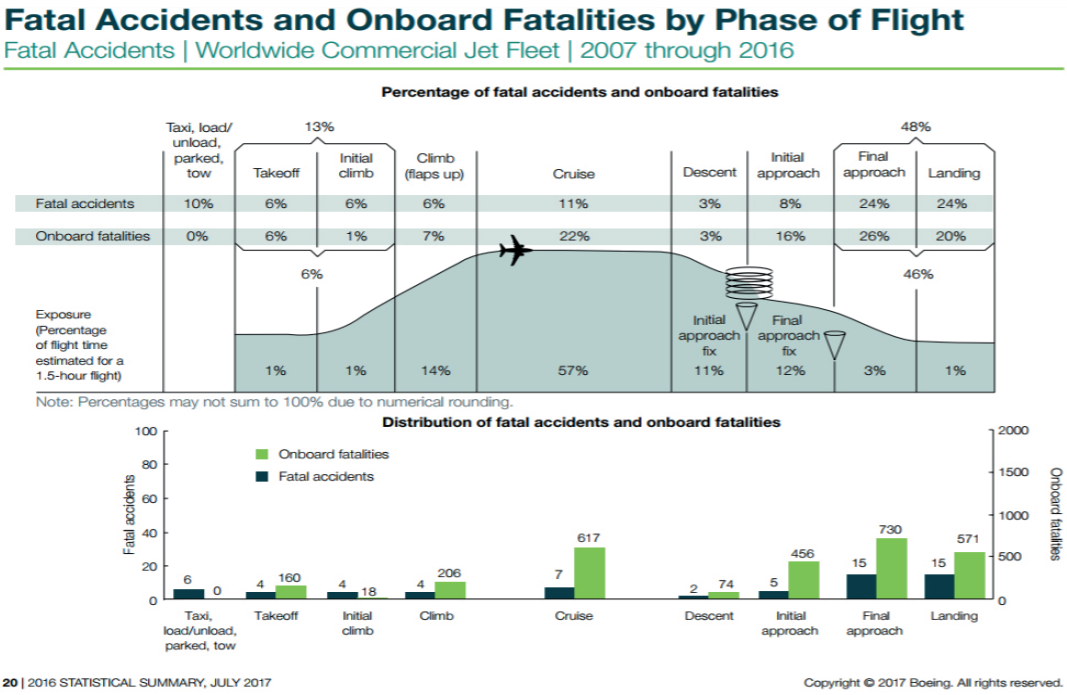


Figure 2 *Analysis of different phases of flight from the perspective of incidents on board aircraft [6]*

modulated signals that disturb the signal in a larger spectrum or noise interference - randomly generated signal (white or pink noise)).

- Relatively according to the mean frequency of interference to the position of our signal where the interference itself is located, one can talk about "out of band", "near the band" or "in band" interference.
- Signal interference bandwidth, where signal interferences can be divided into two subgroups of broadband or narrowband interference.
- Interference power, in which case it is the ratio of the carrier signal to the interference signal (interference) and Jammer to Signal J/S ratio.
- The time domain of interference, wherein the interference may be transmitted continuously or discreetly at time intervals or pulses. In the case of the pulse interference, it can be characterized by pulse width or number of pulses per second.

Interference on the L1 GPS frequency can manifest in the transmitted spectrum in different ways due to the fact that the signal is spread.

Given that the electromagnetic spectrum is limited, its emission must be regulated. Spectrum regulation is dealt with by the International Telecommunication Union (ITU), with the transmission of the GNSS frequencies being in the RNSS frequency bands of 1164 to 1215 MHz, 1240 to 1330 MHz and 5010 to 5030 MHz.

Potential GNSS signal interference is particularly problematic in the final approach phase of an aircraft. Because in this phase of flight, the greatest navigational performance of the aircraft is required. If one wants to establish a critical phase of flight, one must use the air traffic accident statistics. One such analysis was carried out by Boeing in 2017 and further elaborated by Geoffrey 2018 [6]. In this analysis authors were mainly interested in the flight phase with the highest number of incidents and accidents. The largest percentage is in the Final approach and Landing phase.

As is apparent from the analyses of Zizka (2019) and Kalasova (2015), it is precisely the areas for the use of precision approach that are intersected by transport infrastructure (highways or high-traffic roads) or in densely built areas (industrial zones, production plants and specialized businesses) that are the most problematic [7-8]. According to Curran (2017), accidental signal interference occurs right above zones that are industrially active or where high-intensity transport infrastructures exist [2]. A typical example is Frankfurt Airport, where the GNSS signal interference occurs mainly from cars located on the A3 motorway, which runs parallel to the runway and airport infrastructure for precision instrument approach GBAS. Based on the results of experiments using the method of analogy, it is decided to test this model for the Zilina airport and thus identify a critical approach point for it. Based on results of measurements (experiment) analysed by Novak 2018, it is possible to say, that from the personal jammers and car jammers point of view, the critical point is where the road infrastructure (highway) intersects the glide path for the instrument approach [5]. The height was calculated by applying a numerical method based on propagation of the signal above the terrain, assuming the on-board satellite unit has a conventional directional antenna with a signal pre-amplifier to receive the SBAS signal, in this case from the EGNOS satellite.

Based on practical measurements, which were carried out at the Zilina airport, it was found that areas where the

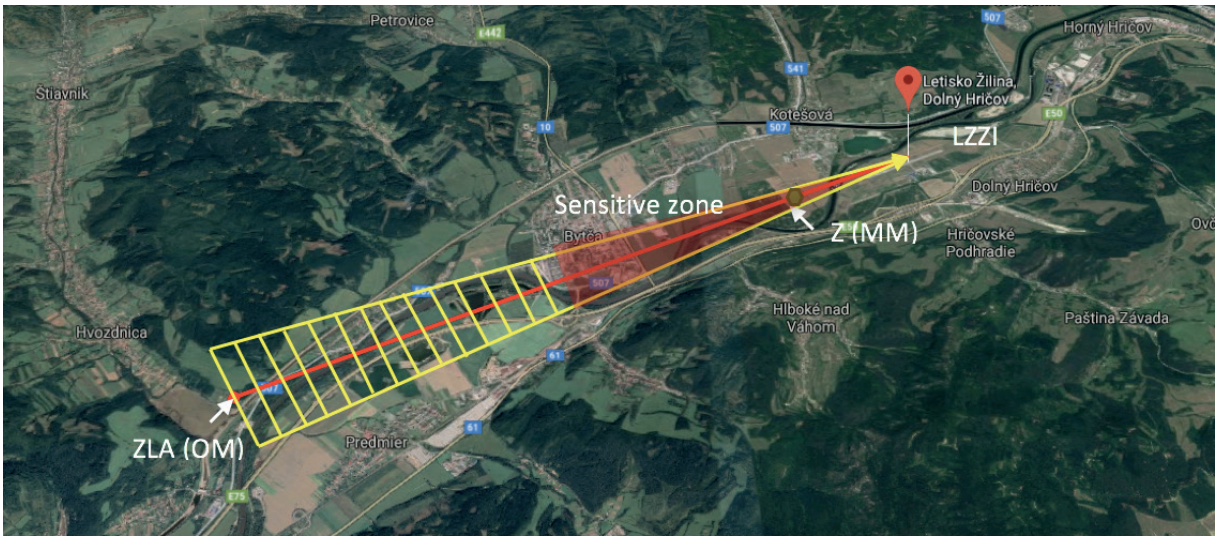


Figure 3 *LZZI approach glide path for runway 06*

aircraft is low above the terrain are critical [9]. Based on the above mentioned knowledge, it is possible to define a problematic place as a place or space where the aircraft drops below a critical height above the terrain (Figure 2). Figure 3 shows the critical phase of the flight in red and is defined around the Z (MM) point, as shown in Figure 3. The critical phase is from point ZLA (OM) to RWY06 LZZI, but the sensitive zone is defined around the Z (MM) ±1NM. The minimum safe altitude is defined in Aeronautical Information Publication of the Slovak Republic, part Aerodromes (AD 2 LZZI-7-3) for this situation.

3 Analysis of possible locations for detector placement

Based on statistical outputs of measurements, carried out by the Police of the Slovak Republic in cooperation with the Regulatory Authority for Electronic Communications and Postal Services from 2012 to 2017, it was found that occurrence of the GNSS interference on GPS L1 frequency is mainly on highway sections, expressways and 1st class roads, which are subject to a toll system that uses the GPS L1 [10]. Therefore, it is very important to analyse these types of roads and their parallelism with the approach axis for the LZZI 06 runway. The analysis of possible locations was based on terrain orography, electromagnetic wave propagation model, sensitive zone and glide axis profile for the 06 LZZI instrument runway.

For this analysis, the Zilina airport was chosen, since it by its nature represents a typical case of a regional airport located in rugged mountain terrain. Irregular air transport, charter flights as well as flight school training flights, are carried out at the airport and an introduction of an airline scheduled service is planned. At the same time, research flights of the University of Zilina in Zilina, focusing on safety, meteorological phenomena and earth exploration are carried out at the airport. During those flights, the GNSS signal interference has been registered several times through the flight laboratory, which has also been documented and described quote. This fact led to the need to build the ground-based interference detectors. When selecting a suitable location, it is necessary to analyse the potential sources of interference and their impact on the glide path of the instrument approach, as shown in Figure 4. In this case, points were identified (marked in Figure 4) on the existing transport infrastructure, as well as on the glide path, which are becoming critical for the safe approach execution. In terms of overall flight, this is the final approach phase defined by the ZLA (OM) point and the point of contact on the runway. This phase of the flight has to be divided into two further sub-sections, ZLA (OM) to Z (MM), which for the airport is defined as the final decision or Missed Approach Point (MAPt). The second part is from the Z (MM) point to the touchdown point on the runway. This division will allow to identify the risks that the GNSS signal interference could cause at this stage of the flight more accurately (Figure 3). Although there are several LPS SR or Eurocontrol studies that describe this risk, none of them explicitly anticipates deliberate interference and its increase in the period after the introduction of the LPV instrument approach, and even after subsequent calibration flights, the repeat studies are not considered, despite the fact that up to 48% of the total number of fatal incidents occurred during this phase of the flight [6, 11-12]. Based on the above analyses, findings and modelling of the situation, it was decided to identify a suitable location for the placement of the GNSS interference detector. The conditions that such a point should satisfy are: technical infrastructure (connection to the electricity grid and the possibility of data transmission) and section on the final approach track. This analysis therefore shows that such a point at the Zilina Airport should be in the centre of the descent plane for the 06 LZZI instrument runway. The point Z (MM) was chosen for this simulation, location of which is shown in Figure 3.

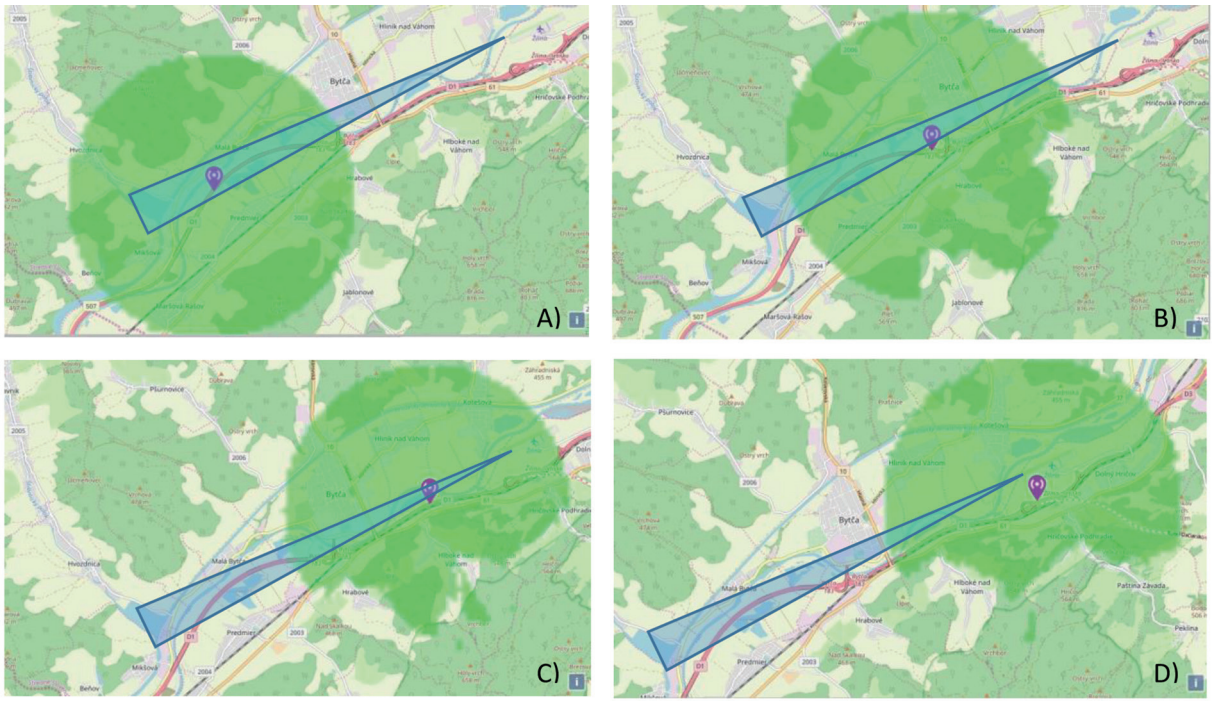
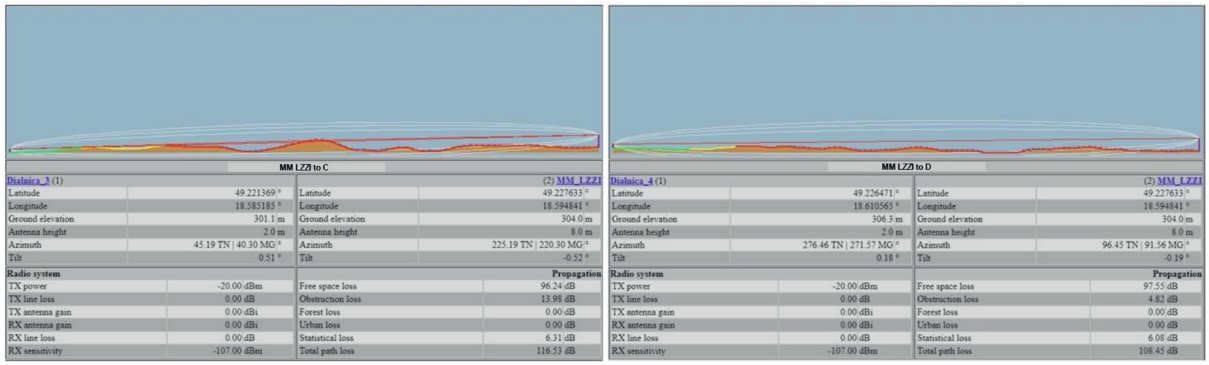


Figure 4 *Point analysis of disturbance signal propagation on the GPS L1 frequency from D1 highway for 06 LZZI instrument runway*



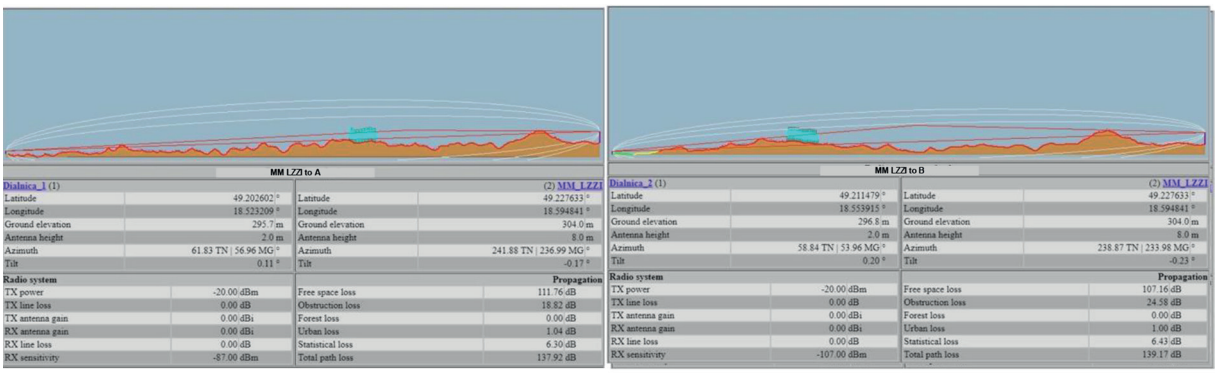
MM LZZI to C

Dialnica_3 (1)		(2) MM_LZZI	
Latitude	49.221369 °	Latitude	49.227633 °
Longitude	18.585185 °	Longitude	18.594841 °
Ground elevation	301.1 m	Ground elevation	304.0 m
Antenna height	2.0 m	Antenna height	8.0 m
Azimuth	45.19 TN | 40.30 MG °	Azimuth	225.19 TN | 220.30 MG °
Tilt	0.51 °	Tilt	-0.52 °
Radio system			**Propagation**
TX power	-20.00 dBm	Free space loss	96.24 dB
TX line loss	0.00 dB	Obstruction loss	13.98 dB
TX antenna gain	0.00 dBi	Forest loss	0.00 dB
RX antenna gain	0.00 dBi	Urban loss	0.00 dB
RX line loss	0.00 dB	Statistical loss	6.31 dB
RX sensitivity	-107.00 dBm	Total path loss	116.53 dB

MM LZZI to D

Dialnica_4 (1)		(2) MM_LZZI	
Latitude	49.226471 °	Latitude	49.227633 °
Longitude	18.610565 °	Longitude	18.594841 °
Ground elevation	306.3 m	Ground elevation	304.0 m
Antenna height	2.0 m	Antenna height	8.0 m
Azimuth	276.46 TN | 271.57 MG °	Azimuth	96.45 TN | 91.56 MG °
Tilt	0.18 °	Tilt	-0.19 °
Radio system			**Propagation**
TX power	-20.00 dBm	Free space loss	97.55 dB
TX line loss	0.00 dB	Obstruction loss	4.82 dB
TX antenna gain	0.00 dBi	Forest loss	0.00 dB
RX antenna gain	0.00 dBi	Urban loss	0.00 dB
RX line loss	0.00 dB	Statistical loss	6.08 dB
RX sensitivity	-107.00 dBm	Total path loss	108.45 dB

Figure 5 *Signal interference simulation output for two placement variants C and D*



MM LZZI to A

Dialnica_1 (1)		(2) MM_LZZI	
Latitude	49.202602 °	Latitude	49.227633 °
Longitude	18.523209 °	Longitude	18.594841 °
Ground elevation	295.7 m	Ground elevation	304.0 m
Antenna height	2.0 m	Antenna height	8.0 m
Azimuth	61.83 TN | 56.96 MG °	Azimuth	241.88 TN | 236.99 MG °
Tilt	0.11 °	Tilt	-0.17 °
Radio system			**Propagation**
TX power	-20.00 dBm	Free space loss	111.76 dB
TX line loss	0.00 dB	Obstruction loss	18.82 dB
TX antenna gain	0.00 dBi	Forest loss	0.00 dB
RX antenna gain	0.00 dBi	Urban loss	1.04 dB
RX line loss	0.00 dB	Statistical loss	6.30 dB
RX sensitivity	-87.00 dBm	Total path loss	137.92 dB

MM LZZI to B

Dialnica_2 (1)		(2) MM_LZZI	
Latitude	49.211479 °	Latitude	49.227633 °
Longitude	18.553915 °	Longitude	18.594841 °
Ground elevation	296.8 m	Ground elevation	304.0 m
Antenna height	2.0 m	Antenna height	8.0 m
Azimuth	58.84 TN | 53.96 MG °	Azimuth	238.87 TN | 233.98 MG °
Tilt	0.20 °	Tilt	-0.23 °
Radio system			**Propagation**
TX power	-20.00 dBm	Free space loss	107.16 dB
TX line loss	0.00 dB	Obstruction loss	24.58 dB
TX antenna gain	0.00 dBi	Forest loss	0.00 dB
RX antenna gain	0.00 dBi	Urban loss	1.00 dB
RX line loss	0.00 dB	Statistical loss	6.43 dB
RX sensitivity	-107.00 dBm	Total path loss	139.17 dB

Figure 6 *Signal interference simulation output for two placement variants A and B*

4 Simulation of interference in real operation

For the Zilina airport, the Z (MM) was chosen as the best point where the GNSS signal receiver in the simulation was placed, where the height of the antenna above the ground corresponds to the height of the existing mast 8m above the ground. To simulate an interfering signal, a car moving on a highway was chosen with an antenna height that interferes the signal 2m above the terrain, which represents normal delivery van or truck.

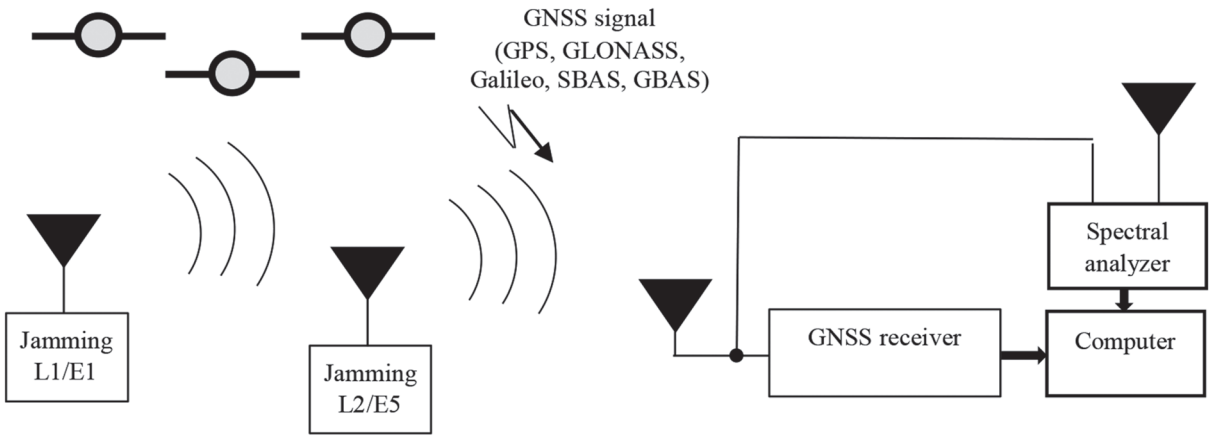


Figure 7 *Ground station interference measurement, block diagram rs)*

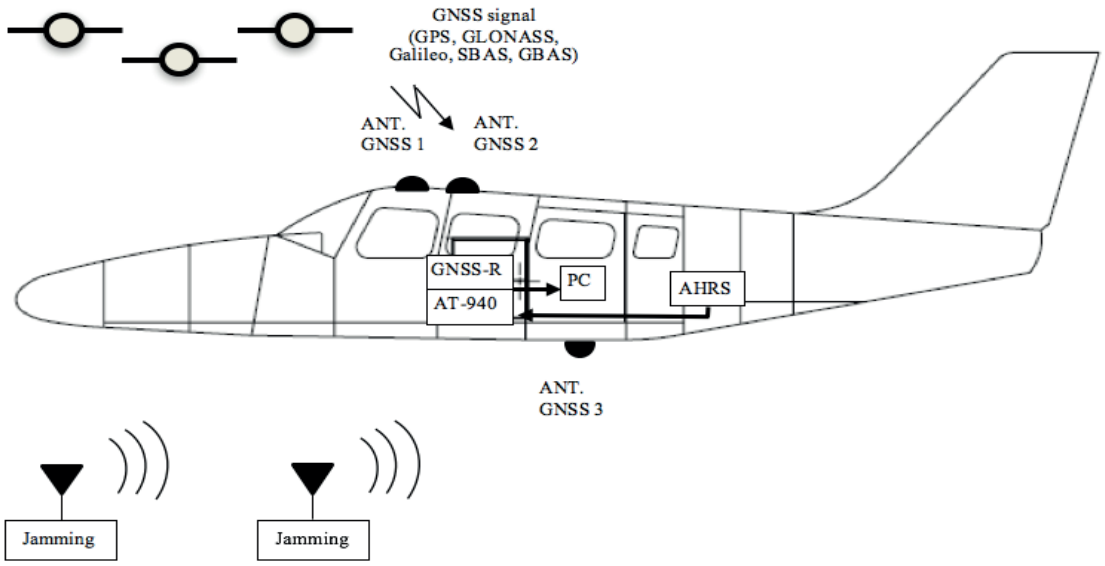


Figure 8 *Measurement of interference by flight laboratory, block diagram*

The jammer parameters were based on the average value of the available signal generators and measurements of commercially available jammers sold on the Internet. The power of the available jammers oscillates around -18 dBm ± 2 dBm. Antenna gain and attenuation was set to 0 dB to simplify the calculation, similarly the receiver sensitivity and antenna gain were set to - 107 dBm.

Subsequently, the propagation of the signal around selected points was calculated by means of the program for simulation of RF signal propagation and a model of point connection was created. Vertical models were created between the interference source and the interference detector (Figures 5 and 6). The calculation confirmed that the signal from the interference source could not be easily identified if the detection device was not properly positioned.

The simulation suggests that the GNSS interference detector is able to identify interference at points C and D and the detection of the interference signal will be in the critical phase of flight, which is the L (MM) up to the Touchdown point (Figure 5). From the points A and B simulation point of view it is obvious that the receiver antenna is low above the ground and therefore there is no direct connection (Figure 6). The interference may affect the first phase of ZLA (OM) approach to Z (MM), where the aircraft is relatively high above the ground, and therefore the GNSS antenna is isolated by the aircraft fuselage from a ground jammer whose antenna height is relatively low above the ground (2 m). Disturbance or failure may only occur if the aircraft falls below the minimum safe altitude above the ground. The minimum safe altitude is defined in Aeronautical Information Publication of the Slovak Republic, part Aerodromes (AD 2 LZZI-7-3) for this situation.

5 Practical verification

The procedure for practical verification of the simulation has been carried out by ground and subsequent aerial verification measurement. When the GNSS signal detector, which represents a GNSS receiver, a spectrum analyser, a computer and a suitable antenna system, was placed at the Z (MM) point (Figure 7). The antenna must

be at a minimum height of 8 m above the ground for detection of interference from ground-based transport infrastructure and the GNSS receiver antenna has to have the same directional characteristics as the aircraft antenna. The spectrum analyzer antenna should be omnidirectional with a pre-amplifier to be able to monitor the interference signal from a distance of at least 5 km, corresponding to the critical zone [13].

The flight measurement takes place along the glide path for the 06 LZZI instrument runway. Final approach from the ZLA (OM) point to the threshold point on the 06 LZZI runway is particularly important for our experiment. A block diagram of flight measurement using the AeroLab 1 flight laboratory is shown in Figure 8.

The flight and ground measurement output is compared and the ground system is calibrated so that the interference sensitivity threshold is set for the failure of the GNSS signal on board the aircraft. Because the threshold for the activation of the alarm in case of the interference is one most important parameter for safety. The value of the interference threshold must be minimum -14 dBm less to GNSS receiver sensitivity. This is important for the future operation of the monitoring equipment in order to differentiate the impact of the GNSS interference on the final phase of the final approach flight [7, 14-15]. This practical procedure described the future measuring test for validation.

6 Discussion

Safety, reliability and continuity are among the main priorities of aviation. It is important to realize that the ground infrastructure for provision of the radio navigation services must meet strict conditions in these three areas. The navigation devices operate in 24H continuous operation, their reliability level must reach a minimum of 98% during their planned lifetime. Where terrestrial systems, such as VOR/DME, ILS are devices that operate at dedicated frequencies and are protected, satellite navigation equipment was originally designed for military use and only partially for civilian use, but not with priority for aviation. It was only later released and approved by the International Civil Aviation Organization (ICAO) for worldwide use, provided that there must be a backup in the event of a system malfunction or failure. The problem of dependence of the GNSS systems on only one data source, GPS L1C/A, is very restrictive in terms of safety, reliability and continuity of air transport. From the economic efficiency point of view, such a monopoly in the provision of navigation information at first glance may seem advantageous, but has its limits in terms of the need to maintain large-scale terrestrial and on-board infrastructures based on conventional VOR/DME or DME/DME navigation. This GNSS monopoly built on GPS L1C/A has only little resistance to accidental and deliberate frequency interference.

Nevertheless, our simulation points out that from the economic efficiency point of view it is also possible to operate the GNSS approach systems with a high degree of safety and continuity in an environment with a high rate of interference and jamming. But only under condition of solution status monitoring and with a possibility of crew early warning for the loss of navigation performance incident. The authors of this article have shown that the current solution is not at a sufficient level of safety in terms of current knowledge of the issue and technical development of terrestrial and satellite navigation systems.

7 Conclusion

Introduction of the GNNS procedures for precision approach at airports in the Slovak Republic has raised several questions about the safety and reliability of the use of these procedures. Despite the established quality management systems and risk analysis, to date, not all the issues of safety, reliability and continuity of service provision at airports have been answered. Based on numerous reports on the GNSS interference and reduced navigation performance of GNSS GPS L1C/A receivers, we have decided to carry out an analysis of the location of the GNSS jamming detector for Zilina Airport (LZZI) and propose a practical verification of its location by an experiment. It should be noted that safety, reliability and continuity are the main priorities for the provision of services in air transport. Based on results of analyses, as well as the discussion itself, recommendations can be made to increase the safety and reliability of GNSS approaches.

Suggested recommendations are as follows. In the case of the aerodrome approval for approach using GNSS, it is required for the risk analysis to be only of limited validity period and to be carried out at regular intervals. The installation calibration measurement should establish an initial state and map out the possibilities and threats from the risk analysis of GNSS deployment.

Acknowledgments

This paper is published as one of the scientific outputs of the project: „New technologies and best practices in education in the Air Transport and Professional Pilots", KEGA 011ZU-4/2018.

References

[1] Eurocontrol helps build consensus on how dual frequency multi-constellation GNSS will be used in aviation - Eurocontrol [online] [Viewed 2019-10-07]. 2018. Available from: https://www.eurocontrol.int/news/moving-one-gps-signal-eight-signals-four-constellations

[2] CURRAN T. J., BAVARO, M., CLOSAS, P., NAVARRO, M. A Look at the threat of systematic jamming of GNSS. *Insied GNSS* [online]. 2017, September/October, p. 46-53. ISSN 1559-503X. Available from: http://insidegnss.com/auto/sepoct17-CURRAN.pdf

[3] CATLOS, M., KURDEL, P., NOVAK SEDLACKOVA, A., LABUN, J., CESKOVIC, M. Continual monitoring of precision of aerial transport objects. In: 13th International Scientific Conference - New Trends in Aviation Development NTAD 2018 : proceedings [online]. 2018. p. 76-81. Available from: https://doi.org/10.1109/NTAD.2018.8551683

[4] NOVAK, A., BUGAJ, M, KOVACIK, L., LUSIAK, T. GNSS signal interference and its impact on safety of air transport. In: Increasing Safety and Quality in Civil and Military Aviation : proceedings. Zilina: University of Zilina, 2019. ISSN 2644-495X, ISBN 978-80-554-1549-9, p. 81-84.

[5] NOVAK, A., HAVEL, K., BUGAJ, M. Measurement of GNSS signal interference by a flight laboratory. In: INAIR 2018, Aviation on the Growth Path : proceedings. Hainburg: Elsevier, 2018. ISSN 2352-1465, p. 271-278.

[6] GEOFFREY, T. Landing the most dangerous phase of flight, airlinereatings [online] [Viewed 2019-10-07]. Available from: https://www.airlineratings.com/news/passenger-news/landing-dangerous-phase-flight

[7] ZIZKA, J., NOVAK, A. Complex evaluation of risks associated with the use of new GNSS procedures In: *Studies*. Vol. 6. Zilina: Air Transport Department, Faculty of Operation and Economics of Transport and Communications, University of Zilina, 2019. ISBN 978-80-554-1564-2, p. 194-196.

[8] KALASOVA, A., FAITH, P., MIKULSKI, J. Telematics applications, an important basis for improving the road safety. In: Tools of Transport Telematics TST 2015 : proceedings [online]. Vol 531. Communications in Computer and Information Science. Cham: Springer, 2015. ISBN 978-3-319-24576-8, eISBN 978-3-319-24577-5. Available from: https://doi.org/10.1007/978-3-319-24577-5_29

[9] NOVAK, A., SKULTETY, F., KANDERA, B., LUSIAK, T. Measuring and testing area navigation procedures with GNSS. *MATEC Web of Conferences* [online]. 2018, **236**, 01004. eISSN 2261-236X. Available from: https://doi.org/10.1051/matecconf /201823601004

[10] BRIDA, P., MLYNKA, M., MACHAJ, J. How to solve GNSS problem in critical environment? In: IEEE 17th International Conference on Intelligent Engineering Systems INES 2013 : proceedings [online]. IEEE, 2013. ISSN 1543-9259, p. 27-31. Available from: https://doi.org/10.1109/INES.2013.6632835

[11] KRAUS, J. Determining acceptable level of safety of approach to landing. In: 20th International Conference Transport Means 2016 : proceedings. 2016. ISSN 1822-296X, eISSN 2351-7034, p. 230-235.

[12] ZAGORECKI, A., RISTVEJ, J., KLUPA, K. Analytics for protecting critical infrastructure. *Communications - Scientific Letters of the University of Zilina* [online]. 2015, **17**(1), p. 111-115. ISSN 1335-4205, eISSN 2585-7878. Available from: http://komunikacie.uniza.sk/index.php/communications/article/view/402

[13] BRIDA, P., MACHAJ, J., BENIKOVSKY, J. A modular localization system as a positioning service for road transport. *Sensors* [online]. 2014, **14**(11), p. 20274-20296. ISSN 1424-8220. Available from: https://doi.org/10.3390/s141120274

[14] BREZONAKOVA, A., SKVAREKOVA, I., PECHO, P., DAVIES, R., BUGAJ, M., KANDERA, B. The effects of back lit aircraft instrument displays on pilots fatigue and performance. *Transportation Research Procedia* [online]. 2019, **40**, p. 1273-1280. ISSN 2352-1465. Available from: https://doi.org/10.1016/j.trpro.2019.07.177

[15] ROSTAS, J., SKULTETY, F. Are today's pilots ready for full use of GNSS technologies? *Transportation Research Procedia* [online]. 2017, **28**, p. 217-225. ISSN 2352-1465. Available from: https://doi.org/10.1016/j.trpro.2017.12.188

Annex - Nomenclature

The abbreviation	The full name
ABAS	Airborne Based Augmentation Systems
ADS-B	Automatic Dependent Surveillance–Broadcast
DME	Distance Measuring Equipment
DFMC	Dual Frequency Multi Constellation Services
FAA	United States' Federal Aviation Administration
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GBAS	Ground Based Augmentation System
ITU	International Telecommunication Union
ICAO	International Civil Aviation Organization
ILS	Instrument Landing System
LPV	Localizer Performance with Vertical guidance
MAPt	Missed Approach Point
NAVSTAR	Navigation Signal for Timing and Ranging
PBN	Performance Based Navigation
RNSS	Radionavigation Satellite Service
RF	Radio Freqency
SBAS	Satellite Based Augmentation System
TAWS	Terrain Avoidance Warning Systems
VOR	VHF Omni-Directional Range

Dhani Setyawan

DECOMPOSING THE INFLUENCING FACTORS OF ENERGY INTENSITY IN THE PASSENGER TRANSPORTATION SECTOR IN INDONESIA

Indonesia's transport sector has experienced rapid growth that has caused excessive fossil fuel energy consumption. Over 2000 to 2016 total final energy consumption in Indonesia's transport sector has grown by 10% per annum so that transport now provides a large and rapidly growing component of total energy use. This study analyzes the specific characteristics of energy intensity in the transportation sector in Indonesia from 2000 to 2016 by employing a multiplicative Log Mean Divisia Index-II. The passenger transport sector in Indonesia, including the four modes of road, rail, water and air is examined in this study. Overall, the decline in energy intensity in passenger transport is attributed to the intensity effect. In passenger transport, the improvement of intensity effect was found to have significantly reduced the overall aggregate energy intensity, while the change in structural effect was found to have a relatively small reduction in the aggregate energy intensity.

Keywords: *energy efficiency, passenger transportation, energy intensity, Indonesia*

1 Introduction

Indonesia's transportation energy consumption is rapidly increasing, primarily due to rising economic activity and population growth. As an emerging economy and the fourth most populous country in the world [1], Indonesia's economic growth has had around 6% growth every year since 2010, leading to an increase in the mobility of the middle class [2]. Indonesia's transportation sector has gone through the rapid development, causing a significant use of fossil fuel energy consumption. This sector uses more than 60% of Indonesian's total oil use, approximately 70% of which is consumed in road transportation [2].

In the last decade, Indonesia had experienced several increases in fuel prices[2] (see Table 1). Subsidies for fossil fuels were increasing as a portion of the national budget, although the recent price increase was the lowest in the world, particularly for a net importing country [3]. Starting from 2005, the Indonesian Government had cut subsidies for energy and increased fossil fuel prices more than threefold. The purpose of this energy price reform was not only to limit the difference between international and domestic prices but also to bring decrease the burden on the state budget; as the budget for fuel subsidies accounted a substantial percentage in the national state budget [4].

Some studies [7-8] show that an increase in energy price can significantly improve energy efficiency. Therefore, the energy price instrument is one of the essential tools for energy reform and subsidies for energy is a key determinant of energy prices. The most common definition of an energy subsidy is a payment from government to consumer or producer in order to control energy prices [9-10]. In Indonesia, energy subsidies are mainly used by the government to control the energy price lower rather than the economic production cost.

Since the transportation sector in Indonesia is very complex and the statistical energy data in the transportation sector is limited, this study only investigates Indonesia's passenger transportation for a limited period. This study examines all the four modes of the transportation system in Indonesia: road, rail, water and air. Those modes of transport are used to measure the level of energy intensity in the passenger transport sector. This study is crucial for Indonesia's government in order to develop policies to improve energy efficiency, as well as to investigate the driving factors that affect the changes in transportation energy usage.

2 Literature review

The decomposition indices method has been extensively employed in measuring the driving factors behind the changes in energy used in transport. Authors of [11] investigated changes in the structure of passenger transport energy consumption in eight OECD countries from 1970 to 1987 by employing Laspeyres Divisia Indices.

[1] Thus far, the energy prices including electricity, natural gas and fossil fuel have been fully regulated by the government. The Indonesian government oversees the price of fossil fuel products with adjusting periodically following a formula in which the international price plays an important variable.

Dhani Setyawan
Ministry of Finance of the Republic of Indonesia, DKI Jakarta, Indonesia
E-mail of corresponding author: dhanisetyawan83@gmail.com

Table 1 *Fossil fuel subsidy reforms since 2005*

Year	Fuel Type	Fuel price policy reform
2005	Diesel and gasoline	Manufacturing industries are no longer able to get subsidized diesel. In March, the price increased by 29% and further increased in October by 114%.
2006	LPG	LPG price increase targeted to manufacturing industries
2007	Kerosene and LPG	In order to encourage LPG use, the government introduce the kerosene to LPG conversion program
2009	Diesel and gasoline	In January, Diesel and gasoline decreased by 7% and 11%, respectively.
2013	Diesel and gasoline	Both diesel and gasoline are increased by around 40%
2013	Electricity	Electricity base tariff increased by 15%
2014	Diesel and gasoline	Both Diesel and gasoline are increased by 36% and 31%, respectively
2015	Diesel and gasoline	Gasoline subsidies are removed and diesel subsidies are reduced by Rp 1,000/ litre.

Rp = Rupiah, Kg = kilogram, LPG = liquefied petroleum gas.
Source: [5-6].

Their study observed that shifting to more energy intensive transport mode and increasing passenger-kilometres has become the main reason for the increase in energy use.

As the economic theory forecasts that a higher price will affect in decreasing consumption. Thus, an increase in fuel price is considered in this paper as a factor affecting the changes in energy intensity. In [12-13] is argued that reducing the energy price will decrease energy intensity and lead to an improvement in energy efficiency. Additionally, by employing firm level data, in [14] is demonstrated that a rise in different type of energy products ameliorate energy efficiency. Furthermore, a strong evidence of energy savings, resulting from a higher energy price in China's paper industry, is revealed in [15], as well. Increases in the energy price result in decreasing energy intensity by means of efficient usage and structural adjustment, [16-17].

Nevertheless, there are contrasting points of view regarding the influence of price on energy intensity. By employing provincial level data, in [17] is suggested that energy price had a lesser effect on energy efficiency. Further, authors of [18] investigated at China's provincial level and revealed that the role of energy price to energy intensity is weak compared to other factors.

Other research also reveals some drawbacks of high subsidies that lead to inefficiency and failure in providing an affordable energy price for the poor. In [19-21] is observed that the high subsidized energy prices are more benefitting the non-poor households rather than the poor. Further, author of [22] argued the needs of reducing the cross-subsidies of electricity rates in India that less are optimal for several consumer groups. With regard to the high subsidies of energy in Indonesia in the previous few decades, the influence of energy subsidy to the overall trend of energy intensity in the transport sector is essential as one of the factors that will be measured in this study.

Besides those studies, other researches also relate the drawbacks of transportation mobility to the environment. In [23] are investigated changes of the Green House Gases (GHG) level from the electric mobility implementation and found varies impact in different countries. Furthermore, the impact of the road traffic on the environment is observed in [24] and importance of information and education in reducing the negative impacts of transport is emphasized. Authors of [25] have focused on road freight transport and the GHG emissions and concluded that biofuels are more favourable compare to fossil fuels. Another study from [26] also investigated the advantages and disadvantages of electromobility in the freight transport and suggested the importance of transport policy implementation in EU countries.

3 Methodology

This study investigates the driving forces of and examines the transportation energy consumption trend for, Indonesia by applying the Log Mean Divisia Index (LMDI) method for period 2000 to 2016. The transportation energy consumption is classified as passenger transportation, covering the four modes: road, water, rail and air. To determine changes in energy intensity trends, the following approach is employed [27]:

$$I_t = \frac{E_t}{Y_t} = \sum_{k=1}^{n} \frac{Y_{kt}}{Y_t} \frac{E_{kt}}{Y_{kt}} = \sum_{k=1}^{n} S_{kt} I_{kt}\,, \tag{1}$$

where:
I_t is the aggregate energy intensity at time t,
E_t is the energy consumption in all transport sectors at time t,
Y_t is the economic activity of all transport sectors at time t,
Y_{kt} is the economic activity in sector k at time t,
E_{kt} is the energy consumption in sector k at time t,
I_{kt} is the energy intensity of sector k at time t,
S_{kt} is the share of sector k in economic value of all transport sectors at time t.

Using multiplicative decomposition, the relation of the two time periods are described as below:

$$D_{Tot,T} = \frac{T_T}{I_0} = D_{Int,T} \times D_{Str,T}\,, \tag{2}$$

Table 2 *Variables in the transport sector*

Subsector	Mode of transport	Structural factors	Intensity factors
Passenger	Road Rail Water Air	Share of total passenger kilometre	Energy per passenger kilometre[2]

where:
$D_{Str,T}$ is the structural effect at time T (an index that determines the effect of the structural shift),
$D_{Int,T}$ is the intensity effect at time T (an index that determines the changes in sectoral energy intensity effect), which are computed as:

$$D_{Str} = \exp\left\{\sum_{k=1}^{N} \frac{L\frac{E_{k,T}E_{k,O}}{E_T E_O}}{\sum_k L\frac{E_{k,T}E_{k,O}}{E_T E_O}} \ln\left(\frac{S_{k,T}}{S_{k,O}}\right)\right\}, \quad (3)$$

and

$$D_{Int} = \exp\left\{\sum_{k=1}^{N} \frac{L\frac{E_{k,T}E_{k,O}}{E_T E_O}}{\sum_k L\frac{E_{k,T}E_{k,O}}{E_T E_O}} \ln\left(\frac{I_{k,T}}{I_{k,O}}\right)\right\}. \quad (4)$$

Energy intensity for passenger transport was computed separately in this study, to compare the energy efficiency improvement in each type of transport mode. The structural and intensity effects are the two driving forces that are attributed for the changes in aggregate energy in transportation. The energy intensity effect is a ratio that relates to energy consumption and turnover, which represents the energy efficiency of transportation activity. This effect is predicted to decrease over time, owing to development of more energy efficient technologies. In addition, the structural effect captures the change of turnover for each transport mode. This effect measures changes of energy use in due to the changes in modes of transports share of the economy. All the things being equal, the changes in the level of transportation turnover modal shares directly affect the level of transportation energy consumption modal share.

4 Data

Due to the lack of energy statistical data in the transportation sector in Indonesia, this study only takes into account the four modes of transport: road, water, rail and air. The data for this study comes from a peer-reviewed database of transport data: The Transport Databank [28]. This data is a collaboration of the Asian Development Bank (ADB), the Clean Air Asia's Transport Research Laboratory, University of California Davis' Institute of Transportation Studies and the Partnership of Sustainable, Low Carbon Transport (SLoCaT). The collaboration published a database on transport with a focus on Asia and the Pacific. The historical data was constructed from national surveys and international statistics [28]. Table 2 below describes the variables in the passenger transport sector in this study.

In terms of moving passengers, this study measured activity by the passenger-kilometre (PKM). The PKM measures efficiency of the passenger transport by calculating how much energy is needed to move one passenger per kilometre. The energy consumption for passenger transport is measured in Kilo Ton of Oil Equivalent (KTOE). Changes in energy consumption are decomposed in terms of the structural effect and intensity effect using the relative contribution of the passenger tasks. This study separates the analysis of passenger and freight sectors differently as the underlying economic factors shaping for both sectors differed [29].

5 Analysis of energy consumption in the transportation sector

As increasing energy consumption aligns with economic growth, which is defined as the GDP per capita [30], it is worth comparing the transport energy use and GDP. Figure 1 shows the increasing trend both in GDP and transport energy consumption in Indonesia from 2000 to 2016. Thus, in general, as the GDP and energy use increase, it can be assumed that it will be aligned with increasing mobility and improved standard of living. Turnover in the passenger transport is influenced by several factors, for instance, population growth, urbanization and changes in income, where these factors are expected to change the level of energy consumption in the transport sector [31-32].

While the value added data are available for total transport, the data for road passenger transport is not available separately. Given that this variable is important (each providing about half of energy use in transport by 2015) but also so distinctive, this study analyzes it separately, using dominators for the energy intensity measures (passenger kilometres), for the period 2000 to 2016.

5.1 The passenger transport

In passenger transport, the aggregate turnover and aggregate energy consumption have grown approximately four-fold from 2000 to 2016, generated substantially by an increase in the distance travelled per passenger [33]. The

[2] A passenger kilometre is defined as one passenger being moved one kilometre.

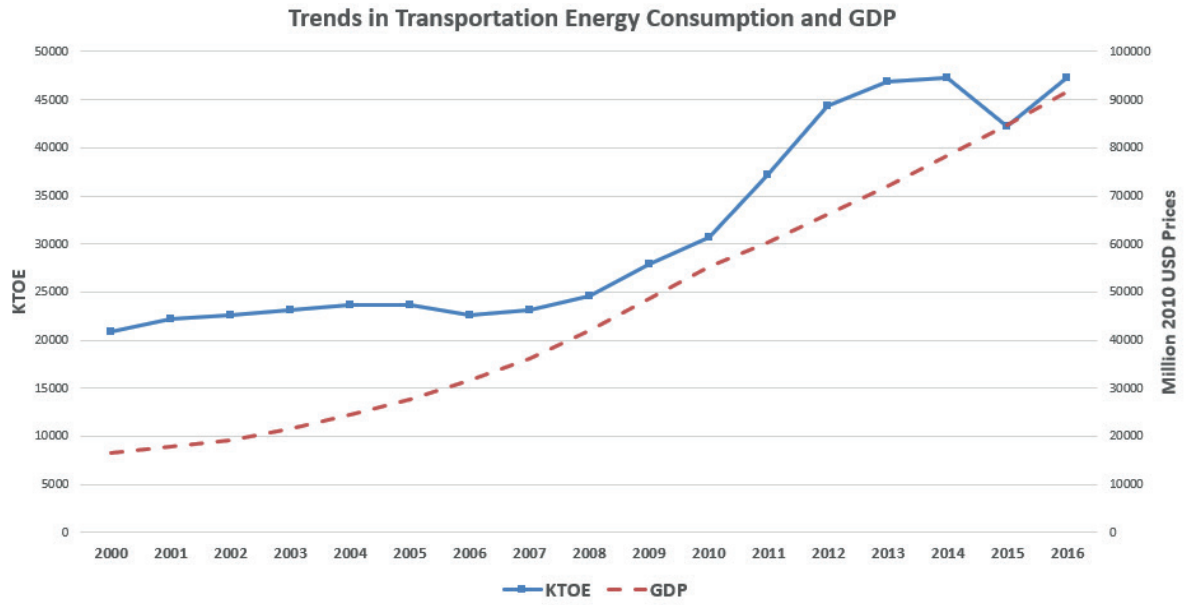


Figure 1 *Trends in the transportation energy consumption and GDP*

Table 3 *Share and Growth of Turnover (Passenger Transport - PKM)*

Sectors	Annual Growth Rate of Passenger Turnover (%)					Share of Total (%)				
	00-05	05-08	08-14	14-16	00-16	00	05	08	14	16
Road	11.6	21.6	5.4	4.8	10.1	94.6	95.4	96.4	95.7	95.6
Rail	-5.7	7.7	-0.3	1.1	-0.4	2.7	1.2	0.8	0.6	0.5
Water	0.5	8.3	8.1	5.8	5.4	0.5	0.3	0.2	0.2	0.3
Air	19.5	13.1	11.2	5.4	13.3	2.2	3.2	2.6	3.5	3.6
Aggregate	11.4	21.2	5.5	4.8	10.1					

passenger transport turnover increased from 710 billion PKM in 2000 to 3,286 billion PKM in 2016, with an average annual growth rate of 10.1%. For the same period, the total energy consumption in the passenger transport increased from 12,957 KTOE in 2000 to 54,384 KTOE in 2016 with an average annual growth rate of 9.4%. It is apparent that the increasing energy consumption growth and its turnover tracked each other very closely, with average yearly growth rates of 9.38% and 10.05%, respectively. This growth illustrates that Indonesia's transportation sector experienced an increasing energy usage and steady development, driven by many forces, including economic growth and improved standard of living.

The largest turnover share in the passenger transportation sector came from the road transport with approximately around 95.6% of the aggregate turnover in the passenger transportation sector in 2016 (see Table 3). Following the road transport, the air, rail and water transport accounted for around 3.6%, 0.5%, and 0.3%, respectively. It is not surprising that the road transport has the largest share given the vast development of Indonesia's highway networks over the last decade.

The road transport has quite a steady turnover share from around 94.6% in 2000 to about 95.6% in 2016. The turnover share of the rail and water modes decreased quite significantly during the study period from around 2.7% and 0.5% in 2000 to 0.5% and 0.3% in 2016, respectively. The turnover in the air transport increased from around 2.2% in 2000 to about 3.6% in 2016. The change in the air travel activity suggests that travellers are increasingly choosing to fly rather than to use other modes of transport. Domestic aviation was the fastest growing mode of the passenger transport, increasing by around 19% per year from 2000 to 2005 and approximately 13% annually from 2000 to 2016. Activity in other modes of the passenger transport also expanded, except for the rail transport mode that had the lowest growth over the period. Indeed, the rail transport significantly slowed its growth during the study period.

For a closer look at the structural proportion of energy consumption in the passenger transport, Table 4 illustrates the energy consumption of passenger transportation from 2000 to 2016. The road transport was the highest energy-consuming sector, responsible for around 91.8% of that consumed by the total passenger transport in 2000 but decreasing to around 88% in 2016. The share of the water transport and railway energy consumption decreased from 0.9% in 2000 to 0.5% and 0.2% in 2016, respectively, while the share of civil aviation almost doubled from approximately 6.4% in 2000 to 11.4% in 2016.

The primary reason for changes in the transport sector energy consumption and turnover is the shift in use of different modes of transport. This shift occurred as

Table 4 *Share and Growth of Energy Consumption*

Sectors	Annual Growth Rate of Passenger Energy Consumption (%)					Share of Total (%)				
	00-05	05-08	08-14	14-16	00-16	00	05	08	14	16
Road	10.3	19.2	4.9	4.7	9.1	91.8	89.7	91.1	88.1	88.0
Rail	-6.3	7.0	0.0	2.4	-0.5	0.9	0.4	0.3	0.2	0.2
Water	-0.9	8.4	8.3	9.3	5.4	0.9	0.5	0.4	0.5	0.5
Air	19. 5	13.3	11.1	5.4	13.3	6.4	9.4	8.2	11.2	11.4
Aggregate	10.8	18.6	5.5	4.8	9.4					

Table 5 *Energy intensity in passenger transport sector*

Sectors	Passenger Energy Intensity (KTOE/ Billion Passenger Kilometre)					Energy Intensity Changes (%)
	00	05	08	14	16	00 - 16
Road	17.7	16.7	15.7	15.2	15.2	-14.0
Rail	6.0	5.9	5.7	5.8	6.0	-1.0
Water	31.9	29.6	29.7	30.0	31.9	-0.2
Air	52.9	52.9	53.1	52.9	52.9	-0.1
Aggregate	18.2	17.7	16.6	16.5	16.6	-9.3

changes from less intensive energy consumption mode, for instance, the railway transport, to more intensive energy consumption modes the road and air transport.

Table 5 shows the energy intensity of each transport mode in 2000, 2005, 2008, 2014 and 2016. From 2000 to 2016, the aggregate energy intensity in the passenger transport sector decreased by around 9.3%. The declining trend in this period was mainly driven by the declining trend of energy intensity in the road transport for about 14%, followed by the rail, water and air transport sector for approximately 1%, 0.2% and 0.1%, respectively.

Overall, the energy intensity of the passenger transport sector declined steadily over the period from 2000 to 2016. In Table 5 can be seen that the least efficient transport mode or the highest energy intensity in the passenger transport is civil aviation, followed by the water transport with road and rail transport in the third and fourth place. The aviation sector is the least efficient mode, as it requires around 53 KTOE per billion passenger kilometres, accounting to nearly three times more than that by the road transport and almost nine times with respect to the railway. The most efficient transport mode in passenger transport is railway, which requires only around six KTOE per billion passenger kilometres.

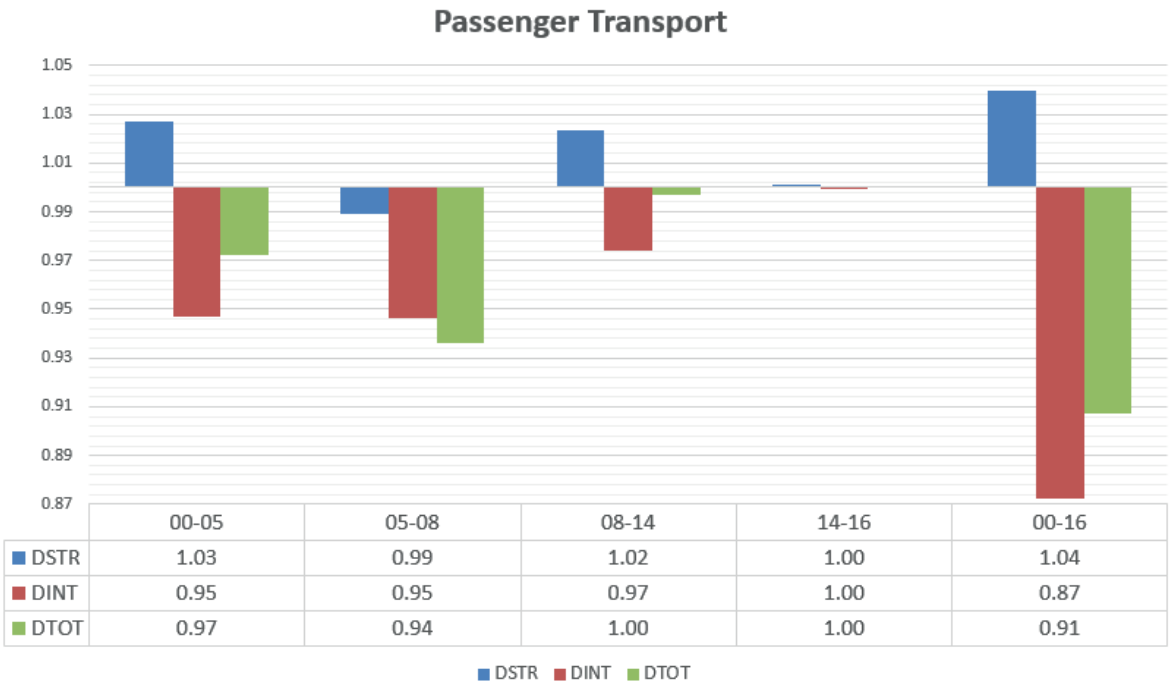
6 Decomposition analysis

The final energy use in the passenger transport sector rose during the study period. From 2000 to 2016, the energy intensity in the passenger transport declined from approximately 18.2 KTOE/billion tonne PKM in 2000 to 16.5 KTOE/billion PKM in 2016 (decreasing around 9.3%). By employing the LMDI method, this study estimates the driving forces of the passenger transport-related energy consumption in Indonesia, including the structural effect (DSTR) and intensity effect (DINT). As the energy source for the transportation sector mainly comes from oil and oil products, the decomposition analysis is divided into several sub-periods following the rising of the oil price in Indonesia. During the study period, the fuel price had been increased periodically. Diesel price was increased twofold in 2005, while in 2008 both diesel and petrol were raised[3] by one-third [4]. In 2013, the Indonesian government further increased diesel and petrol prices by 22% and 44%, respectively. Further, in November 2014, petrol and diesel were again increased by 31% and 36%, respectively.

Results of decomposition for the passenger transport in Indonesia from 2000 to 2016 are presented in graphical form in Figure 2.

The decomposition results in the passenger transport in Figure 2 show that the transportation energy intensity effect (DINT) significantly contributed directly to reduce the aggregate energy intensity (DTOT), while the role of the structure effect (DSTR) is found to be relatively small in most sub-periods. The energy intensity effect contributes strongly in decreasing the total energy intensity in the passenger transport. From 2000 to 2016, this effect induced a decrease in the passenger transport energy intensity for around 13%, while the structural effect increased the energy intensity for only 4%. The structural effect has had a negative effect on energy intensity in the passenger transport sector over this period. Overall, between 2000 to 2016, the aggregate energy intensity (DTOT) in the passenger transport decreased by around 9% as compared to its base level in 2000.

[3] In 2005, premium gasoline has been raised from Rp 1800 to Rp 4500 /litre and kerosene from Rp 700 to Rp 2000 /litre. Further, in 2008, premium gasoline has been raised from Rp 4500 to Rp 6000 /litre and kerosene from Rp 2000 to 2500 /litre.



	00-05	05-08	08-14	14-16	00-16
DSTR	1.03	0.99	1.02	1.00	1.04
DINT	0.95	0.95	0.97	1.00	0.87
DTOT	0.97	0.94	1.00	1.00	0.91

Figure 2 *Decomposition result of the passenger transport*

From 2005 to 2008, the structural effect contributed an 1% reduction in the energy intensity, while the intensity effect decreased it by around 5%. Overall, the aggregate energy intensity in the passenger transport decreased by around 6% during this period. In the period of 2008 to 2014, the aggregate energy intensity in the passenger transport was stable, due to the increase in the structural effect that was offset by the intensity effect.

The graph reveals that the energy efficiency in the passenger transport improved over the period from 2000 to 2016. This means that improving the transport intensity effect can significantly influence energy savings. This effect may relate to policy measures that have been enacted in Indonesia such as developing new green vehicle technology, improvement in fuel quality, advances in the transportation system and promoting energy alternatives, while, on the other hand, the increases in the structural effect potentially due to modal shifting, from less energy intensive mode, like rail transport, to more energy intensive modes such as the air and road transport. As noted earlier, the road transport sector has been a prominent transportation mode.

7 Benchmarking Indonesia to other ASEAN countries

This study provides an analysis of Indonesia's energy use compared to other selected ASEAN countries from 2000 to 2016. Due to a limit in the available data, in this study, the group of six ASEAN countries includes Indonesia, Malaysia, the Philippines, Thailand, Cambodia and Myanmar.

Figure 3 illustrates the passenger turnover share for the six selected ASEAN countries. The road transport has the largest share in the passenger transportation over the study period in all of the six selected ASEAN countries. From 2000 to 2016 this mode of transport contributes for more than 70% of the share for Malaysia, Thailand, the Philippines, Myanmar and Cambodia, where in Indonesia this mode contributed for more than 90%. The second highest turnover contributor came from the air transport. In Cambodia, air transport almost doubled from 14.7% in 2000 to around 28.6% in 2016, where the increase of aviation was followed with a decrease of the road transport. Another significant increase in the air transport can also be seen in Malaysia, where this sector contributed for around 26.3% in 2016. The rail transport has quite a significant share in Myanmar for around 16.5% in 2000 but fell to around 3.4% in 2016. The water transport in the total turnover was found the lowest in all of the six ASEAN countries for less than 1%, although in Cambodia it increased to around 4.4% in 2016.

Similarly, to the share of turnover, the energy consumption in the road transport mode also played a key role in the overall energy consumption in the passenger transport from 2000 to 2016 (Figure 4).

This mode of transport consumed more than 70% in all the six ASEAN countries, except Cambodia, where its energy consumption share decreased for almost a half from 63.9% in 2000 to 35.9% in 2016. The share of energy consumption in air transport increased quite significantly in Malaysia, Thailand, Philippines, Cambodia and Indonesia which increased from 27.4%, 20.3%, 15.1%, 36.1%, 6.4% to around 35.7%, 28.7%, 30.6%, 58.4%, 11.4%, respectively. Similar to the share of turnover, the share of energy consumption in railways and waterways were found to be insignificant in all of

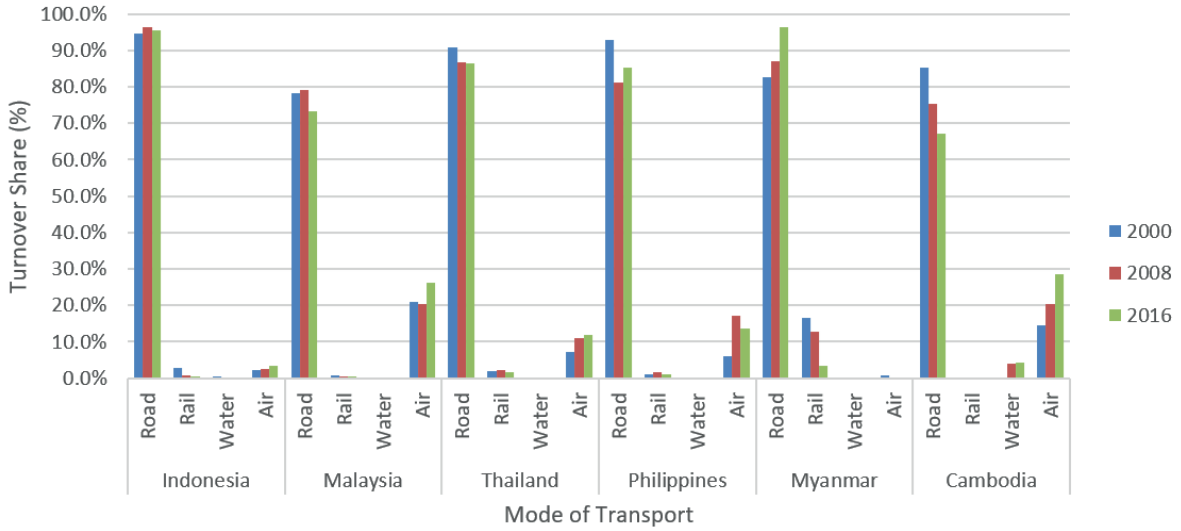


Figure 3 *The passenger turnover share in ASEAN countries*

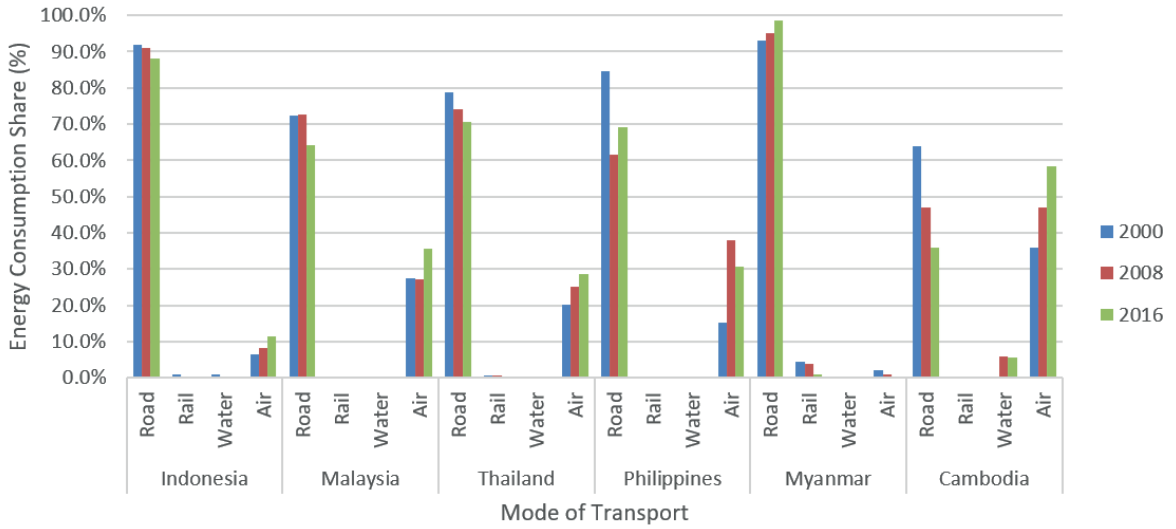


Figure 4 *The passenger energy consumption share in ASEAN countries*

the six ASEAN countries, which only covered less than 6% during the study period.

8 Conclusions

Regarding the turnover and energy use, the passenger transport in Indonesia experienced a substantial growth. The main factors affecting this trend were changes in the energy intensity, transport choices and turnover growth. For the passenger transport, the highest energy consuming mode and the major transport mode was the road transport. In spite of being a country with many archipelagos, the road transport contributed the largest turnover in passenger transport, whereas air, waterways and rail transport did not play a significant role in the passenger transportation in Indonesia. Over the study period from 2000 to 2016, the road transport accounted for the largest share, with more than 90% of the share in terms of passenger-kilometres travelled.

In the passenger transport, the structural effect (DSTR) had an increasing impact on the aggregate energy intensity in most years, although its magnitude was insignificant compared to the decreasing intensity effect. The increasing structural effect was associated with the modal shifting, from the less intensive energy modes, like railway, to more energy consumption modes, like road and aviation. In the passenger transport (see Table 3), the share of railways to total turnover dropped from 2.7% in 2000 to 0.5% in 2016; the air transport increased from 2.2% in 2000 to 3.6% in 2016; while the road transport increased only slightly by around 1% in 2016 compared to the base year. Based on [33], the increasing demand of the air transport services in Indonesia recently has been encouraged by the substantial development of airways transport infrastructure, specifically airports and the growing of the low-cost airlines. Between

2015 and 2016, the total number of airline passengers grew 10.5%, to 95.2 million [33]. In line with the increasing turnover share in the aviation, the ratio of the air transport energy consumption to other modes also increased from around 6.4% in 2000 to approximately 11.4% in 2016. Moreover, the energy consumption in the road transport decreased from 91.8% in 2000 to 88% in 2016; while the rail and water transport also had a slight decrease in energy consumption (see Table 4). The road transport has been the prominent transportation mode, which explains more than 90% of the total turnover and total energy consumption during the study period.

Over the period 2000 to 2016, the energy intensity of the passenger transport sector declined steadily. In the passenger transport, the decreasing aggregate energy intensity was attributable to the intensity effect that lowered the overall energy intensity in the passenger transport to around 9% from 2000 to 2016. Further, the findings also indicate that the transportation energy intensity effect had a greater contribution in reducing overall energy intensity in the passenger transport; while the role of the structural effect increased the aggregate energy intensity. The reduction of the intensity effect can be associated with better and more effective policy measures, for instance encouraging the use of cleaner fuels, upgrading the traffic equipment and supporting development of the new technology. In addition, since the fuel price reform, Indonesia's energy prices have increased substantially, although it is lower than international fuel prices. Therefore, energy price changes potentially have had a small impact on the overall energy intensity changes in the transport sector.

9 Policy implications

Several features of Indonesia's transport system are distinctive in an international context. Almost all the passenger kilometres are travelled on roads (95.6% in 2016), with virtually no rail (0.5%) or water (0.3%) passenger travel, but with a small but rising share of air travel (3.6% in 2016). A good transportation system is one of the main components for strong economic growth [34]. However, energy consumption in this sector has grown continuously over the last decade and will potentially continue to increase in the future. Attempts to reduce the energy use in the transportation sector require significant attention, where the clear and coherent government policy needs to be established to conduct a suitable energy policy in this sector. Therefore, the findings in this study are aimed to establish a scientific evidence for developing a policy on energy efficiency measures in the transportation sector.

The structural effect of the passenger transport indicated a negative influence (that is, increasing) aggregate energy intensity over the period of the study. This seemingly indicates contradictory efforts in reducing overall energy intensity. Based on the observation of energy usage performance across the transportation modes in the study period, the breakdown of the road and air transport turnover for passenger has increased substantially, employing more energy per unit of turnover compared to other modes of transport. Therefore, to lower the level of energy use in the transport sector, one of the key policy measures required is to promote different modes of transport besides the road and air transport, such as railways, waterways transport and other modes of transport that consume less energy.

The road transportation system is a crucial sector that not only moves people and goods, but is also essential for industry and local trade. Therefore, the key strategy to further reduce the energy intensity growth in the transport sector is to lower the growth of energy consumption. Enacting green energy policies, such as shifting to cleaner energy fuels [35] and encouraging the use of public transport [36], including the rail and water transportation would help to reach this goal. Regulatory measures, including vehicle efficiency and occupancy standards [37] are needed to reduce the energy intensity in the transport sector. Moreover, fiscal measures, such as providing tax incentives and subsidies for development of the public transportation [36], cleaner fuels or development of green technology could be considered as an effective way to encourage awareness of the people to use energy more efficiently that leads towards a modal shifting to more energy efficient transport modes.

Overall, Indonesia's transport structure needs to be focused on the waterway and railway transports that have lower energy use and greater transport capacity. The Indonesian Government needs to enact various policy measures to expand investment in the waterways and railways. Improving development of the inland waterways and railways network can be done gradually by providing more effort in optimizing networks across different transport modes and enhancing the road infrastructure to make all the other transport modes closely connected [38]. These efforts are expected to create an efficient transport system, where railways and waterways become more dominant in the Indonesia's transport system.

References

[1] WorldBank 2014. *Indonesia Economic Quarterly*. Jakarta: World Bank Group, 2014.

[2] DEENDARLIANTO, WIDYAPARAGA, A., SOPHA, B. M., BUDIMAN, A., MUTHOHAR, I., SETIAWAN, I. C., LINDASISTA, A., SOEMARDJITO, J., OKA, K. Scenarios analysis of energy mix for road transportation sector in Indonesia. *Renewable and Sustainable Energy Reviews* [online]. 2017, **70**, p. 13-23. ISSN 1364-0321. Available from: https://doi.org/10.1016/j.rser.2016.11.206

[3] IEA 2015, Southeast Asia Energy Outlook 2015, Directorate of Global Energy Economics, International Energy Agency, France.
[4] HOWES, S., DAVIES, R. Survey of recent development. *Bulletin of Indonesian Economic Studies* [online]. 2014, **50**(2), p. 157-183. ISSN 0007-4918, eISSN 1472-7234. Available from: https://doi.org/10.1080/00074918.2015.1016565
[5] ADB 2015, *Fossil fueal subsidies in Indonesia, trends, impacts and reforms*. Asian Development Bank, Manila, Philippines: Asian Development Bank, 2015.
[6] BEATON, C., LONTOH, L. *Lessons learned from Indonesia's attempts to reform fossil-fuel subsidies* [online]. Geneva, Switzerland: International Institute for Sustanaible Development, 2010. Available from: https://www.iisd.org/pdf/2010/lessons_indonesia_fossil_fuel_reform.pdf
[7] CHEN, K.-H., YANG, H.-Y., LEE, J.-M., CHI, C.-F. The impact of energy prices on energy consumption and energy efficiency: evidence from Taiwan. *Energy Efficiency* [online]. 2016, **9**(6), p. 1329-1349. ISSN 1570-646X, eISSN 1570-6478. Available from: https://doi.org/10.1007/s12053-016-9426-y
[8] HE, L., DING, Z,. YIN, F., WU, M. The impact of relative energy prices on industrial energy consumption in China: a consideration of inflation costs. *SpringerPlus* [online]. 2016, **5**(1), p. 1-21. eISSN 2193-1801. Available from: https://doi.org/10.1186/s40064-016-2661
[9] IEA 1999. *World energy outlook insights, looking at energy subsidies: getting the prices right*. Paris: International Energy Agency (IEA).
[10] OECD 1998. *Improving the environment by reducing subsidies*. Paris: Organization for Economic Development and Cooperation, 1998.
[11] SCHIPPER, L., STEINER, R., DUERR, P., AN, F., STROM, S. Energy use in passenger transport in OECD countries: changes since 1970. *Transportation: Planning - Policy - Research - Practice* [online]. 1992, **19**(1), p. 25-42. ISSN 0049-4488, eISSN 1572-9435. Available from: https://doi.org/10.1007/BF01130772
[12] KEPPLER, J., BIROL, F. *Prices, technology development and the rebound effect*. IDEAS Working Paper Series from RePEc, 2000.
[13] HANG, L., TU, M. The impacts of energy prices on energy intensity: evidence from China. *Energy Policy* [online]. 2007, **35**(5), p. 2978-2988. ISSN 0301-4215. Available from: https://doi.org/10.1016/j.enpol.2006.10.022
[14] FISHER-VANDEN, K., JEFFERSON, G. H., LIU, H., TAO, Q. What is driving China's decline in energy intensity? *Resource and Energy Economics* [online]. 2004, **26**(1), p. 77-97. ISSN 0928-7655, eISSN 1873-0221. Available from: https://doi.org/10.1016/j.reseneeco.2003.07.002
[15] LIN, B., MOUBARAK, M. Estimation of energy saving potential in China's paper industry. *Energy* [online]. 2014, **65**, p. 182-189. ISSN 0360-5442, eISSN 1873-6785. Available from: https://doi.org/10.1016/j.energy.2013.12.014
[16] FARAJZADEH, Z., NEMATOLLAHI, M. A. Energy intensity and its components in Iran: determinants and trends. *Energy Economics* [online]. 2018, **73**, p. 161-177. ISSN 0140-9883, eISSN 1873-6181. Available from: https://doi.org/10.1016/j.eneco.2018.05.021
[17] SONG, F., ZHENG, X. What drives the change in China's energy intensity: combining decomposition analysis and econometric analysis at the provincial level. *Energy Policy* [online]. 2012, **51**, p. 445-453. ISSN 0301-4215. Available from: https://doi.org/10.1016/j.enpol.2012.08.044
[18] YANG, G., LI, W., WANG, J., ZHANG, D. A comparative study on the influential factors of China's provincial energy intensity. *Energy Policy* [online]. 2016, **88**, p. 74-85. ISSN 0301-4215. Available from: https://doi.org/10.1016/j.enpol.2015.10.011
[19] DUBE, I. Impact of energy subsidies on energy consumption and supply in Zimbabwe. Do the urban poor really benefit? *Energy Policy* [online]. 2003, **31**(15), p. 1635-1645. ISSN 0301-4215. Available from: https://doi.org/10.1016/S0301-4215(02)00229-X
[20] GANGOPADHYAY, S., RAMASWAMI, B., WADHWA, W. Reducing subsidies on household fuels in India: how will it affect the poor? *Energy Policy* [online]. 2005, **33**(18), p. 2326-2336. ISSN 0301-4215. Available from: https://doi.org/10.1016/j.enpol.2004.04.024
[21] KEBEDE, B. Energy subsidies and costs in urban Ethiopia: the cases of kerosene and electricity. *Renewable Energy* [online]. 2006, **31**(13), p. 2140-2151. ISSN 0960-1481, eISSN 1879-0682. Available from: https://doi.org/10.1016/j.renene.2005.10.005
[22] CHATTOPADHYAY, P. Cross-subsidy in electricity tariffs: evidence from India. *Energy Policy* [online]. 2004, **32**(5), p. 673-684. ISSN 0301-4215. Available from: https://doi.org/10.1016/S0301-4215(02)00332-4
[23] SKRUCANY, T., KENDRA, M., STOPKA, O., MILOJEVIC, S., FIGLUS, T., CSISZAR, C. (2019) Impact of the electric mobility implementation on the greenhouse gases production in central european countries. *Sustainability* [online]. 2019, **11**(18), 4948. eISSN 2071-1050. Available from: https://doi.org/10.3390/su11184948
[24] GOGOLA, M., VETERNIK, M. The analysis of the environmental impact of road traffic in city of Martin. *Logi - Scientific Journal on Transport and Logistics*. 2016, 7(1), p. 49-60. ISSN 1804-3216.

[25] LIZBETIN, J., HLATKA, M., BARTUSKA, L. Issues concerning declared energy consumption and greenhouse gas emissions of FAME biofuels. *Sustainability* [online]. 2018, **10**(9), 3025. eISSN 2071-1050. Available from: https://doi.org/10.3390/su10093025

[26] CEMPIREK, V., RYBICKA, I., LJUBAJ, I. (2019). Development of electromobility in terms of freight transport. *LOGI - Scientific Journal on Transport and Logistics* [online]. 2019, **10**(2), p. 23-32. ISSN 1804-3216. Available from: https://doi.org/10.2478/logi-2019-0012

[27] ANG, B. W. LMDI decomposition approach: a guide for implementation. *Energy Policy* [online]. 2015, **86**, p. 233-238. ISSN 0301-4215. Available from: https://doi.org/10.1016/j.enpol.2015.07.007

[28] ZHANG, X., EMMERSON, P. *Transport databank*. Clean Air Asia, 2016.

[29] LIPSCY, P. Y., SCHIPPER, L. Energy efficiency in the Japanese transport sector. *Energy Policy* [online]. 2013, **56**, p. 248-258. ISSN 0301-4215. Available from: https://doi.org/10.1016/j.enpol.2012.12.045

[30] MASIH, A. M. M., MASIH, R. Energy consumption, real income and temporal causality: results from a multi-country study based on cointegration and error-correction modelling techniques. *Energy Economics* [online]. 1996, **18(**3), p. 165-183. ISSN 0140-9883, eISSN 1873-6181. Available from: https://doi.org/10.1016/0140-9883(96)00009-6

[31] CHAI, J., LU, Q.-Y., WANG, S.-Y., LAI, K. K. Analysis of road transportation energy consumption demand in China. *Transportation Research Part D Transport and Environment* [online]. 2016, **48**, p. 112-124. ISSN 1361-9209, eISSN 1879-2340. Available from: https://doi.org/10.1016/j.trd.2016.08.009

[32] TIWARI, P., GULATI, M. An analysis of trends in passenger and freight transport energy consumption in India. *Research in Transportation Economics* [online]. 2013, **38**(1), p. 84-90. ISSN 0739-8859. Available from: https://doi.org/10.1016/j.retrec.2012.05.003

[33] IEA 2017. *Energy Efficiency 2017*. International Energy Agency, 2017.

[34] PRADHAN, R. P., BAGCHI, T. P. Effect of transportation infrastructure on economic growth in India: The VECM approach. *Research in Transportation Economics* [online], 2013, **38**(1), p. 139-148. ISSN 0739-8859. Available from: https://doi.org/10.1016/j.retrec.2012.05.008

[35] SCHIPPER, L., NG, W. *Rapid motorization in China: environmental and social challenges*. ADB-JBIC-World Bank East Asia and Pacific Infrastructure Flagship Study, 2004.

[36] ADB 2006. *Energy efficiency and climate change considerations for on-road transport in Asia*. Philippines: Asian Development Bank (ADB), 2006.

[37] AN, F., SAUER, A. Comparison of passenger vehicle fuel economy and GHG emission standards around the world [online]. Arlington, VA: Pew Center on Global Climate Change, 2004. Available from: https://www.c2es.org/site/assets/uploads/2004/12/comparison-passenger-vehicle-fuel-economy-ghg-emission-standards-around-world.pdf

[38] WANG, Y., LI, K. - P., XU, X. Transport energy consumption and saving in China. *Renewable and Sustainable Energy Reviews* [online]. 2014, **29**, p. 641-655. ISSN 1364-0321. Available from: https://doi.org/10.1016/j.rser.2013.08.104

Jaromir Siroky - Katarina Magdechova - Petr Nachtigall - Stephan Schroder - Pavlina Siroka

IMPLEMENTATION SYSTEM OF TECHNICAL SPECIFICATION FOR INTEROPERABILITY FOR THE ROLLING STOCKS

In the paper, the present state of scientific knowledge in authorising the types of railway vehicles in Slovakia and abroad is analysed by the authors. The aim of the authorisation of the types of railway vehicles is to take into account the interests of the whole society in the field of transport in the manufacturing of these vehicles or their import from abroad. These interests mainly include a greater security of transport, as well as a greater quality, reliability and lifespan of railway vehicles. The authors also analyse the Commission Recommendation 2014/897/EU of 5 December 2014 on matters related to the placing in service and use of structural subsystems and vehicles under Directives 2008/57/EC and 2004/49/EC of the European Parliament and of the Council, which should contain instructions related to the matters in question. At the same time, it is necessary to take into account Directive (EU) 2016/797 of the European Parliament and of the Council of 11 May 2016 on the interoperability of the rail system within the European Union, and Directive (EU) 2016/798 of the European Parliament and of the Council of 11 May 2016 on the railway safety. In addition to ensuring the applicability of interoperability regulations (new directives, regulations, decisions, recommendations, etc.), the aim of this paper is to present a methodology of authorising the types of railway vehicles and a software model (SW) of authorising the types of railway vehicles, with European legislation as a possible tool to speed up and simplify the entire process. The paper is concluded by three types of authorisation processes for different railway vehicles and their accessories, comparing the length and complexity of processing using the standard authorisation process and using the SW tool proposed.

Keywords: *Single European Area, interoperability, methodology of authorisation type of vehicles, software model, authorisation type of vehicles, TSI*

1 Introduction

The main task of the authorisation process of the vehicles types is to attract wider interest of the public in production and import in the area of transport. It mainly applies to safety but also to ensuring the quality, reliability and lifecycle of the rolling stocks at the required level. These requirements need to be met and at the same time they should be a tool to satisfy transport needs. The vehicles should reflect the progress in science and technology and their construction should be in harmony with the development of transport [1-2].

Achieving the aims of interoperation within the spectre of the rail system in the European Union should result in setting up the optimal level of technical harmonization and mitigate, improve and further develop services that are offered within international railway transport [3]. The aim is to create an inner market with facilities and services targeting at building, innovation, modernization and operation of the rail system in the Union. Within the fourth railway packet, a new procedure is proposed of authorising a new type of railway vehicles and placing them in service. The aim is to transfer the competences of authorising and certifying the types of railway vehicles to the European Union Agency for Railways (EUAR) [4]. Individual activities should be carried out on the basis of dividing the competences between the EUAR and the National Safety Authority (NSA). Objective activities will be carried out on the basis of the legal relationship between these two bodies. Fundamental is the elimination of national regulations and direct exercisability of technical specification for interoperability (TSI) as internal regulations for the entire railway network [5].

2 Theoretical background

Authorising the types of railway vehicles is an integral part of the transport process. The main goal should be securing the safety of transport. The aim of every authorising body is to ensure that the authorisation process is carried out in accordance with the European legislation and in a non-discriminatory way [6].

Development of authorising the types of railway vehicles has gone through several changes since the

Jaromir Siroky[1,*], Katarina Magdechova[2], Petr Nachtigall[1], Stephan Schroder[3], Pavlina Siroka[1]
[1]Department of Transport Technology and Control, Faculty of Transport Engineering, University of Pardubice, Czech Republic
[2]Wagon Service Travel, s.r.o., Bratislava, Slovakia
[3]Bahnzentrum Ingenieurburo, Syntrans GmbH, Germany
*E-mail of corresponding author: Jaromir.Siroky@upce.cz

establishment of the Slovak Republic. Every development stage was aiming to make the process easier and ensure the continuity of particular activities in every area [7].

Authorising the types of railway vehicles has changed a lot since the Slovak Republic was established in 1993 [8]. It is possible to divide this process into several development stages that are described in detail in this paper. These are:

- authorising the types of railway vehicles in the years 1993-1996,
- authorising the types of railway vehicles in the years 1996-2009,
- authorising the types of railway vehicles in the years 2009-2014.
- authorising the types of railway vehicles in the years after the possible shift of competences to the EUAR (after 2020) [9-10].

In individual countries, authorising the types of railway vehicles and placing them in service is carried out by the authorising bodies or the NSA (National Safety Authority) in accordance with the Directive (EU) 2016/797 of the European Parliament and of the Council on the interoperability of the rail system within the European Union and also in accordance with particular related regulations, decisions and recommendations, and upon their transposition into the national legislation [11].

To analyse the current situation in the area of authorising the types of railway vehicles, the knowledge acquired at different conferences was used, as well as the available literature and studies from universities and research institutes. The analysis was carried out in six countries - Poland, Austria, Hungary, the Czech Republic, Spain and Germany [12-13].

Within the analysis, the legislation in the area of authorising the types of railway vehicles was analyzed and competences of particular bodies. With help of administrative outputs, an overview was made of registered vehicles in individual member states, notified bodies carried out in member states and types of vehicles registered in the European Register of Authorised Types of Vehicles (ERATV) [14-15].

The Transport Research Institute (Vyskumny ustav dopravny, a. s.) [16] has made a study called Implementation of interoperability of conventional railways of the Slovak Republic. The study describes the current interoperability and safety situation in terms of guidelines implementation, technical specifications for interoperability and assessing conformity of the individual components. To effectively address interoperability, an organisational proposal was prepared defining the key roles of state administration, railway undertakings, infrastructure manager and other stakeholders. The conformity assessment part describes the baselines of conformity assessment, requirements to be met by the authorised bodies, notification requirements and specific requirements in the process of assessing conformity of interoperability components and subsystems.

The study [17] addresses the use of TSIs as technical regulations, which are an integral part of authorising the types of railway vehicles and as such, the study was used mainly in the first part of the work in analysing individual TSIs related to railway vehicles. However, it only served as an aid as much of the data needed to be updated.

At the Department of Transport and Handling Machines of the Faculty of Mechanical Engineering of the University of Zilina [18], a computation tool for railway vehicle gauge was created with emphasis on interoperability. This program facilitates the work of manufacturers in using the TSIs and assessing conformity with TSIs. Its output is an input for the Notified Body (NoBo) and designated legal person (PPO) in terms of meeting gauge requirements which are provided to the authoriser of the type of railway vehicle based on this input.

In his speech on the international ERTMS conference, Guido [19-20] from the European Railway Agency specified the possible ways to ensure interoperability, analysing related legislation and the responsibilities of the individual bodies: NoBo, DeBo (designated bodies), etc. The knowledge is mainly used in the graphic representation of the hierarchy of legislation and distribution of competences, which is also used in creating the SW model.

The scientific work by Cech [21] focuses on benefits of introducing the interoperability into the European rail system. The author analyzed selected TSIs and focused on modelling the costs and benefits of interoperability, without a specific focus on railway vehicles though.

Jindra [22] is focused on proposing a complex tool for creation of the wagon load reference plans and for data exchange in accordance with the Telematics applications in freight transport TSI (TAF TSI) and analysed the TAF TSI in detail.

Type tests can have a significant impact on the general construction of the railway vehicle. As mentioned above, in manufacturing the railway vehicles, the manufacturer abides by the TSI. The vehicles also have to meet the TSI requirements in the “railway vehicles - noise” subsystem, setting out individual limits for railway vehicle noise [23]. Based on available sources of information, a suitable noise assessment tool can be the simulation using the TWINS software. This software assesses the noise sources, which can be different on the testing track and on the track in operation. The software must be updated in accordance with the amendments to noise legislation. Such update would significantly simplify the definition of manufacturing requirements and provide inputs for NoBos and authorized legal entity (PPOs) [24].

Commission Recommendation 2014/897/EU defines particular important elements in the area of authorising structural subsystems and placing vehicles in service. It is a broad guideline on how to proceed in this area in line with the provisions of Directive 2008/57/ES. There is no graphically shown procedure in the area of authorising the types of railway vehicles that are in accordance with technical specifications for interoperability (TSI) or are not in accordance with the TSI and their consequent placing in service. This was a part of the suggested methodology [25-26].

Individual scientific studies focus only on partial activities in the area of authorising the types of railway vehicles. They do not define the authorisation process as a complex of activities, they do not say how to simplify or speed it up, as a whole. To shorten the duration of the authorisation process, it was necessary to review and analyse particular problems that could emerge during the process. Because of this, it was inevitable to come up with operative solutions of problems that can occur. Not to foresee these problems could lead to extending the time necessary for authorising. Creating a convenient methodology for authorising the types of railway vehicles is an appropriate solution of these drawbacks, preventing the extension of the authorisation process.

3 Methodology

Based on the found differences and problems that could emerge during the process of authorising the types of railway vehicles, this paper outlines a methodology of authorising the types of railway vehicles as a supporting tool for setting up the SW model.

Apart from the basic methods (analysis, synthesis, brainstorming, etc.), other methods that are necessary for solving the stated problems were used, as well, [27].

The aim of the outlined methodology in the area of authorising the types of railway vehicles with the emphasis on using interoperability, new European regulations, rules and decisions, was to set up particular possible ways of carrying out the authorisation. The methodology is graphically shown using a development diagram and is afterwards reflected in an SW model, which is the main tool to minimise the administrative difficulties of the whole process.

The SW model can also be considered a tool to minimise the administrative problems of the whole process. It was created using the Visual Basic for Applications (VBA) macros.

The Microsoft Office package – Microsoft Excel with VBA editor have been applied by creating the Form. This decision was based on several criteria [28]:

1. MS Excel is part of the Microsoft Office package, which is currently installed on most computers used in administration.
2. MS Excel includes VBA editor, which is actually a programming language for Microsoft Office and this opens a lot more other options.
3. It is also possible to open other files from the MS Excel, which allows us to have a database in a separate file.
4. Within the VBA, it is possible to create files in MS Word.
5. Possibility to create pleasant and user-friendly graphical environment.
6. Possibility to create keys and sheets corresponding to a standard the user is familiar with from MS Windows.

Programming

The creation or generation of the Form and the final decision about authorising the type of railway vehicle were carried out by the VBA macros - using programming. Both global variables and functions were used.

Form (Sheet)

The main exercise book that is used by the user was named the “Form”. It contains one page with the same name. The whole Form was named “The Form about Decision on Authorising the Railway Vehicle”. It is pre-programmed and locked and the user can fill it in by the keys used for this purpose. In the first part, there are “Fill in” keys that are used to fill in information and a “Delete” key that is used to delete the filled in data. The latter is of course secured by the warning whether the Form should really be deleted. Filling in the Form can be postponed at any stage and particular steps can be filled in separately. For filling in the Form, sheets created in VBA are always used. There are programmed functions and automatic filling in to make the work for the user as easy as possible, effective and without mistakes. The Form adapts itself according to filled in data so the original and the final ones can look completely different. The outcome of the Form is a prescribed decision in the MS Word. Another advantage of the Form is that it is automatically saved in a file and as such, it is possible to go back to filling it in any time in the future [29].

VBA

Programming in VBA is divided into the following files:

1. Modules

It contains only one module (General), which includes global variables and functions for the work with MS Word and Outlook.

2. Microsoft Excel Objects

It contains one “Workbook”, which also has a function that starts right after opening the Form. This function should serve to find the way to the file. It is a key for the next steps, allowing for opening the database and prescribed documents in Word.

It also contains one ”Worksheet”, which includes:

- a function for opening the database,
- a function that controls whether particular cells in the Form have not been changed and based on these changes, it saves the Form with a new name or adapts the Form according to filled in data (shows and hides parts of the Form),
- it also contains serving functions that are activated by pressing particular keys. Note: these functions open further particular forms and functions according to the key pressed.

Forms

There are 15 forms that are used to fill in the Form. Those forms are activated automatically or after pressing a key.

Every form consists of two parts:

- ”Object“, which is its graphic draft,
- ”Code“, which includes the source code for a particular ”Object”. The ”Code“ contains functions that are

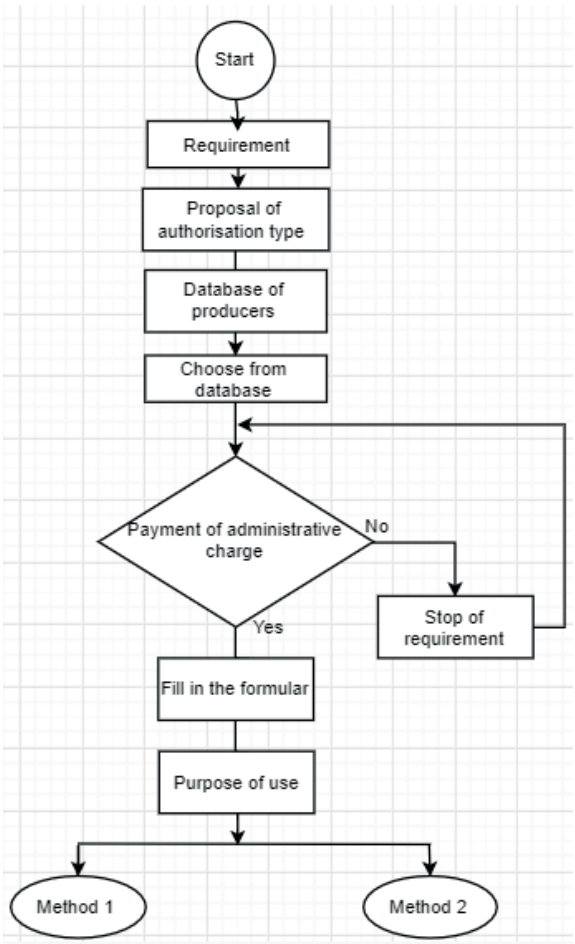


Figure 1 *Basis for the VBA I*

activated when opening a particular form (it serves for the starting initialisation and filling in the form based on the data already stated in the form). It also includes functions that are activated while filling in particular cells (these serve mainly to ensure that the data have been filled correctly or to make sure that data have been filled automatically). Functions that are shown after pressing a button mainly serve to open other forms or to close the form.

Parts in a graphic interface for the Form were set up using "Properties" for particular elements. All the functions and methods used when programming are described in the Help Excel - VBA (option displayed after pressing F1 when the cursor is on this command).

Filling in the Form leads to generating a complete decision that is filled in with the data from the database, as well.

As shown below in Figures 1 and 2, the process starts with submitting a request for authorising the type of a railway vehicle. At first, it is necessary to state or decide whether it will be authorisation of the type of a railway vehicle, its modernisation or innovation. Sending the request electronically will lead to a faster submission and the authorising body can thus start the process of its evaluation sooner. Afterwards, the authorising body can find an applicant in a database of producers, or it can generate it for the next use. The database is, or will be, constantly updated according to the needs of the authorising body, which will speed up the process of decision making, where the date will be automatically reflected.

After setting the purpose of use, it is possible to state which way the authorisation will go, or which method will be used.

In method 1, the authorisation according to specifications of regulations, mainly with the national legislation, is considered. That is why in this methodology the fact that the user will have an updated version of the regulation is considered. After meeting the requirements, it is possible to proceed to generating the decision itself.

The second version of authorisation is verifying the compliance with the TSI. Here it is necessary to get a review from the Notified Body (NoBo). That is why the option of connecting to NoBo is considered. The system also takes into account the possibility of submitting information about the type of vehicle, range of vehicle, power, etc.

Figure 3 shows the graphic representation and programming of the Form.

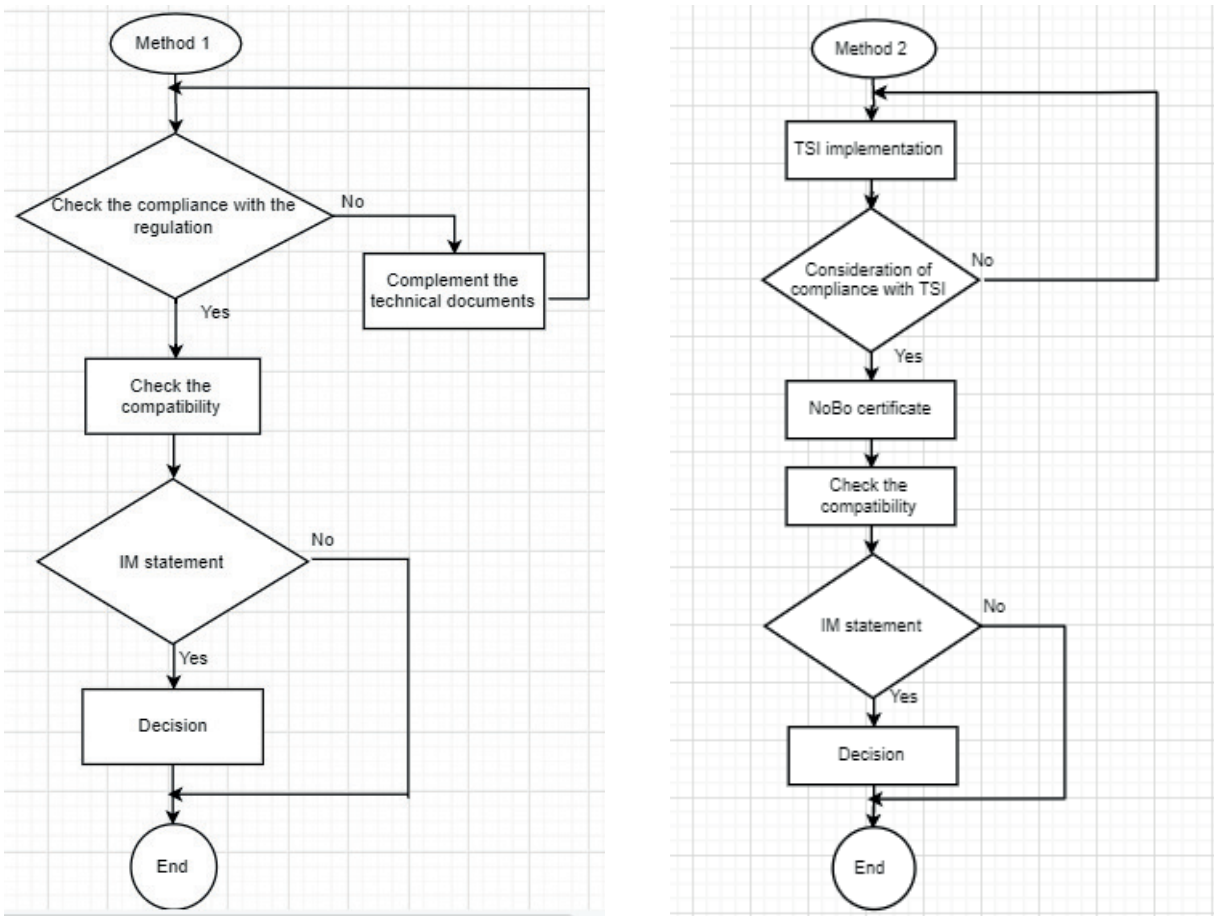


Figure 2 *Basis for the VBA for methods I and II*

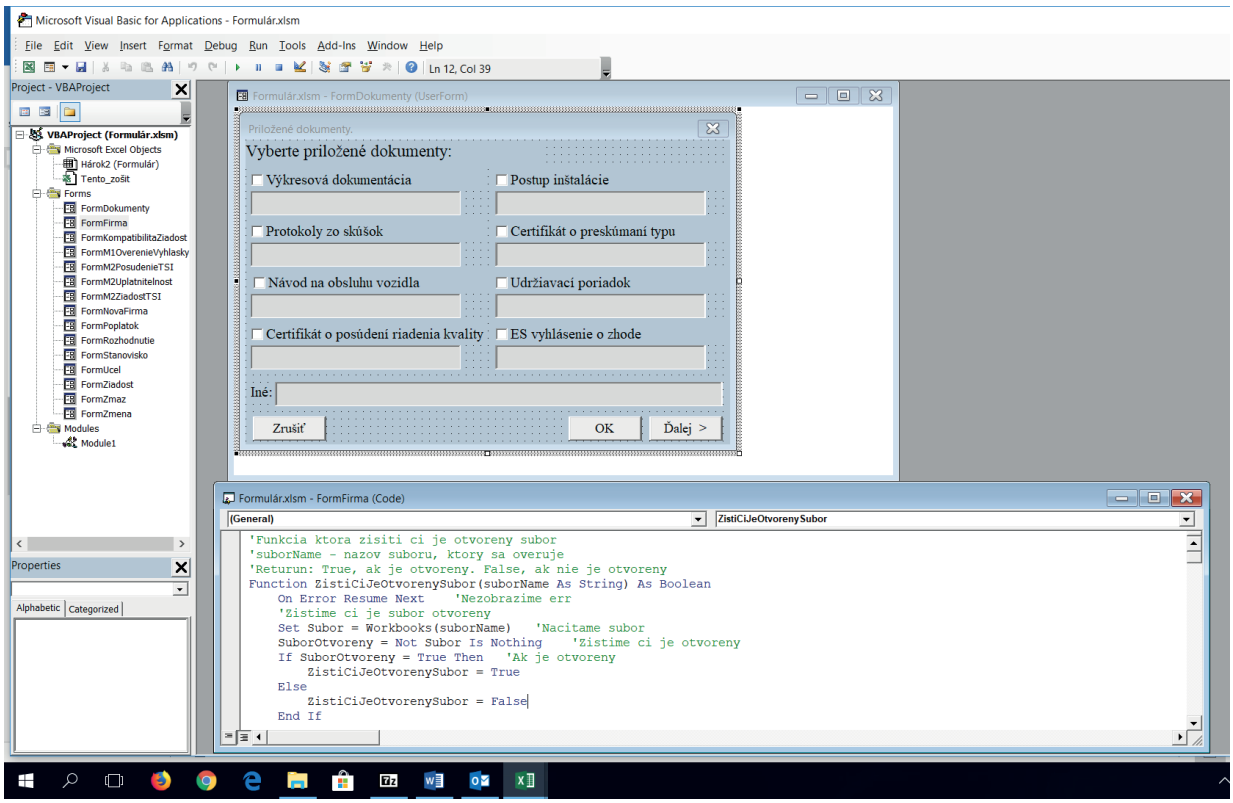


Figure 3 *Graphic representation*

	April	Mai	Juni	Juli	August	September	October	November	December	January	February
Request submission	17.4.										
Request analysis	23.4.		10.6.								
Informing manufacturers about the need of a NoBo opinion			11.6.								
Request prepared by the manufacturer and sent to the NoBo			16.6 20.6.								
Data processed by the NoBo			21.6.		28.8.						
IM opinion needed				31.7.		30.9.					
Suspension						26.9.	16.10.				
Provision of documents from the IM and NoBo							16.10.	18.11.			
Decision processing								18.11.			19.2.
Total duration of 213 days	**8 days**	**27 days**	**11 days**		**44 days**		**123 days**				

Figure 4 *The Gantt chart of the ETCS L1 duration*

	April	Mai	Juni	Juli	August	September	October	November	December
Request submission	17.4.								
Request analysis	23.4.		10.6.						
Informing manufacturers about the need of a NoBo opinion			11.6.						
Data processed by the NoBo			11.6.		20.8.				
IM opinion needed				31.7.		22.9.			
Suspension						18.9.	8.10.		
Provision of documents from the IM and NoBo							8.10.	7.11.	
Decision processing								7.11. 18.11.	19.12.
Total duration of 164 days	**8 days**	**27 days**	**13 days**		**43 days**		**73 days**		

Figure 5 *The Gantt chart of the ETCS L1 using SW*

Table 1 *Conclusion of verifications of cases 1-3*

Case	Without SW [working days]	With SW [working days]	Saving [working days]	Saving [%]
C1	213	164	49	23.00
C2	122	85	37	30.33
C3	153	99	54	35.29

4 Achieved results

The methodology of authorising the railway vehicles and its usage was verified by the use of simulation directly in MS Excel with use of the VBA macros. The aim was to provide an estimate of duration of the authorisation process of the railway vehicles or different cases within this process. The Gantt chart was used to show duration of particular activities.

Case 1: Installing a mobile part of European Train Control System Line 1 into an electrical multiple unit class 671 type 214

The total duration of the process allowing an essential change by installing the ETCS L1 into the electric multiple unit (EMJ) class 671 type 214 was 213 working days. The total duration of the process allowing an essential change by installing EMJ class 671 type 214 into ETCS L1 into a vehicle with the use of SW was 164 working days. The process of authorisation with use of the system was shortened as shown below (Figure 4). The percentage saving was 23%. Creating a prototype board was not considered. Even with this, one can see that the duration of the process decreased significantly. Duration of authorisation is therefore individual and depends on the particular case.

Figure 4 shows the authorisation of an essential change of a train protection system and duration of the individual processes without using the system. The total duration of the authorisation process for 2014 was 213 working days. As can be seen from the Gantt chart, it was necessary to suspend the process in September due to the absence of opinion of the infrastructure manager and the NoBo. The process was suspended due to the absence of documents to be provided by the NoBo and to serve as an input for the infrastructure manager.

Using the SW model shortened the authorisation process by 49 days. If the prototype commission was established as early as in the preparatory phase, i.e. before the submission of the authorisation request, the process could be shortened for duration of the suspension (nearly one month). The benefit of using the system is the streamlining of data entering and processing (Figure 5).

Case 2: Authorising the three-system electrical locomotive class 381 type 109 E2

The total duration of the authorisation process of the locomotive class 381 type 109 E2 was 122 working days. The total net duration of the authorisation process with use of the SW system was 64 working days. To make this time even more effective, a prototype commission was established, which allowed for solving several possible inconsistencies even before submitting the application for authorisation. The process of authorisation with use of the system was shortened as shown below (Table 1). The percentage saving of the authorisation process, including duration of the prototype commission, was 30.33%. The prototype commission markedly influences the net duration of the authorisation process. As already mentioned, the time saving is very individual. A reduction in duration can be secured by establishing the prototype commission.

Case 3: Authorising the diesel unit class 861 type VR - 24 - 2010 - DMJ

The total duration of the authorisation process of DMJ class 861 type VR - 24 - 2010 - DMJ was 153 working days. The total net duration of the authorisation process with use of the SW was 40 working days. To make this time even more effective, a prototype commission was established, which allowed for solving some of the inconsistencies even before submitting the application for authorisation. The process of authorising with use of the system was shortened as shown below (Table 1). The percentage saving of the authorisation process, including the duration of the prototype commission, was 35.29%. As mentioned before, the time saving is very individual and changes with each case. To further reduce the duration, it is also possible to establish the prototype commission, which can influence the duration of the authorisation process.

5 Discussion

Based on the verification of the outlined methodology of authorising the types of railway vehicles with the use of the SW system, one can conclude that duration of the authorisation process is very individual and changes with each case.

In verifying process the stated hypothesis in practice in the first case - implementing a mobile part ETCS L1 into an electric multiple unit class 671 type 214 - the hypothesis was not confirmed. Duration of the authorisation decreased by 23%. However, even this result is a significant shortening.

In the second case - authorising the three-system of an electric locomotive class 381 type 109 E2 - duration of authorisation was significantly shorter. It is very important to note that a prototype commission was established, which can directly influence the net duration of the authorisation process. Since the duration was 30.33% shorter, the hypothesis was confirmed.

The third case - approving the diesel unit class 861 type VR - 24 - 2010 - DMJ - also confirmed the stated hypothesis. A prototype commission was established in this case as well, which influenced the total duration of authorisation to such an extent that the saving was 35.29%.

During the verification of the particular cases, simulation, stated estimation of duration of the particular stages (mostly when using the SW system) were used as well as the method of brainstorming. Durations were estimated based on discussions with the representatives of the authorising bodies or related subjects.

6 Conclusions

Complex mapping and summarising of the current situation in the authorisation process of the types of railway vehicles, both at home and abroad is taking a part of the paper. It gives information on the applicable legislation related to train railway vehicles. It also provides an overview of these matters to both the wide public and the academic sphere.

The methodology, which is based on an SW model, makes the authorisation process significantly easier and faster. It provides clear evidence and ensures more effective cooperation. It also brings the possibility of using the new modern SW tools in the authorisation process. Most of the communication happens online, including the issuing of the particular documents, which are a basis for the decision on the authorisation. Therefore, it saves postage costs, as well as office material.

International application of outlined methodology is possible. It is possible to change it according to individual needs. After adding further functions, it can be adjusted to needs of the EUAR and can speed up the communication and information exchange between the EUAR and national authorisation bodies. This can be ensured by a contact treaty between the EUAR and NSA. The aim is to use the SW model in accordance with the European legislation. This would result in reducing both duration and costs of the authorisation process.

The methodology is an asset also to producers and owners of the railway vehicles, as well as to the wide public. This methodology would bring a faster exchange of information, time and cost savings and a possibility of an operative solution of possible problems.

Acknowledgement

The work was supported by the ERDF/ESF "Project PosiTrans - University of Pardubice cooperation and application sphere in application oriented research of localization, detection and simulation systems for transport processes" (No. CZ.02.1.01/0.0/0.0/17_049/0008394).

References

[1] MAGDECHOVA, K. The analysis of the development of types of authorisation processes for different railway vehicles in Slovak Republic (in Slovak). In: International conference Horizons of Railway Transport 2014 : proceedings. 2014. ISBN 978-80-554-0918-4, p. 140-149.

[2] MAGDECHOVA, K., NACHTIGALL, P. Changes in types of authorisation processes for different railway vehicles (in Slovak). In: International conference Horizons of Railway Transport 2013 : proceedings. 2013. ISBN 978-80-554-0764-7, p. 242-248.

[3] European Commission, DG Move. *Report on the work performed by the task force on railway vehicles authorisation*. Brussel, 2012.
[4] EBA, Schweizerische Eidgenossenschaft, ANSF, BMVIT, HETI. Cross acceptance and corridor guideline. In: Control Command and Railway Communication Conference 2012 : proceedings. 2012.
[5] Reming Consult, a. s., Sudop Praha, a. s., AM Sudop spol. s. r. o. *Technical and economic study for the preparation and implementation of ERTMS on Corridor E* (in Slovak). 2010.
[6] RARDIACA, F., LO YACONO, L. *Authorisation type of vehicles*. Final report. 2013.
[7] IHNAT, P. Railway interoperability in EU (in Slovak). *Railway transport and logistics*. 2006, p. 103-109. ISSN 1336-7493.
[8] DANIS, J., DOLINAYOVA, A., CERNA, L., ZITRICKY, V. Impact of the economic situation in the Slovak Republic on performances of railway transport. *Periodica Polytechnica Transportation Engineering* [online]. 2019, **47**(2), p. 118-123. ISSN 0303-7800, eISSN 1587-3811. Available from: https://doi.org/10.3311/PPtr.11185
[9] Commission Decision No. 2011/314/EU concerning the technical specification for interoperability "relating to the operation and traffic management" subsystem of the trans - European conventional rail system [online]. Brussel, 2011. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:144:0001:0112:EN:PDF
[10] Commission Decision No. 2012/757/EU concerning the technical specification for interoperability relating to the "operation and traffic management" subsystem of the rail system in the European Union and amending Decision 2007/756/EC [online]. Brussel, 2012. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:345:0001:0076:EN:PDF
[11] GHAZEL, M. Formalizing a subset of ERTMS/ETCS specifications for verifications purposes. *Transportation Research Part C: Emerging Technologies* [online]. 2014, **42**, p. 60-75. ISSN 0968-090X, eISSN 1879-2359. Available from: https://doi.org/10.1016/j.trc.2014.02.002
[12] GASPARIK, J., SIROKY, J., PECENY, L., HALAS, M. Methodology for assessing the quality of rail connections on the network. *Communications: Scientific Letters of the University of Zilina* [online]. 2014, **16**(2), p. 25-30. ISSN 1335-4205, eISSN 2585-7878. Available from: http://komunikacie.uniza.sk/index.php/communications/article/view/503
[13] KAMPF, R., GASPARIK, J., KUDLACKOVA, N. Application of different forms of transport in relation to the process of transport user value creation. *Periodica Polytechnica Transportation Engineering* [online]. 2012, **40**(2), p. 71-75. ISSN 0303-7800, eISSN 1587-3811. Available from: https://doi.org/10.3311/pp.tr.2012-2.05
[14] ERA, EK. Progress on the vehicle authorisation task force recommendations. In: EC ERA Workshop on Vehicle Authorisation: proceedings. 2013.
[15] LAVOGIEZ, H., DECHAMPS, J.-M., ARDIACA, F. Type of vehicle. In: ERA Workshop on Vehicle Authorisation : proceedings. 2012.
[16] VUD Transport Research Institute. Implementation of interoperability in Slovak Railway (in Slovak). Final report. 2005.
[17] BIERLEIN, H. Certification and placing in service. In: ERTMS World Conference : proceedings. [online]. 2012. Available from: https://www.uic.org/com/uic-e-news/274/article/10th-uic-ertms-world-conference-2648.
[18] BACISIN, M., FUSATY, M., PALUCH, J. Calculation of the gauge of railway vehicles with regard to interoperability (in Slovak). *Railway transport and logistics*. 2012, p. 10-15. ISSN 1336-7943.
[19] GUIDO, P. Specifications for interoperability. In: ERTMS World Conference : proceedings [online]. 2012. Available from: https://www.uic.org/com/uic-e-news/274/article/10th-uic-ertms-world-conference-2648
[20] GUIDO, P. ERTMS baselines. In: UIC ERTMS World Conference : proceedings [online]. 2014. Available from: https://uic.org/com/uic-e-news/366/article/11th-uic-ertms-world-conference?page=modal_enews
[21] CECH, R. *Costs and benefits analysis of interoperability implementation* (in Czech). Thesis DFJP. University of Pardubice. 2012.
[22] JINDRA, P. *System implementation of operational interoperability in freight railway transportation* (in Czech). Thesis DFJP. University of Pardubice. 2010.
[23] FRID, A., LETH, S., HOGSTROM, C., FARM, J. Noise control design of railway vehicles - impact of new legislation. *Journal of Sound and Vibration* [online]. 2006, **293**(3), p. 910-920. ISSN 0022-460X. Available from: https://doi.org/10.1016/j.jsv.2005.12.014
[24] LOCKETT, R. The vehicle authorisation process. In: ERTMS World Conference : proceedings. [online]. 2012. Available from: https://www.uic.org/com/uic-e-news/274/article/10th-uic-ertms-world-conference-2648
[25] GASPARIK, J., MAJERCAK, J., SIROKY, J., ABRAMOVIC, B., MESKO, P., NACHTIGALL, P., ZITRICKY, V. *Railway traffic operation*. Zilina: University of Zilina, 2017. ISBN 978-80-554-1281-8.
[26] MYDIA, S., THOTTAPPILLIL, R.: An overview of electromagnetic compatibility challenges in European Rail Traffic Management System. *Transportation Research Part C: Emerging Technologies* [online]. 2008, **16**(5), p. 515-534. ISSN 0968-090X, eISSN 1879-2359. Available from: https://doi.org/10.1016/j.trc.2007.11.001
[27] PAHOLOK, I. Simulation as a scientific method (in Slovak). *E - LOGOS Electronic Journal for Philosophy*. 2008, **1**, p. 1-19. ISSN 1211-0442.

[28] SIROKY, J., CEMPIREK, V., GASPARIK, J. *Transport technology and control*. Brno: Tribun EU, 2014. ISBN 978-80-263-0711-2.

[29] TOURNIER, N. T., DRILLER, J. The national legal framework (DE). Technical document (reference document part 3) [online]. 2014. Available from: https://rdd.era.europa.eu/rdd/documents/nlf/NLF%20DE%201.6.1%20for%20publication.pdf

Emilia M. Szumska - Rafal S. Jurecki - Marek Pawelczyk

LIFE CYCLE COST (LCC) LEVEL OF AN URBAN TRANSPORT FLEET WITH DIFFERENTIATED SHARE OF BUSES WITH ALTERNATIVE DRIVE SYSTEMS

In recent years, there has been a significant increase in the number of buses operated by urban public transport companies powered by alternative fuels and equipped with alternative drive systems. In addition to economic factors, operators should also take environmental aspects into account when purchasing new vehicles. In this case, a useful criterion for selecting a vehicle is the Life Cycle Cost (LCC), which, in addition to the cost of purchasing a bus, takes into account the necessary expenses associated with its maintenance, operation, decommissioning, as well as emissions costs. This paper presents a study of the LCC values, estimated for the entire bus fleet based on several bus replacement variants, taking into account different shares of alternative buses in the transport fleet. Analyses have shown that replacing conventional buses by the compressed natural gas (CNG) powered buses will reduce life cycle cost by 27% compared to the LCC level in 2019. Increasing the share of electric buses in the fleet will significantly reduce the level of emissions of harmful substances contained in exhaust gases.

Keywords: *LCC, city bus, alternative drive*

1 Introduction

Urban transport vehicles with alternative drive systems are already widely used by city bus companies. According to definition contained in Art. 2(1) of the Act on Electromobility and Alternative Fuels, a zero-emission bus may be a bus powered by electricity generated from hydrogen in fuel cells installed in it or equipped only with an engine operation cycle that does not lead to greenhouse gas emissions, i.e. a vehicle with an electric battery or network drive (trolleybus). It follows from the cited Act that the definition of a zero-emission vehicle is not equivalent to the definition of an alternative drive vehicle. Vehicles powered by alternative fuels, according to Art. 1(11) of the same Act, include vehicles driven by electricity, hydrogen, liquid biofuels, synthetic and paraffin fuels, compressed natural gas (CNG), including biomethane derived, liquefied natural gas (LNG), including biomethane derived and liquefied petroleum gas [1].

The main barriers to a rapid increase in the share of alternative vehicles in fleets of urban public transport companies are the high purchase prices of such vehicles and costly infrastructure. The decision to replace conventional buses by buses running on alternative fuels or equipped with alternative drive systems is supported by lower emission of harmful substances and lower operating costs.

The Life Cycle Cost method is a useful tool to compare the production and operating costs of vehicles with different types of drive systems. This analysis also makes it easier to decide on choosing and buying a bus with a specific drive variant. The LCC estimation allows to distinguish the costs, starting from the production of a vehicle through its use and operation up to the total depreciation (economic or accounting life) or its disposal as waste (technical life). The Life Cycle Cost method shall include a detailed economic analysis, taking into account investment costs (purchase, registration and additional infrastructure), operating costs (fuel, electricity, repair and maintenance, insurance), end-of-life costs (recycling) and emissions costs. In the literature, there are many papers on analysis of the life cycle cost of city buses equipped with different types of drive systems. Papers [2-3] contain an analysis and comparison of the LCC of buses equipped with conventional and alternative drive systems for the assumed service life.

The aforementioned types of costs in the LCC structure may be extended by cost categories important for the researcher. Works [4-5] present an analysis of the Life Cycle Cost of a bus with an electric drive system, which includes different types of electric power supply and different types of charging systems. The presented results show that lower Life Cycle Cost values are obtained by charging the batteries at stations located at end stops (terminuses).

In many cases, the Life Cycle Cost of a vehicle is also considered from environmental aspects. This method estimates the economic factors associated with production and use of a vehicle and the costs of emissions. The environmental impact, manifested in emission of the harmful substances related to the production and distribution of fuel (energy), production of car parts and components, assembly of parts and components, operation

Emilia M. Szumska*, Rafal S. Jurecki, Marek Pawelczyk
Department of Automotive Engineering and Transport, Faculty of Mechatronics and Automotive Engineering, Kielce University of Technology, Poland
*E-mail of corresponding author: eszumska@tu.kielce.pl

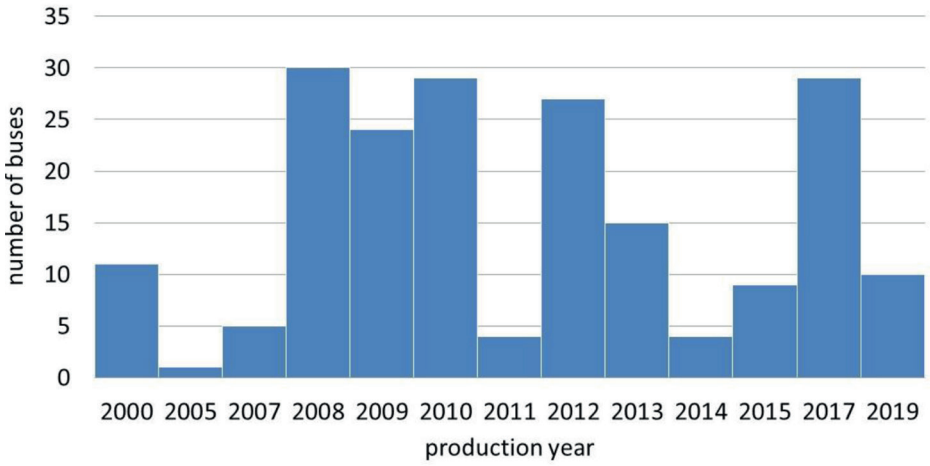


Figure 1 *Age structure of the bus fleet (based on [12])*

and use of a vehicle and its decommissioning, is expressed in monetary units and presented in the form of cost. The Life Cycle Cost analyses, including the costs of emissions of city buses equipped with conventional and alternative drives, are presented in papers [6-8].

The literature provides examples of Life Cycle Cost analyses developed not for a single vehicle but for the entire vehicle fleet. Paper [9] presents an estimation of the share of buses with hybrid and electric drive systems in the bus fleet, using optimization methods by minimizing the LCC, operating costs and emission costs. Paper [10] contains an analysis of the Life Cycle Cost of a truck fleet, in which the alternatively powered vehicles make up for 50% and 75% of the entire fleet. The results show that, despite the higher purchase price, a higher share of alternatively powered vehicles in the fleet leads to lower operating costs and significantly lower emissions of harmful substances in exhaust gases.

This paper presents 5 scenarios for the bus fleet modernization for Kielce in the period 2019-2030, each with different share of alternative buses. According to the adopted variants, the annual life cycle cost values for the entire bus fleet were estimated.

In the analysis presented, the LCC includes the following categories:

- purchase costs,
- operating costs,
- costs of repairs and maintenance,
- infrastructure costs,
- emission costs.

The presented research results may be used by city carriers and contribute to making a decision on purchase of alternative buses.

2 Characteristics of the Kielce public transport fleet

Kielce city transport network includes 66 day lines and two night lines. The total length of the city bus lines is 610 km and the length of the suburban bus lines is 145 km. Over 90,000 passengers use the Kielce public transport every day. The transport performance carried out during a year amounts to 12.54 million passenger kilometers [11].

At present, the Kielce transport fleet consists of 188 buses. The majority of them, i.e. 147, are small and medium buses (9 m, 10 m and 12 m). The number of 18 m articulated buses is 41. The average age of the fleet is 8 years. The age structure of the fleet is shown in Figure 1.

Most vehicles are equipped with diesel drive systems. Since 2017, Kielce has also been operating 25 buses with hybrid drive systems (combustion-electric). Currently, their share in the fleet amounts to 13%. It is planned to purchase 10 buses with engines fueled by compressed natural gas (CNG) in 2020. Currently, works related to the construction of the CNG supply station are underway [12].

3 Assumptions for the analysis

3.1 Bus parameters

In order to ensure that the average age of vehicles in use is not too high (in Kielce it was adopted at 8 years), it is necessary to replace end-of-life buses with new vehicles. The aim of the analysis was to develop variants of bus fleet modernization for Kielce. Based on these, it was determined how a given scenario would affect:

- the structure of the fleet,
- the emission level of harmful substances contained in exhaust gases - carbon dioxide (CO_2), nitrogen oxides (NO_x), particulate matters (PM_x) and volatile organic compounds (VOC),
- the Life cycle cost (LCC) for the entire bus fleet.

This analysis assumes that the buses have a service life of 15 years and an annual mileage of 60,000 km. All cost are displayed in euro according to the conversion rate published by the central bank of the Republic of Poland on 24 February 2020 (1 EUR = 4.28 PLN) [13]. It was assumed that the price of diesel oil was 1.15 €/dm^3, compressed natural gas (CNG) - €0.68/Ndm^3 and electricity - €0.15/kWh [14]. In the paper, the aforementioned prices of fuels,

Table 1 *Bus data adopted for the analysis*

	Diesel	Diesel Euro 6	CNG	EV (200 kWh)	HEV (11.6 kWh)
Purchase price [€]	210 300	210 300	240 000	590 000	350 500
Maintenance cost [€/ year] [15]	3 500	3 500	3 700	3 000	3 700
Cost of replacing a battery pack [€] [16]				140 200	8 130
Average fuel consumption [17]	53.2 dm^3/100km	50.0 dm^3/100 km	50.9 kg/100km	140 kW/100 km	46.5 l/100km

Table 2 *Bus replacement schedule under scenario no. 1*

		2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Number of buses to be exchanged		10	1	0	5	30	24	29	4	27	15	4
Number of purchased buses	Diesel (Euro 6)	-	-	-	-	30	24	29	4	27	15	4
	CNG	10	1	0	5	-	-	-	-	-	-	-
Share of low-emission buses in the fleet		19%	19%	19%	22%	22%	22%	22%	22%	22%	22%	22%

energy and the cost of replacing batteries in EV and HEV vehicles were treated as fixed prices. Table 1 shows the other data adopted for the analysis.

The CNG-fueled vehicles and those equipped with electric drive systems require additional infrastructure. For electric buses, there are two main battery charging methods: fast charging by means of a pantograph located at terminuses or stops and slow charging by means of a plug-in, carried out mainly at depots. The cost of installing a pantograph charger was assumed to be PLN 500,000, while the cost of a plug-in charger was assumed to be €23 365[18]. It was assumed that the cost of building a compressed natural gas supply station would amount to €21 million. A CNG station has fast and slow refueling points [19-20].

The GREET program (Greenhouse gases, Regulated Emissions and Energy use in Transportation Model) was used to estimate emission of the harmful exhaust gas compounds. The program was developed by Argonne National Laboratory (ANL) as a part of a project run by the United States Department of Energy. The GREET program makes it possible to estimate the environmental impact of the life cycle of vehicles equipped with conventional and alternative drive systems. The GREET database contains parameters of 80 different vehicles. Thanks to the interactive interface and the graphical toolkit, simulations can be carried out easily. The program uses data provided by the U.S. Environmental Protection Agency (EPA) [21]. The calculations bring the following data:

- value of energy from combustion of fuel (oil, petrol, gas, coal) and from renewable sources (biomass, wind, solar radiation, water),
- emission level of greenhouse gases (CO_2, CH_4, N_2O), harmful compounds contained in exhaust gases (CO, NO_x, PM_x, SO_x, aliphatic and aromatic hydrocarbons),
- water consumption.

In the GREET program, the life cycle of a vehicle is divided into the stage of fuel production (including acquisition and refinement of crude oil, production, distribution and storage of fuel), the stage of vehicle production (production of parts and components and assembly of the vehicle) and the stage of vehicle operation and use. The program enables emissions of harmful substances and greenhouse gases to be estimated at each of the above-mentioned stages. The emission values of carbon dioxide (CO_2), nitrogen oxides (NO_x), particulate matters (PM_x) and volatile organic compounds (VOC) are presented in the form of costs, in accordance with the rates contained in the European Parliament and of The Council Directive [22].

3.2 Fleet modernization scenarios

The analysis period is 11 consecutive years, i.e. 2019-2030. As mentioned earlier, the vehicle fleet currently (at the end of 2019) consists of 188 buses and this was adopted as a fixed value for the following years. For each of the years covered by the analysis, the total Life Cycle Cost (LCC) and level of emissions for the entire bus fleet were estimated. The assumed service life of vehicles is 15 years, meaning that they must be taken out of service upon reaching this age and new vehicles must be purchased in

Table 3 *Bus replacement schedule under scenario no. 2*

		2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Number of buses to be replaced			1	0	5	30	24	29	4	27	15	4
Number of purchased buses	Diesel (Euro 6)	-	-	-	-	-	-	29	4	27	15	4
	CNG	10	1	0	5	30	24	-	-	-	-	-
Share of low-emission buses in the fleet			19%	19%	22%	38%	50%	50%	50%	50%	50%	50%

Table 4 *Bus replacement schedule under scenario no. 3*

	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Number of buses to be replaced	10	1	0	5	30	24	29	4	27	15	4
Number of purchased CNG buses	10	1	0	5	30	24	29	4	27	15	4
Share of low-emission buses in the fleet	19%	19%	19%	22%	38%	51%	66%	68%	82%	90%	93%

Table 5 *Bus replacement schedule under scenario no. 4*

		2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Number of buses to be replaced		10	1	0	5	30	24	29	4	27	15	4
Number of purchased buses	CNG	5	1	0	2	15	12	15	2	13	8	2
	EV	5	0	0	3	15	12	14	2	14	7	2
Share of low-emission buses in the fleet		19%	19%	19%	22%	38%	51%	66%	68%	82%	90%	93%
Fast chargers (pantograph)		2	2		2	8	8	8	2	8	6	2
Slow chargers (plug-in)		4	1		6	12	12	14	2	14	7	2

their place. The replacement of vehicles was considered according to 5 scenarios.

In **scenario no. 1**, the minimum share of low-emission buses in the fleet by 2025, i.e. 20%, was adopted. Kielce currently operates 25 buses with a hybrid drive system (HEV), which accounts for 13% of the fleet. This variant assumes replacement of 7% of the oldest buses with conventional drive systems by buses with CNG fueled engines. Once the target share of low-emission buses is reached, the remaining vehicles to be replaced can be substituted by conventional buses. However, they will have diesel engines meeting the EURO 6 emission standard (Table 2). Scenario no. 1 also assumes the construction of a compressed natural gas supply infrastructure.

Scenario no. 2 assumes a 50% share (until 2025) of low-emission buses. In 2019, the share of hybrid buses (HEV) in the fleet of the Kielce carrier amounted to 13%. This variant includes the purchase of CNG buses so that they account for 37% of the fleet (Table 3). It is also necessary to build infrastructure designed for refueling buses with compressed natural gas (CNG) engines. Once the 50% share of low-emission buses is achieved, the remaining vehicles to be replaced due their service life end will be substituted by conventional buses with diesel engines meeting the EURO 6 emission standard.

Scenario no. 3 assumes a gradual replacement of buses with conventional drive systems with buses equipped with engines powered by the compressed natural gas (CNG) (Table 4). This variant, like the previous ones, assumes the need to build a CNG refueling station. Under this scenario, at the end of 2030, the share of alternatively powered vehicles will amount to 93%.

Scenario no. 4 assumes the replacement of end-of-life conventional buses with buses with electric drive systems (EV) and CNG buses. The assumption is that in 2030 buses of both types will have an equal share in the fleet (Table 5). Scenario no. 4 includes the construction of a compressed natural gas supply station and charging points for electric buses (chargers in the depot and at bus stops). As a result of the modernization of the fleet at the end of 2030, the share of alternatively driven vehicles will amount to 93%.

Scenario no. 5 assumes a gradual replacement of conventional buses with buses equipped with electric drive systems (EV). This variant assumes the need to adapt the

Table 6 *Bus replacement schedule under scenario no. 5*

	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Number of buses to be replaced	10	1	0	5	30	24	29	4	27	15	4
Number of purchased EV buses	10	1	0	5	30	24	29	4	27	15	4
Share of low-emission buses in the fleet	19%	19%	19%	22%	38%	51%	66%	68%	82%	90%	93%
Fast chargers (pantograph)	4	2	6	4	20	16	20	2	18	10	2
Slow chargers (plug-in)	10	1	2	3	30	24	29	4	27	15	4

battery charging infrastructure, consisting of chargers at bus stops and in the depot (Table 6). According to scenario no. 5, in 2030 the share of alternative vehicles in the bus fleet will amount to 93%.

4 Life Cycle Cost (LCC) model

The presented Life Cycle Cost (LCC) analysis covers the economic and environmental aspects of buses throughout their life cycle. An LCC model in the following form was used for calculations:

$$LCC = \sum_{n=1}^{N} (C_A + C_F + C_M + C_B + C_{Inf} + C_E), \quad (1)$$

where:
$n \in (1,2, \ldots, N)$ - number of vehicles,
LCC - life cycle cost,
C_A - costs of purchase,
C_F - costs of fuel,
C_M - costs of maintenance,
C_B - costs of replacing a battery pack,
C_{Inf} - costs of infrastructure,
C_E - costs of emissions.

The purchase costs C_A were expressed as follows:

$$C_A = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\frac{P_a}{O_i} \right), \quad (2)$$

where:
$k \in (1,2, \ldots, K)$ - type of a vehicle drive system,
$i \in (1,2, \ldots, I)$ - age of a vehicle,
P_a - purchase price,
O_i - service life.

Costs of fuel C_F:

$$C_F = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\frac{f_c}{100} \cdot P_f \cdot D \right), \quad (3)$$

where:
f_c - average fuel consumption [dm^3/100km, kWh/100km],
P_f - unit price of fuel/energy [PLN/dm^3, PLN/kWh, PLN/dm^3],
D - annual mileage [km].

The costs of maintenance C_M include the costs of insurance and periodic inspections, the costs of replacement of tires and service fluids and the costs of required repairs and removal of defects. In the analysis, the costs of repairs and use take the following form:

$$C_M = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\frac{M}{O_i} \right), \quad (4)$$

where:
M - annual costs of repairs and maintenance of the vehicle.

The operating experience gained so far has shown that energy storage devices have a much shorter service life than buses. It was assumed that a battery pack should be replaced every 6 years, therefore, during the service life of a bus it would be necessary to replace the energy storage device twice. The costs of battery replacement C_B were expressed as follows:

$$C_B = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\frac{P_B \cdot B \cdot A_{i,k}}{O_i} \right), \quad (5)$$

where:
P_B - price of battery replacement [PLN/kWh],
B - energy capacity of batteries [kWh],
$A_{i,k}$ - number of vehicles with i - age and k - type of drive system.

Buses equipped with the CNG engines require construction of special infrastructure. It consists of pumps and specially adapted refueling and storage equipment for compressed natural gas. For electric buses, the infrastructure includes battery charging stations. The cost of infrastructure C_{Inf} can be expressed as follows:

$$C_{Inf} = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\frac{P_C \cdot L_{j,k}}{A_{i,k}} \right), \quad (6)$$

where:
$j \in (1,2, \ldots, J)$ - number of charging/refueling stations,
P_C - price of building a charging/refueling station,
$L_{j,k}$ - number of stations intended for k - type of drive system,
$A_{i,k}$ - number of vehicles with i - age and k - type of drive system.

In this paper, Life Cycle Cost includes environmental costs in the form of emission costs calculated in accordance

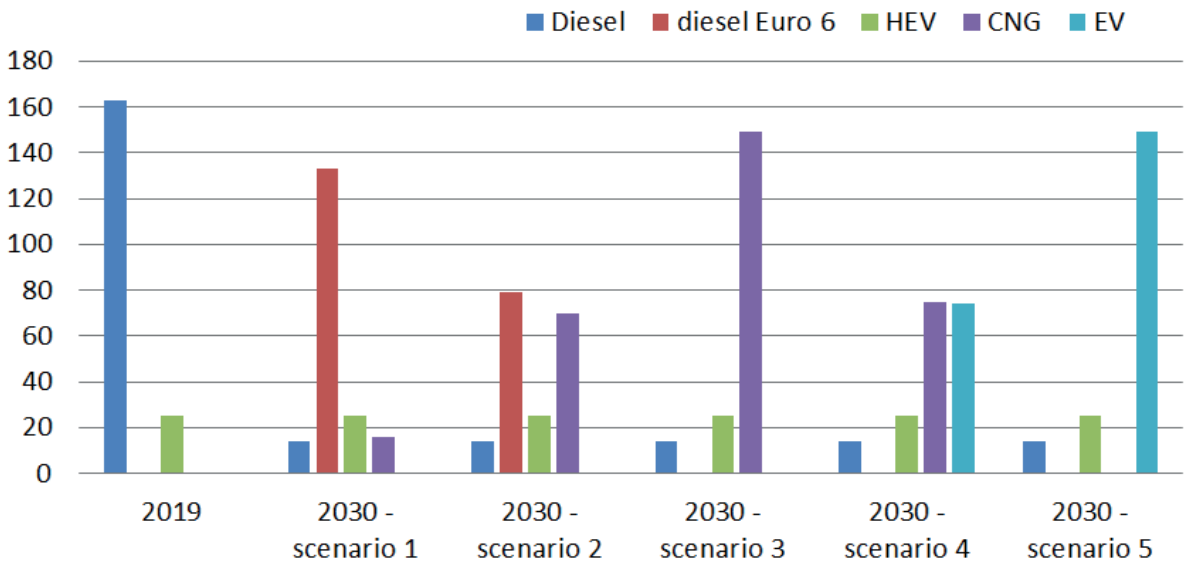


Figure 2 *Fleet structure for the variants described above*

with the rates set out in the European Parliament and of The Council Directive [21]. The costs of emissions C_M take the following form:

$$C_A = \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{z=1}^{Z} \left(\frac{P_z \cdot E_z \cdot D}{O_i} \right), \quad (7)$$

where:
$z \in (1,2, \ldots, Z)$ - harmful substance contained in exhaust gases (e.g. CO_2, NO_x, etc.),
P_z – costs' rate per emission [PLN/kg],
E_z - emission level [kg/km],
D - annual mileage [km].

5 Results of the analysis

5.1 Changes in the structure of the fleet

In 2019, vehicles meeting the European exhaust emission standards of EURO 3 and EURO 4 made up for a significant part of the fleet. Between 2021 and 2024, buses over 10 years of age will account for around 70% of the fleet. In the years 2024-2026, there will be the largest number of buses to be replaced due to their age. During that period it will be necessary to purchase as many as 83 vehicles.

The fleet structure for the analyzed variants is shown in Figure 2.

In scenario no. 1, the assumed 20% share of the low-emission buses in the transport fleet will be reached in 2023. A total of 41 buses with alternative drive systems will be purchased.

In scenario no. 2, the envisaged 50% share of alternative buses will be achieved a bit later, i.e. in 2025. There will be a total of 95 vehicles with alternative drives then. Successive replacement of "old" conventionally powered buses will result in only 14 such buses remaining in 2030 and their share in the fleet will fall to 7%.

According to scenarios no. 1 and 2, in 2030, a significant part of the fleet will still be made up of buses equipped with diesel internal combustion engines. In scenarios no. 3, 4 and 5, the share of buses with alternative drive systems will reach 93% in 2030.

5.2 Emission levels and costs

The main reasons for replacing vehicles with conventional drive systems are economic and ecological aspects. Despite the fact that modern emission reduction systems, such as exhaust gas recirculation system (EGR), selective catalytic reduction system (SCR) with a catalyst, AdBlue (urea water solution) or particulate filter (DPF) are installed in buses, road transport still largely contributes to increase in emissions of nitrogen oxides, carbon dioxide and suspended particulates PM_x in cities.

Figure 3 shows the level of CO_2 emissions in the period under consideration for the simulated scenarios.

The carbon dioxide emission values for the period 2019-2023 are similar for all of the bus replacement variants considered. This is a consequence of the small number of new buses, since only 16 will be replaced by 2023. A faster reduction in the CO_2 emission will take place from 2024 onwards. Between 2024 and 2030, the level of carbon dioxide emission for the analyzed scenarios will vary considerably. The lowest level of CO_2 emission will occur for scenarios no. 4 and 5, which assume the purchase of electric buses. Successive replacement of "old" buses also affects the level of emissions of other harmful compounds, such as NO_x, VOC and PM_x. Figure 4 shows the values of emissions of NO_x, VOC and PM_x.

The emission reduction values of the analyzed exhaust gas components are presented in Table 7.

Scenario no. 1 assumes the lowest share of low-emission vehicles in the transport fleet (20%) among the

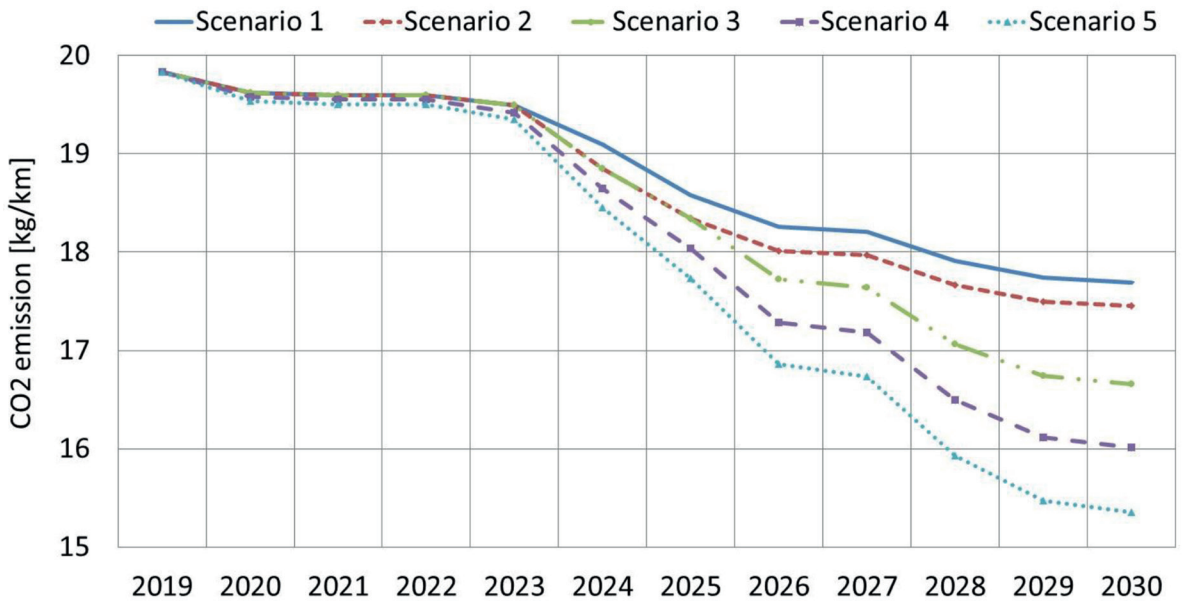


Figure 3 *Level of CO_2 emissions*

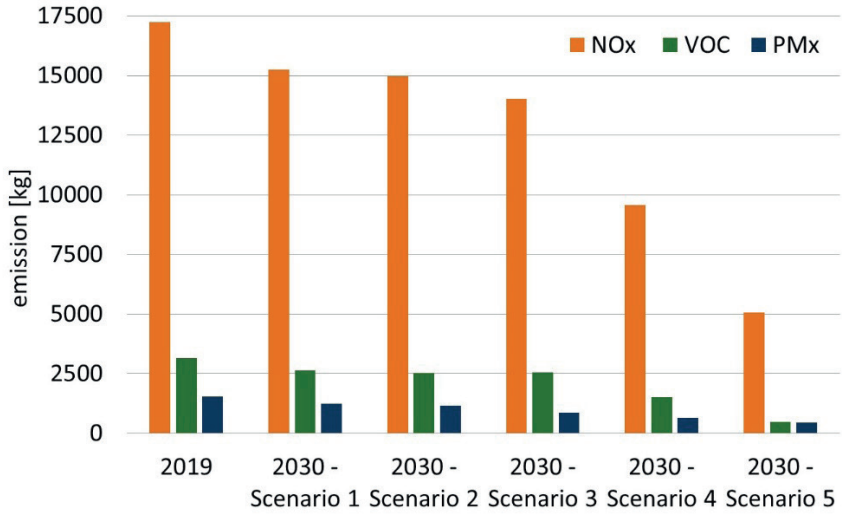


Figure 4 *Emission values of NO_x, VOC and PM_x for the variants considered*

Table 7 *Emission reduction values compared to base year 2019*

	2030 - scenario 1	2030 - scenario 2	2030 - scenario 3	2030 - scenario 4	2030 - scenario 5
CO_2	11%	12%	16%	19%	23%
NO_x	12%	13%	19%	44%	71%
LZO	7%	7%	8%	44%	83%
PM_x	19%	25%	44%	57%	71%

considered variants. In this scenario the CO_2 and NO_x emissions will decrease by relatively small values of 11% and 12%, respectively by 2030, compared to 2019 values. The PM_x and VOC emissions will also decrease by 19% and 7%, respectively (Table 7). Scenario no. 3, which assumes the replacement of buses with a conventional combustion engine by vehicles with CNG engines, will allow a significant reduction in particulate matter emissions. Compared to 2019, the value of PM_x emission in 2030 may be 44% lower.

Scenarios no. 4 and 5, in which the purchase of electric buses is assumed, are characterized by even lower emissions of harmful substances contained in exhaust gases. In 2030, according to scenario no. 4, the level of CO_2 emission will drop by 19%, the level of NO_x and VOC emissions by 44% and the level of PM_x emissions by 57% compared to 2019.

Among the considered variants, the lowest emission of the analyzed harmful substances contained in exhaust gases is provided by scenario no. 5, according to which

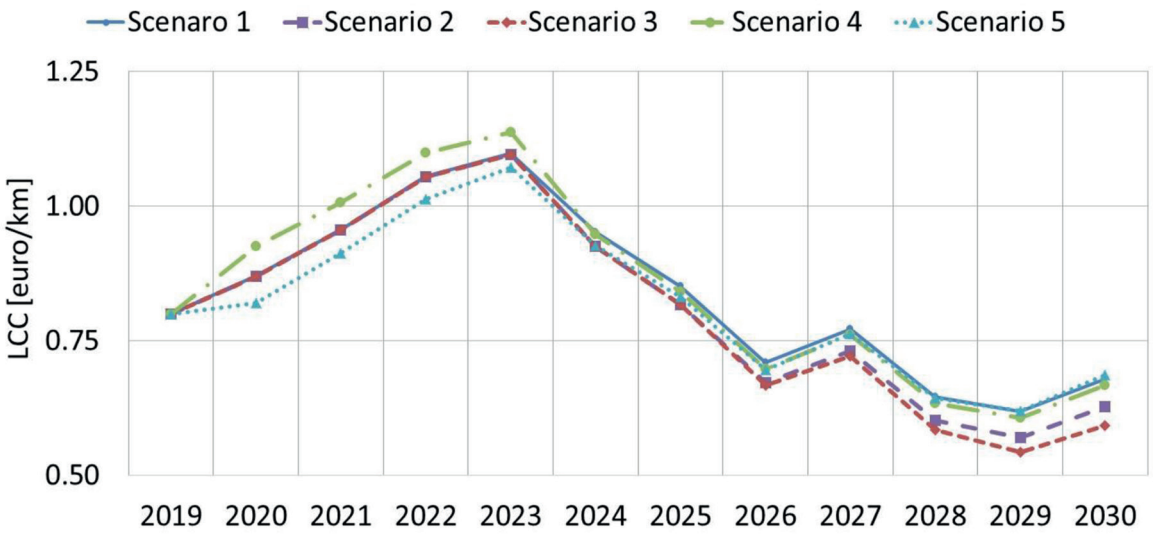


Figure 5 *Change in LCC in the analyzed period*

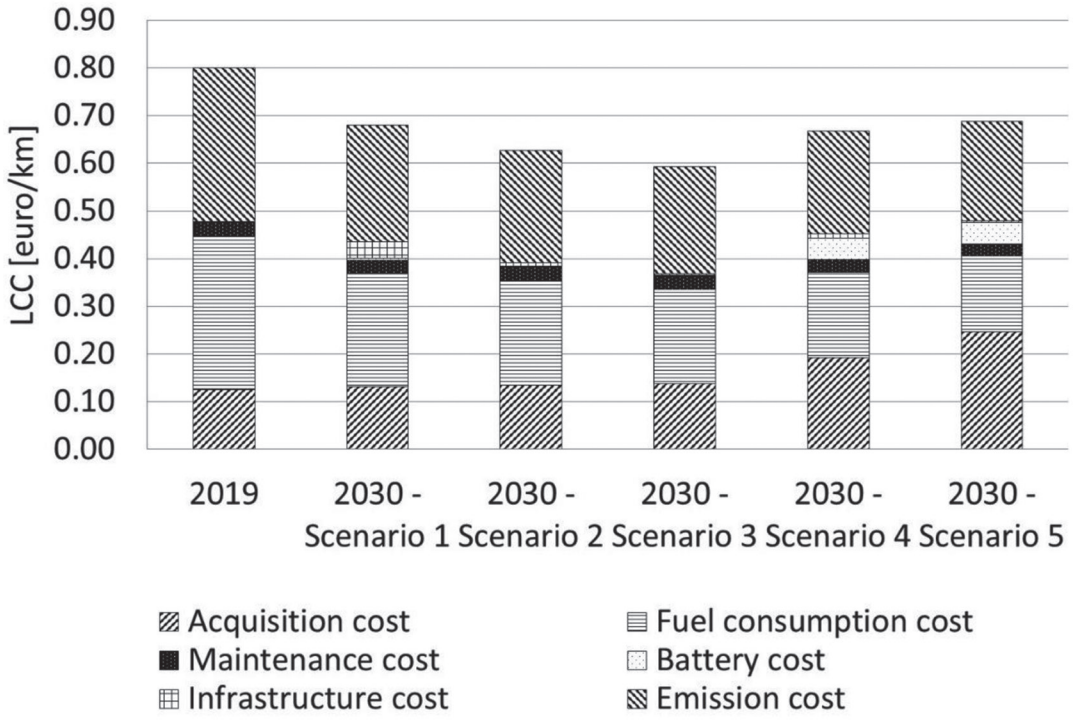


Figure 6 *The LCC summary*

in 2030 CO_2 emissions will drop by 23%, NO_x and PM_x emissions by 71% and the VOC emissions by 83%.

5.3 Life Cycle Cost (LCC) values

An important factor determining the choice of the strategy for modernization of the bus fleet (in accordance with the assumed scenarios) is the cost of purchase of a new vehicle. The cost of purchasing a bus with an engine fueled by compressed natural gas is 10% higher than for a bus with a conventional internal combustion engine. For electric buses, the cost of purchase is 64% higher than for buses with conventional drives.

An important issue related to purchase of alternatively fueled vehicles or vehicles equipped with alternative drive systems is a need to provide appropriate infrastructure. Conventional and hybrid vehicles do not require additional infrastructure outlays.

In the case of operation of electric buses, it is necessary to incur additional costs related to construction of the battery charging stations. In turn, the use of buses with CNG engines requires the construction of storage facilities for compressed gas and refueling stations. This involves significant costs. In the LCC method, infrastructure costs are spread over all vehicles using it. The LCC values

in the analyzed period, estimated in accordance with the considered variants of bus fleet modernization, are presented in Figure 5.

Until 2023, LCC values are similar for all of the scenarios analyzed, as a relatively low number of vehicles would be replaced during that period. The higher LCC values for modernization following scenario no. 3 are due to the need to provide infrastructure for electric and CNG buses. Moreover, in this period the operation of the largest number of vehicles used so far will come to an end. This will result in an increase in the LCC value.

Since 2023, the LCC value has been gradually decreasing due to the replacement of old buses with new ones. Among the analyzed scenarios, the lowest LCC values are found in scenario no. 3, which provides for replacement of buses with conventional engines by vehicles fueled by CNG gas. The LCC value determined for 2030, estimated according to assumptions of variant no. 3, is 27% lower than the LCC level in 2019. The structure of individual LCC components in 2030, based on selected variants, is shown in Figure 6.

The Life Cycle Cost values, estimated based on scenarios assuming a 20% and 50% share of low-emission buses in fleet, set for 2030, are 19% and 20% lower than the LCC level in 2019, respectively. In both variants, the largest share in the LCC is represented by emission costs - 38%. The costs of fuel consumption, which account for 36% of LCC, are also a significant share.

The highest LCC values are found in variants no. 4 and 5, which provide for the purchase of electric buses. The Life Cycle Cost values estimated for variants no. 4 and 5 have the lowest share of fuel consumption costs and emission costs among the LCCs calculated based on the other scenarios. However, the high purchase price, the costly infrastructure and the need to replace batteries make the LCC level estimated for variants no. 4 and 5 significantly higher than for the others. In 2030, according to scenario no. 5, the cost of purchasing a vehicle is 36% of the LCC and the costs of the battery replacement are 7% of the LCC. For variant no. 5, the share of bus purchase costs is 29% of LCC and the cost of the battery replacement is 7% of LCC. The Life Cycle Cost values estimated for variants no. 4 and 5 are 18% and 14% lower than the LCC values in 2019, respectively.

6 Conclusions

With a view to purchasing buses that are alternatively powered or equipped with alternative drive systems, several aspects should be taken into account, such as the purchase costs, operating costs and emission level. Urban public transport companies face the choice of buying alternative or conventional buses and are guided mainly by economic aspects. In general, the costs of purchasing such a vehicle are much higher than the costs of purchasing a bus with a conventional drive system. Unlike the conventional and hybrid vehicles, electric and CNG-fueled buses require infrastructure. However, alternative buses are characterized by the lower operating costs and significantly lower emissions of harmful substances in exhaust gases, such as CO_2, NO_x or PM_x.

The analyses show that the lowest emission level and operating costs can be achieved by increasing the share of electric buses in the fleet. Among the scenarios analyzed, the lowest Life Cycle Cost (LCC) values are represented by the variant that provides for the replacement of buses with conventional drive by the CNG-fueled vehicles.

References

[1] Electro-mobility and alternative fuel act of 11th January 2018. JoL item 317.

[2] POTKANY, M., HLATKA, M., DEBNAR, M., HANZL, J. Comparison of the lifecycle cost structure of electric and diesel buses. *Nase More - International Journal of Maritime Science and Technology* [online]. 2018, **65**(4), p. 270-275. eISSN 1848-6320. Available from: http://www.nasemore.com/vol-65-no-4/

[3] ALLY, J., PRYOR, T. Life cycle costing of diesel, natural gas, hybrid and hydrogen fuel cell bus systems: an Australian case study. *Energy Policy* [online]. 2016, **94**, p. 285-294. ISSN 0301-4215. Available from: https://doi.org/10.1016/j.enpol.2016.03.039

[4] LAJUNEN, A. Lifecycle costs and charging requirements of electric buses with different charging methods. *Journal of Cleaner Production* [online]. 2018, **172**(20), p. 56-67. ISSN 0959-6526. Available from: https://doi.org/10.1016/j.jclepro.2017.10.066

[5] OLSSON, O., GRAUERS, A., PETTERSSON, S. Method to analyze cost effectiveness of different electric bus systems. In: EVS29 - International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium: proceedings. 2016. ISBN 9781510832701, p. 604-615.

[6] TONGA, F., HENDRICKSON, CH., BIEHLER, A., JARAMILLO, P., SEKI, S. Life cycle ownership cost and environmental externality of alternative fuel options for transit buses. *Transportation Research Part D: Transport and Environment* [online]. 2017, **5**(7), p. 287-302. ISSN 1361-9209. Available from: https://doi.org/10.1016/j.trd.2017.09.023

[7] HARRIS, A., SOBAN, D, SMYTH, B. M., BEST, R. Assessing life cycle impacts and the risk and uncertainty of alternative bus technologies. *Renewable and Sustainable Energy Reviews* [online]. 2018, **97**, p. 569-579. ISSN 1364-0321. Available from: https://doi.org/10.1016/j.rser.2018.08.045

[8] SZUMSKA, E., PAWELCZYK, M., PISTEK, V. Evaluation of the Life Cycle Costs for urban buses equipped with conventional and hybrid drive trains. *The Archives of Automotive Engineering - Archiwum Motoryzacji* [online]. 2019, **83**(1), p. 73-86. eISSN 1234-754X. Available from: https://doi.org/10.14669/AM.VOL83.ART5

[9] ISLAM, A., LOWNES, N. When to go electric? A parallel bus fleet replacement study. *Transportation Research Part D: Transport and Environment* [online]. 2019, **72**, p. 299-311. ISSN 1361-9209. Available from: https://doi.org/10.1016/j.trd.2019.05.007

[10] SEN, B., ERCAN, T., TATARI, O., ZHENG, Q. P. Robust Pareto optimal approach to sustainable heavy-duty truck fleet composition. *Resources, Conservation and Recycling* [online]. 2019, **146**, p. 502-513. ISSN 0921-3449. Available from: https://doi.org/10.1016/j.resconrec.2019.03.042

[11] Transport plan for the Kielce commune and adjacent communes forming joint public transport [online]. Available from: http://www.bip.kielce.eu/attachments/1036567/3713969/1.1/Plan_transportowy.pdf

[12] The structure of transport fleet in Kielce [online] [Viewed 2019-09-06]. Available from: http://komunikacja-kielce.pl/organizacja-komunikacji/tabor/

[13] Schedule of exchange rates [online] [Viewed 2020-02-24]. Available from: https://www.nbp.pl/homen.aspx?f=/kursy/RatesC.html

[14] Fuel price in Poland - Information market S. A. [online] [Viewed 2019-09-10]. Available from: www.e-petrol.pl

[15] WYSZOMIRSKI, O., WOLEK, M., JAGIELLO, A., KONIAK, M., BARTLOMIEJCZYK, M., GRZELEC, K., GROMADZKI, M. *Electromobility in public transport* (in Polish) [online]. Polish Development Fund (PFR). Available from: http://pspa.com.pl/assets/uploads/2018/12/Raport_PFR_elektromobilnosc_w_transporcie.pdf

[16] MPK Tarnow tested the electric bus and calculates the disadvantages of such a vehicle (in Polish) [online] [Viewed 2019-09-10]. Available from: https://www.transport-publiczny.pl/wiadomosci/mpk-tarnow-przetestowalo-elektrobus-i-wylicza-wady-takiego-pojazdu-59229.html

[17] Fuel consumption of urban buses equipped with conventional and alternative drives (in Polish) [online] [Viewed 2019-09-10]. Available from: http://wde.warszawa.pl/wp-content/uploads/2019/04/Prezentacja_elektromobilno%C5%9B%C4%87-Vo.pdf

[18] Alternative fuels in public transport [online]. Report 208/2019. Polish Alternative Fuels Association (PSPA). Available from: http://pspa.com.pl

[19] Analysis of costs and benefits related to the use of zero-emission buses in the provision of public transport services (in Polish) [online]. City Transport Board in Gdansk. Available from: https://ztm.gda.pl

[20] GRZELEC, K., OKROJ, D. Perspectives for the use of electric buses in the public transport in the chosen city: Sopot / Perspektywy obslugi miast autobusami elektrycznymi na przykladzie Sopotu (in Polish). *Autobusy: Technika, Eksploatacja, Systemy Transportowe*. 2016, **17**(11), p. 26-32. ISSN 1509-5878.

[21] GREET Life-cycle model. User guide. Center for Transportation Research Energy Systems. Division Argonne National Laboratory, 2016.

[22] Directive 2009/33/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of clean and energy-efficient road transport vehicles.

Antanas Fursenko - Arturas Kilikevicius - Kristina Kilikeviciene - Jonas Skeivalas -Albinas Kasparaitis
Jonas Matijosius - Dariusz Wieckowski

ANALYSIS OF DYNAMIC PARAMETERS OF THE SYSTEM OF RASTER FORMATION AND CONTROL

The presented research work analyzes the sensing system, the main aim of which is a raster formation and controlling this process using the optical measuring equipment and high precision angle encoders. Optical measuring equipment are used for the raster position detection, meanwhile angle encoders for controlling the tape speed. The main parameter of raster formation process is fixed transportation speed, but there are difficulties to realize it, because there is imperfection of the device elements. The article analyzes the dispersion of vibration accelerations of the raster formation device and tape in the two directions (transverse and longitudinal) and presents an analysis of their parameters in application of the theory of covariance functions. The results of the measurements of vibration accelerations at the fixed points of the device constructions and the tape were recorded on a time scale in the form of digital arrays (matrices). Values of auto-covariance and inter-covariance functions of digital arrays of the vibration accelerations measurement data were calculated by changing the quantum interval in a time scale. The developed software Matlab 7 in operator package environment was used in the calculations.

Keywords: *raster formation and controlling sensing system, vibration signal accelerations, dynamic parameters, covariance function, quantum interval*

1 Introduction

Steel tape, on the surface of which certain symbols, such as rasters, special marks and the like are formed, are often used for metrological and technological needs, such as for measuring the displacement. Those symbols are formed by the movement of a steel tape and the beams of the source of light, which affects either the band surface or its special coatings.

The undesired deviations in the position and the shape of symbols formed in that way mainly depend on uncontrolled deviations of the said movement of the tape and the light beam from the set parameters.

One of the most important components of such deviations are vibrations of the band resulting from the sources of excitation of the tape transporting system mechanism and other internal and external vibrations. Uncontrolled changes in their position and the displacement speed, as well as the band deformation that have an adverse effect on the quality of the formed structures, occur as a result of their impact. It also affects the accuracy and reliability of the active control of the process carried out in real time. This effect on the system of defined parameters depends on vibration parameters - frequencies, amplitudes and other statistical characteristics. Knowing these parameters is important for designing the construction and optimizing the processes of formation the construction formation and of the structure of symbols.

The presented research examines the tape transport system consisting of electromechanical tape tractions and its fixed tensioning mechanisms as well as a tape tilt mechanism operating on sliding friction. This system is mounted on a massive granite base built on the foundation using passive vibration insulators. The research and data processing method and results of the experimental study of the layout system are proposed in the article. Experience of this article authors in construction, analysis and optimization of mechatronic precision devices is presented in papers [1-4].

The formation process takes place in the dynamic mode because both the steel tape and the laser raster formation head are continuously moving during the process.

In assessment of the significant points of the construction being examined, analysis of the vibrations dispersion of construction of the device and its tape point was carried out by application of the covariance function theory.

The Gaussian process emulator, with separable covariance function assumption of separability, imposes constraints on the emulator and may negatively affect its performance in some applications where separability may not hold. Zhang et al. [5] propose a multi-output Gaussian

Antanas Fursenko[1], Arturas Kilikevicius[1], Kristina Kilikeviciene[2], Jonas Skeivalas[3], Albinas Kasparaitis[1], Jonas Matijosius[1], Dariusz Wieckowski[4,*]
[1]Institute of Mechanical Science, Vilnius Gediminas Technical University, Lithuania
[2]Department of Mechanical and Material Engineering, Vilnius Gediminas Technical University, Lithuania
[3]Department of Geodesy and Cadastre, Vilnius Gediminas Technical University, Lithuania
[4]Lukasiewicz R&D Network Automotive Industry Institute, Warsaw, Poland
*E-mail of corresponding author: d.wieckowski@pimot.eu

process emulator with a non-separable auto-covariance function to avoid limitations of using the separable emulators.

Guella et al. [6] provide characterization theorems for high dimensional spheres, as well as for the Hilbert sphere.

Kithulgoda et al. [7] present a strategy for incremental maintenance of the Fourier spectrum to changes in concept that take place in data stream environments. Singh and Singh [8] generalized Chebyshev-Fourier moments (G-CHFMs) and generalized pseudo - Jacobi-Fourier moments (G-PJFMs) have been extended to represent color images using quaternion algebra.

Zhang et al. [9] used the spectral densities to calculate the average entropy of mixtures of random density matrices and show that the average entropy of the arithmetic-mean-state of n qubit density matrices randomly chosen from the Hilbert-Schmidt ensemble was never decreasing with the number n. Kurt and Eryigit [10] studied dynamics of the two-state system coupled to an environment with peaked spectral density.

Alotta et al. [11] provided a complete characterization of normal multivariate stochastic vector processes. Band et al. [12] focused in fractal Fourier analysis in which the graphs of the basic functions are Cantor sets, discontinuous at a countable dense set of points, yet have good approximation properties. Zhang et al. [13] simulated fractal signals that were generated under varying signal lengths and scaling exponents that indicate a range of physiologically conceivable fractal signals. Liu et al. [14] aimed to improve the computational accuracy of the fractal dimension of the profile curve.

Florindo and Bruno [15] proposed a novel technique for the numerical calculus of the fractal dimension of fractal objects, which can be represented as a closed contour. Xie et al. [16] proposed a layered online data reconciliation strategy based on a Gaussian mixture model, which is proposed for the complex industrial processes with multiple modes. Odry et al. [17] demonstrated the use of a special test bench that both enables simulations of various dynamic behaviors of wheeled robots and measures their real attitude angles along with the raw sensor data. Song el al. [18] proposed a structure-preserving bilateral texture filtering algorithm to smooth the texture while preserving salient structures.

Du et al. [19] proposed a correlation filter based visual tracking approach that integrates spatial-temporal adaptive feature weights into the DCF method to derive an efficient solution. Li et al. [20] proposed an algorithm for adaptive multiscale morphological filtering (AMMF), and its effect was evaluated by a simulated signal. Nathen et al. [21] presented an extension of a Large Eddy Simulation based approach in the framework of Lattice Boltzmann Methods.

During the research, an analysis of the raster formation device construction and dispersion of vibration accelerations of its tape points was carried out.

The tape transport system consists of electromechanical tape traction and its fixed tensile mechanisms, as well as a tape thrust mechanism operating on slip friction. The main parameter of the raster formation process is fixed transportation speed, but there are difficulties to realize it, because there is some imperfection of the device elements. This method may be used to produce a precision metrological scale on a stainless steel tape. The covariance function theory was used in analysis of the vibration parameters. Graphical expressions of the covariance functions show change of probabilistic inter-dependence of vibration parameters in the timeline. This degree of change of the vibration parameters spatial relations depends on technological and dynamic properties of the tape device being examined.

During the experiment, four sensors were laid out on the structural elements of the device and the fifth sensor was placed on the tape of the device. Vibration accelerations of all the five sensors were measured in the longitudinal and transverse directions of the tape movement.

2 Simulation of vibration parameters

The research presents an option for calculating the most reliable trend value of the vibrations vector applying the method of the least squares. It is assumed that the vibration vector trend is a discrete quantity with a fixed value. Use of the least squares method partially eliminates random vibration errors. When processing the large-scale measurement data, the least squares method also provides asymptotically effective values of the calculated parameters in the case where the measurement data distribution is not normal.

The array of measurements of vibration accelerations consists of 5 vectors φ (columns). Each vector is understood as a random function with random measurement errors. The most reliable trend value $\tilde{\varphi}$ which is also called a weighted average, is calculated for each vector φ in application of the least squares method. An assumption that the trend value of the vector changes according to the harmonic law, where the forecasted wave length corresponds to the length of the vibration accelerations vector, is used. The parameter equation of the individual vector value φ_i reads:

$$\varepsilon_i = \varphi_i - a_i\tilde{\varphi}\,, \tag{1}$$

where ε_i is a random acceleration error, φ_i is an acceleration value and $\tilde{\varphi}$ is an acceleration vector trend. Coefficient a_i is expressed as:

$$a_i = \cos\Delta_i \tag{2}$$

where $\Delta_i = \Delta \cdot i, \Delta = 2\pi/n$ is the value of the unit of measure, rad, $i = 1,2,\ldots,n$.

The following is the equation (1) in the matrix form:

$$\varepsilon = \varphi - A\tilde{\varphi}, \tag{3}$$

where ε is a random error vector, $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)^T$ is a vibration accelerations vector, $A = (a_1, a_2, \ldots, a_n)^T$ is a matrix of coefficients of parameter equations $(n \times 1)$.

The most reliable trend value of the vibration accelerations vector φ is calculated in application of the condition of the least squares method:

$$\phi = \varepsilon^T P \varepsilon = \min, \tag{4}$$

where P is the diagonal matrix of weights of vibration accelerations $\varphi_i (n \times n)$

Weights of individual acceleration values φ_i are calculated according to formula

$$p_i = \frac{\sigma_0^2}{\sigma_{\varphi i}^2}, \tag{5}$$

where σ_0 is the standard deviation of the measurement result φ_0 the weight whereof is accepted as equal to one $p_0 = 1$. Thus, the σ_0 value is selected freely, because it has no impact on the calculation results. Value of the measurement result φ_0 is selected so that weights p_i were close to one (to reduce the scope of calculations).

Out of t equation

$$u_i = \ln \varphi_i \tag{6}$$

the following is derived:

$$\sigma_{\varphi i} = \sigma_{u i} \varphi_i, \tag{7}$$

Equation (7) shows that $\sigma_{\varphi i}$ value depends on value of the vibration accelerations φ_i. Thus, the acceleration of a higher value is of lower accuracy, because $\varphi_i \gg \sigma_{ui}$.

Using Equation (5), the following expression is obtained:

$$p_i = \frac{\sigma_0^2}{\sigma_{ui}^2 \varphi_i^2} = \varphi_i^{-2} \cdot 10^{-7}, \tag{8}$$

where the accepted average value . $\frac{\sigma_0^2}{\sigma_{ui}^2} = 10^{-7}$

The extremity of the function (4) is determined having calculated its partial derivatives according to the parameter $\tilde{\varphi}$, equated it to zero and solved the equation received:

$$\frac{\partial \Phi}{\partial \tilde{\varphi}} = 2\left(\frac{\partial \varepsilon}{\partial \tilde{\varphi}}\right)^T P \cdot \varepsilon = 0 \tag{9}$$

Then, the following is obtained:

$$-A^T P \varepsilon = 0 \tag{10}$$

and

$$A^T P A \tilde{\varphi} - A^T P \varphi = 0. \tag{11}$$

The solution is equal to:

$$\tilde{\varphi} = (A^T P A)^{-1} A^T P \varphi = N^{-1} \omega, \tag{12}$$

where

$$N = (A^T P A)^{-1}, \tag{13}$$

$$\omega = A^T P \varphi. \tag{14}$$

The accuracy of the parameter values calculated in application of the least squares method is assessed using their covariance matrix $K_{\tilde{\varphi}}$. Since the trend of cases of this task is a quantity (a scalar), the following may be written:

$$K'_{\tilde{\phi}} = \sigma'^2_{\tilde{\phi}} = \sigma'^2_0 N^{-1}, \tag{15}$$

where σ'_0 is a value of the standard deviation σ_0. It is assessed according to the formula

$$\sigma'^2_0 = \frac{1}{n-1} \varepsilon^T P \varepsilon. \tag{16}$$

3 Object of the research and measurement results

This article examines a raster formation device (Figure 1). The research work analyzes the tape transport system consisting of the electromechanical tape traction and its fixed tensile mechanisms, as well as a tape tilt mechanism operating on sliding friction. Vibrations in two directions (longitudinally and transversely) were measured at 4 points of the raster device (point 1 - fixed tension mechanism of the tape; point 2 - tape displacement measurement system; point 3- tilting node; point 4 - traction - tension system) and on the metal tape.

The block diagram of the metal scale production device presented in Figure 1 comprises the following: 1 - tape retractor mechanism; 2 - granite base; 3 - system measuring the tape displacement (point 2); 4 - tilting node (point 3); 5 - error monitoring node; 6 - traction - tension system (point 4); 7 - tape retractor mechanism; 8 - mechanism covering the tape in protective film; 9 - tape tension control and management system; 10 - fixed tension mechanism of the tape (point 1); 11 - optic system for turning laser beam microscopes with CCD camera; 12 - scanner; 13 - laser.

The articles analyzes vibration signals of significant points of the metal scale production device, when the device is in operating mode. Measurements were conducted in the two directions X and Y (Figure 1). Figure 2 (Y direction) and Figure 4 (X direction) illustrate results of acceleration of these measurement points and the spectral density graphs of these accelerations are respectively presented in Figures 3 and 5.

Analysis of results, illustrated in Figures 2-5, reveals that the highest level of vibrations was observed in the traction - tension system (Figure 1, point 4) when assessing the vibrations of the longitudinal direction Y (up to 32 mm/s^2); when assessing the vibrations of the transverse direction X, accordingly, the highest level of vibrations was observed in the fixed tension mechanism of the tape (Figure 1, point 1) and was about 190 mm/s^2 (the main components

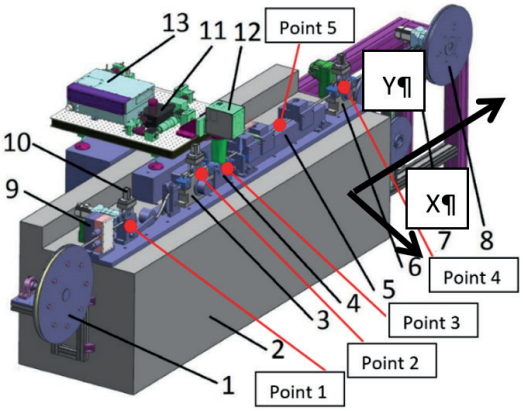
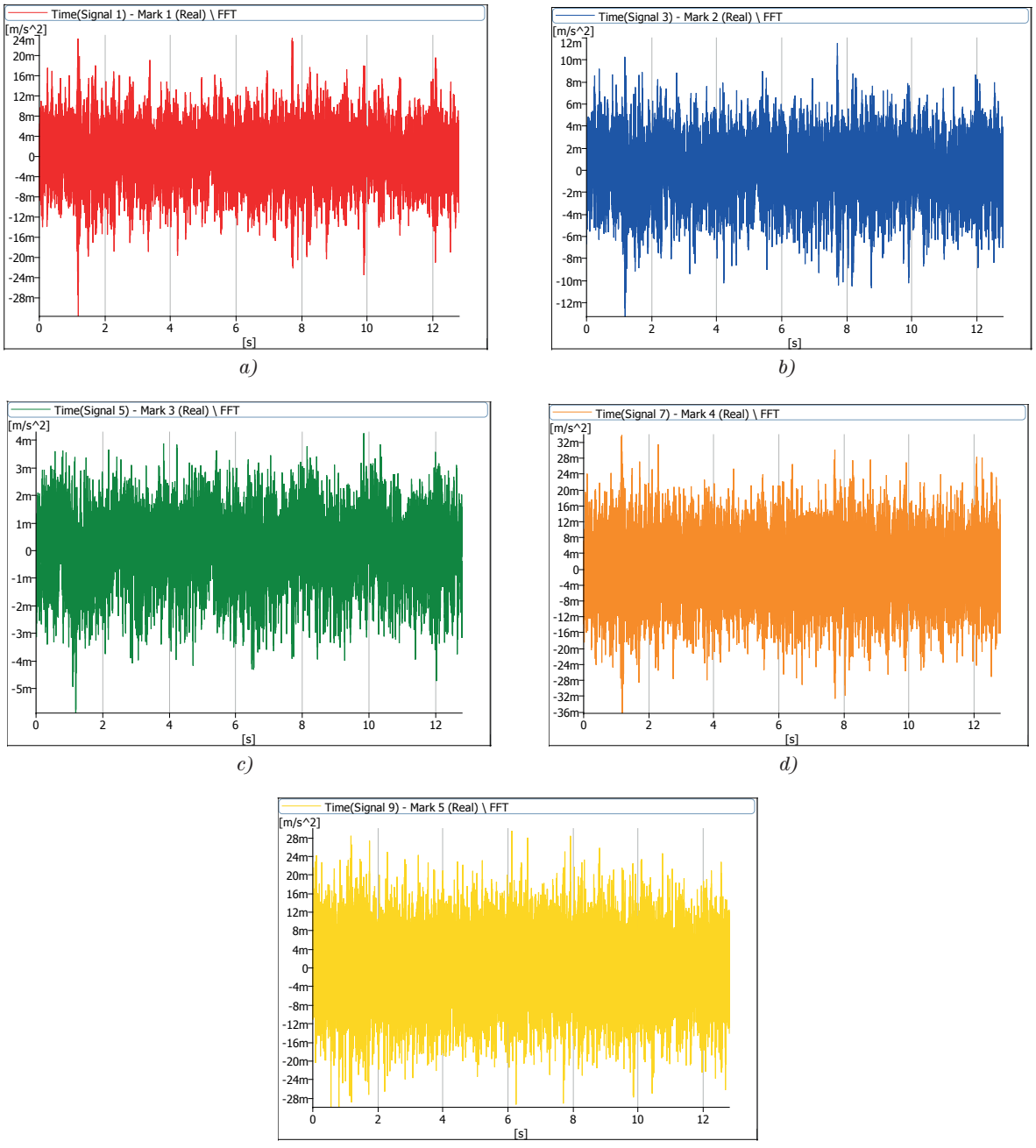


Figure 1 *Block diagram of the metal scale production device with five measurement points*



a) *b)* *c)* *d)* *e)*

Figure 2 *Graphs of time signal for the Y direction accelerations of measurement points (Figure 1): a) red - point 1; b) blue - point 2; c) green - point 3; d) orange - point 4; e) yellow - point 5*

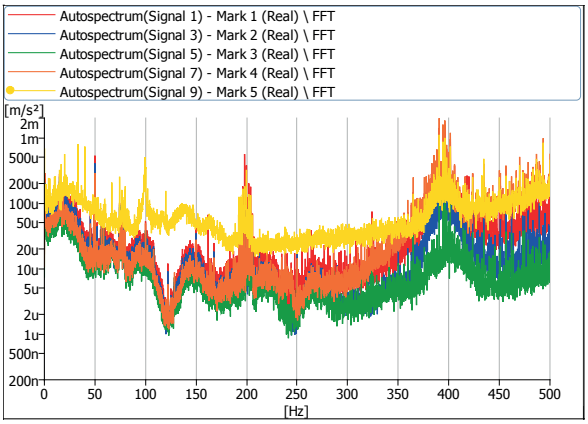


Figure 3 *Graphs of spectral density for accelerations (Y direction in Figure 1) of measurement points (red - point 1; blue - point 2; green - point 3; orange - point 4; yellow - point 5)*

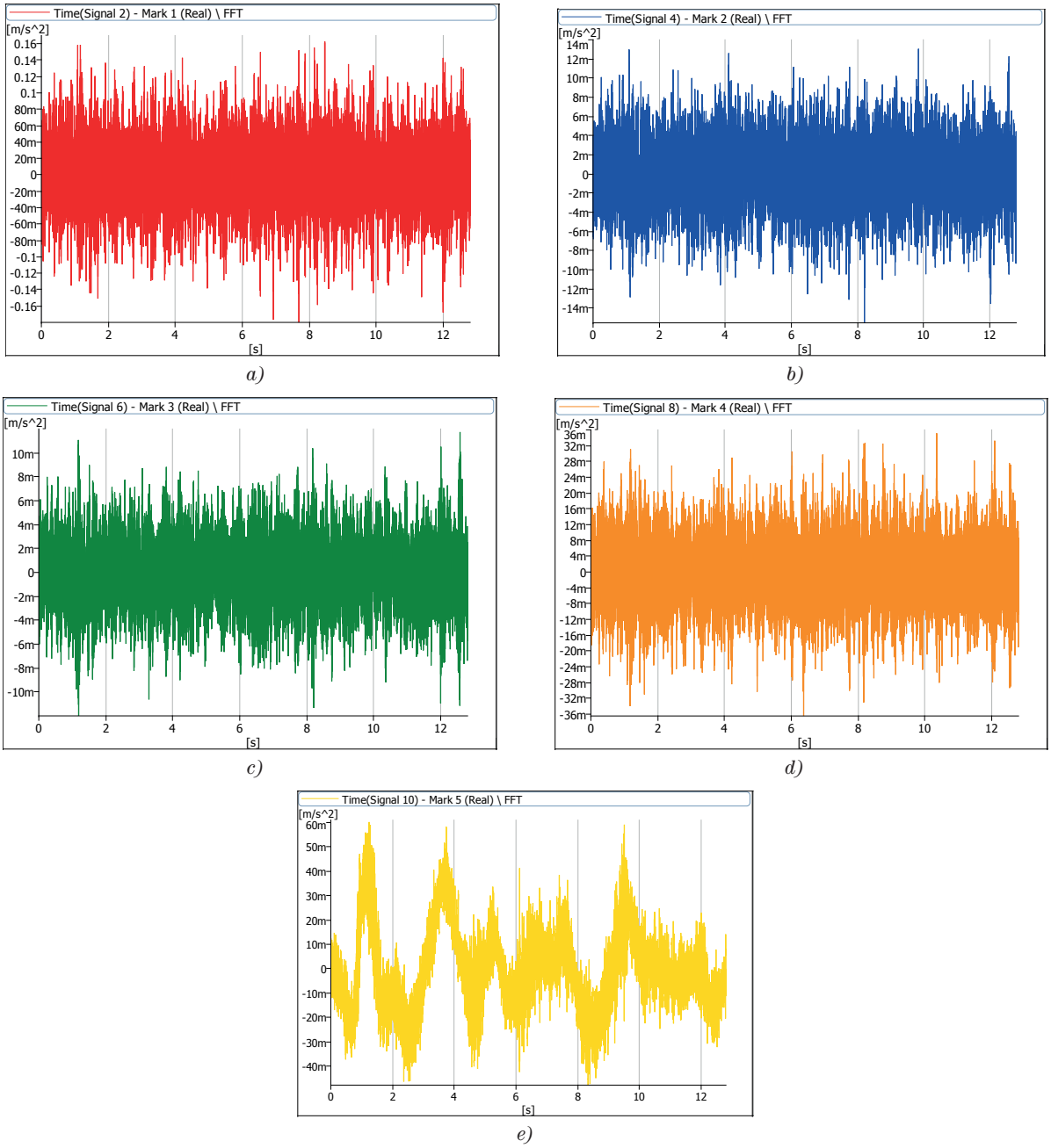


Figure 4 *Graphs of time signal for the X direction accelerations of measurement points (Figure 1): a) red - point 1; b) blue - point 2; c) green - point 3; d) orange - point 4; e) yellow - point 5*

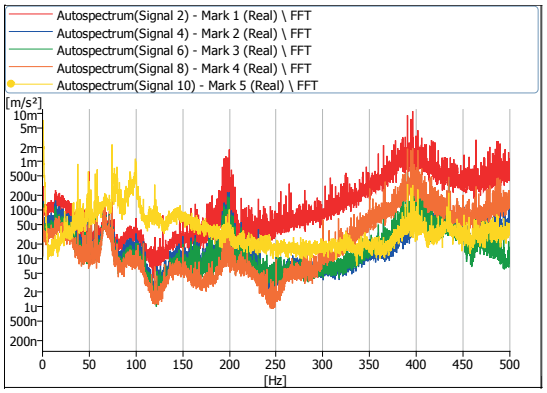


Figure 5 *Graphs of spectral density for accelerations (X direction in Figure 1) of measurement points (red - point 1; blue - point 2; green - point 3; orange - point 4; yellow - point 5)*

Table 1 *Numbering of array vectors in the sequence order*

Acceleration vector No.	Direction of the measurement of accelerations relative to the tape movement direction
1	longitudinal
2	transverse
3	longitudinal
4	transverse
5	longitudinal
6	transverse
7	longitudinal
8	transverse
9	longitudinal
10	transverse

were in the frequency range from 280 to 420 Hz from Figure 5). The lowest level of vibrations in the Y and X directions was in the tape tilting node (Figure 1, point 3), where the formation of rasters takes place and it was 6 (in the Y direction) and 12 mm/s² (in the X direction).

4 Covariance model of vibration signal parameters and the results of the analysis of the vibrations acceleration model

Guchhait and Banerjee [22] proposed a variant of constitutive equation error based material parameter estimation procedure for linear elastic plates, which was developed from partially measured free vibration signatures.

The conception of a stationary random function is the scientific ground of this theoretical model. Considering that, errors of measuring the bridge vibration parameters are random and of the same accurateness, ie. the average error $M\Delta = const \to 0$ the dispersion $D\Delta = const$. The covariance function of digital signals depends on the difference between the arguments only, ie. on the quantisation interval on the time scale [23].

The theoretical model was a subject of an assumption that errors of digital signals of vibrations are random and possibly systematic. The measurement data trend of the vector is eliminated in each vector of measurement arrays of vibration parameters. Vibration dispersion time interval is used as one of parameters. Tables 1 and 2 illustrate the schemes of the numbering of measurement arrays of the vibration accelerations vector. The quality of data of accelerations of all the five vectors was assessed using their accuracy indicator, a standard deviation. Standard deviation values are presented in Table 2.

Auto-covariance function of one data array or inter-covariance function of the two arrays $K_\varphi(\tau)$ is expressed as [24-25]:

$$K_\varphi(\tau) = M\{\varepsilon\overline{\overline{\varphi}}_1(u)\cdot\varepsilon\overline{\overline{\varphi}}_2(u+\tau)\} \tag{17}$$

or

$$K_\varphi(\tau) = \frac{1}{T-\tau}\int_0^{T-\tau}\varepsilon\varphi_1(u)\varepsilon\varphi_2(u+\tau)du \tag{18}$$

where $\varepsilon\varphi_1 = \varphi_1 - \widetilde{\varphi_1}$, $\varepsilon\varphi_2 = \varphi_2 - \widetilde{\varphi_2}$ are centred measurement vectors of vibration parameters φ with eliminated trend, u is a vibrations parameter, $\tau = \tau\cdot\Delta$ is a variable quantum range, k is a number of the measurement units, Δ is the value of the measurement unit, T is time and M is a symbol of the mean.

The covariance model of vibration signal parameters which authors used in this article is described in detail in papers [3, 23].

Arrays of the vibration accelerations measurements' data were obtained using 1-axis accelerometers 8344 and a

Table 2 *Accuracy indicators of the installation construction and vectors of vibration accelerations of the tape points*

Name of vectors	Standard deviation value of single measurement result of vectors σ'_i, m/s^2
	Array
1	0.006
2	0.041
3	0.003
4	0.003
5	0.001
6	0.003
7	0.008
8	0.017
9	0.008
10	0.017

	Vector parameter	Value m/s^2
1	Standard deviation value σ'_0 of the result with the weight of p = 1	0.01
6	Standard deviation value σ'_0 of the result with the weight of p = 1	0.01
1	Standard deviation value σ'_v of the weighted average of vectors	$1{\cdot}10^{-7}$
6	Standard deviation value σ'_v of the weighted average of vectors	$2{\cdot}10^{-7}$

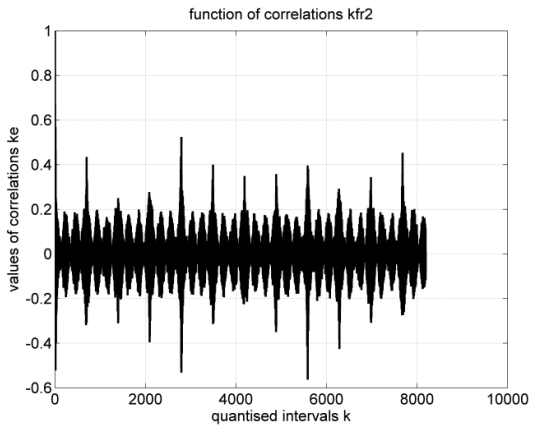


Figure 6 *Simplified auto-covariance function of the 2nd vibration accelerations vector of the array*

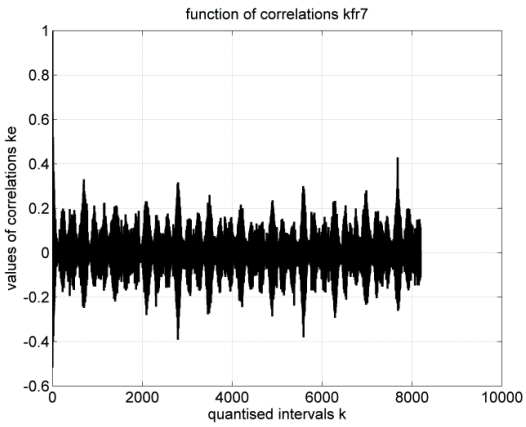


Figure 7 *Simplified auto-covariance function of the 7th vibration accelerations vector of the array*

tri-axial accelerometer 4506. Signals were recorded in the course of the time intervals $\tau_\Delta = 0.00078\,s$ and 12.8 s. Values of $n = 16386$ vibration signal accelerations were fit in each array vector.

Expression of each measurement vector is a random function due to random measurement errors when bringing its expression closer to the shape of a stationary function, trend components, which were calculated in application of the method of least squares, were eliminated.

Measurement data arrays were processed using the developed computer software in application of Matlab programme package operators.

Values of the quantised interval of simplified covariance functions range from 1 to $n/2 = 8000$. The value $K'_\varphi(\tau)$ of the simplified auto-covariance function $K_\varphi(\tau)$ was calculated for each vibrations vector and graphical expressions of 16 simplified auto-covariance functions were obtained.

Values of inter-normalized covariance functions $K'_\varphi(\tau)$s were also calculated based on 10 vibration array vectors and their graphical expressions were obtained according to their respective combinations.

Figures 6-17 illustrate more important graphical expressions.

Figure 18 presents the graphical image of the generalized (spatial) correlation matrix of the array of 8 vectors of vibration accelerations. Expression of the correlation matrix gets the block shape of 8 pyramids where the correlation coefficients values are shown in shades of the colour spectrum. The horizontal plane shows the colour projection of the pyramids.

Graphical expressions of the simplified auto-covariance and inter-covariance functions are of the two-tier shape. This is related to the fact that the structure of vibration signals is a composite function with vibrations of large amplitudes with long intervals and vibrations of small

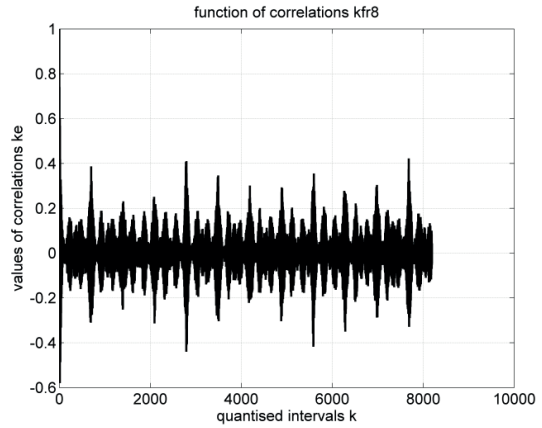


Figure 8 *Simplified auto-covariance function of the 8th vibration accelerations vector of the array*

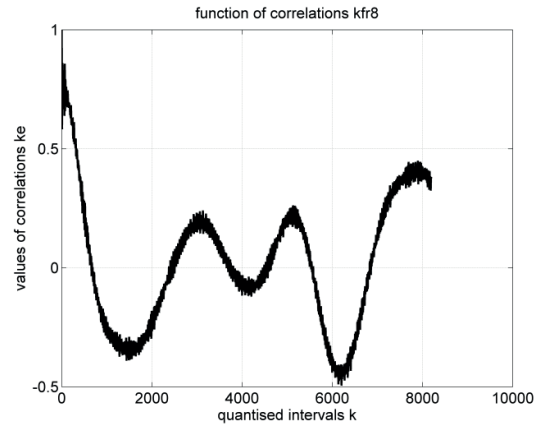


Figure 9 *Simplified auto-covariance function of the 10(8) th vibration accelerations vector of the array*

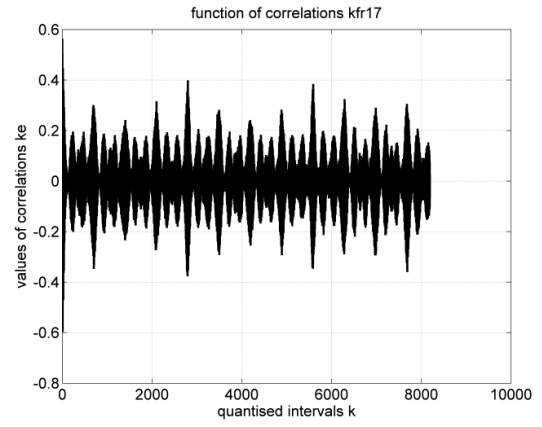


Figure 10 *Simplified inter-covariance function of the 1st and 7th vibration accelerations vectors of the array*

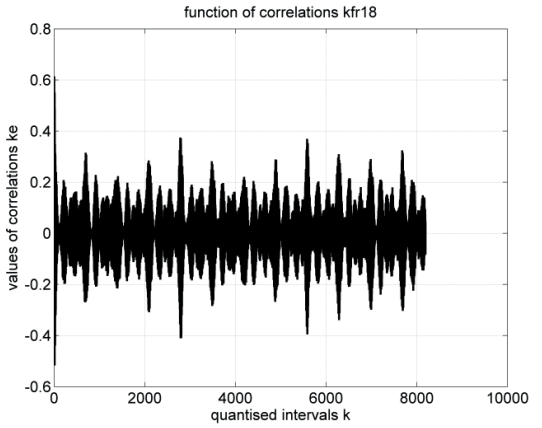


Figure 11 *Simplified inter-covariance function of the 1st and 8th vibration accelerations vectors of the array*

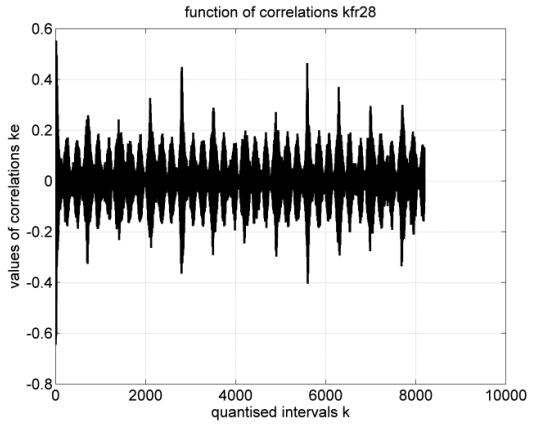


Figure 12 *Simplified inter-covariance function of the 2nd and 8th vibration accelerations vectors of the array*

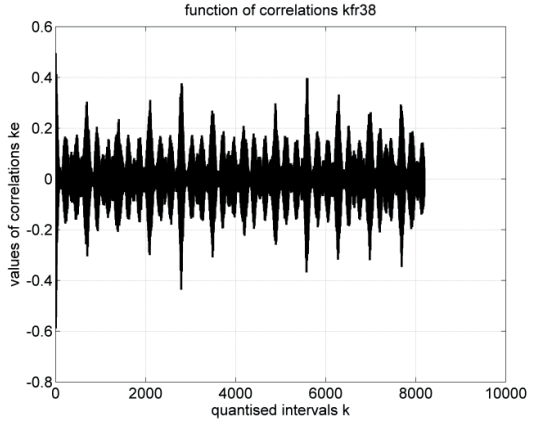


Figure 13 *Simplified inter-covariance function of the 3rd and 8th vibration accelerations vectors of the array*

amplitudes. This determines a two-tier expression of the covariance functions.

The simplified auto-covariance functions of vectors of all accelerations of the arrays get the highest correlation coefficient values $r \to 1$ at values $(\tau_k \to 0\,s)$ of the quantised interval and then decrease to $r \to 0$ at $k \to 10(\tau_k \to 0.0078\,s)$ as a result of vibrations of small amplitudes and at $k \to 1000(\tau_k \to 0.78\,s)$ as a result of vibrations of large amplitudes. Expressions of simplified auto-covariance functions of acceleration vectors of the arrays are similar.

The 10th simplified array auto-covariance function of the transverse vibrations accelerations vector of the tape has an expression that is close to the theoretical shape.

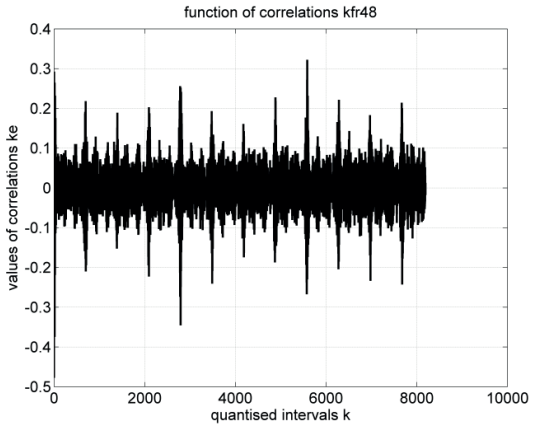


Figure 14 *Simplified inter-covariance function of the 4th and 8th vibration accelerations vectors of the array*

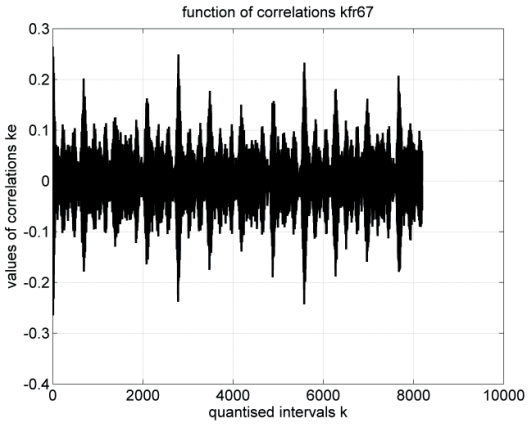


Figure 16 *Simplified inter-covariance function of the 6th and 9(7)th vibration accelerations vectors of the array*

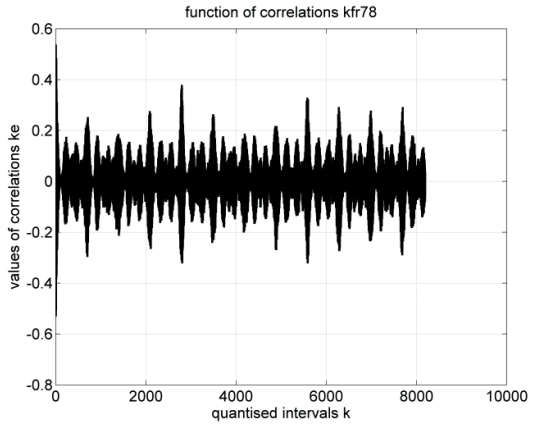


Figure 15 *Simplified inter-covariance function of the 7th and 8th vibration accelerations vectors of the array*

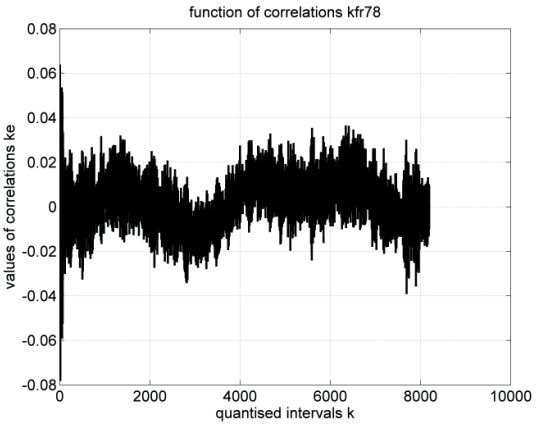


Figure 17 *Simplified inter-covariance function of the 9(7)th and 10(8)th vibration accelerations vectors of the array*

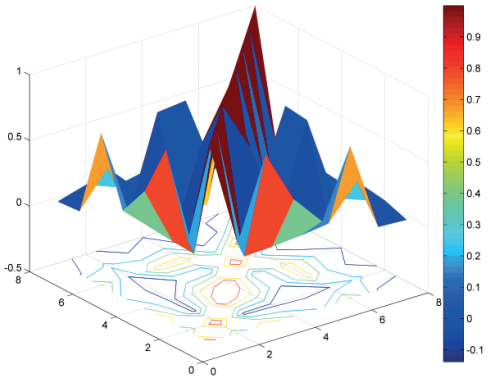


Figure 18 *Graphical image of the generalized (spatial) correlation matrix of the 1st array of 8 vibration accelerations vectors*

The simplified inter-covariance functions of the acceleration vectors of the array have average correlation coefficient values for the majority of pairs of vectors ranging in the 0.4-0.8 interval, which is illustrated by the data of inter-correlation of the 1st acceleration vector with vectors of all the other 7 accelerations and between vectors 2 and 3, 2 and 7, 2 and 8, 3 and 6, 3 and 7, 3 and 8, and 6 and 8. Thus the correlation between the majority of accelerations vectors is significant enough, which shows potentially inter-related vibrations of nodes of the device construction.

5 Conclusions

Analysis of the dynamic parameters of the sensing system construction, the main aim of which is the raster formation and controlling this process, using the optical

measuring equipment and high precision angle encoders, revealed that the highest level of vibrations of up to 32 mm/s^2 was in the traction - tension system, when assessing the longitudinal vibrations; when assessing vibrations in the transverse direction the highest level of up to 190 mm/s^2 was formed in the fixed tension mechanism of the tape. Main components of the vibration acceleration amplitudes were in the frequency range from 280 to 420 Hz.

The simplified auto-covariance and inter-covariance functions show the change of the probabilistic inter-dependence of vibration parameters of the raster formation tape device in the time scale. This degree of changes depends on the relation between vibrations of the structural nodes of the device being examined and the dynamic properties.

Graphical expressions of simplified auto-covariance and inter-covariance functions are two-tiered, which is related to the fact that the structure of vibration signals is a composite function that has vibrations of large amplitudes with long intervals and vibrations of relatively small amplitudes. Expressions of simplified auto-covariance functions of accelerations vectors of the arrays are similar.

The 10th simplified auto-covariance function of the vector of transverse vibration accelerations of the tape has a different expression than other auto-covariance functions.

The simplified inter-covariance functions of accelerations vectors of the array have the average correlation coefficient values ranging within the 0.4-0.8 interval in the majority of pairs of vectors. Thus, the correlation between the majority of pairs of accelerations vectors is significant enough, which shows the possibly inter-related vibrations of construction nodes of the device.

References

[1] KILIKEVICIUS, A., KASPARAITIS, A. Dynamic research of multi-body mechanical systems of angle measurement. *International Journal of Precision Engineering and Manufacturing* [online]. 2017, **18**(8), p. 1065-1073. ISSN 2234-7593, eISSN 2005-4602. Available from: https://doi.org/10.1007/s12541-017-0125-1

[2] SIAUDINYTE, L., KILIKEVICIUS, A., SABAITIS, D., GRATTAN, K. T. V. Modal analysis and experimental research into improved centering-leveling devices. *Measurement* [online]. 2016, **88**, p. 9-17. ISSN 0263-2241. Available from: https://doi.org/10.1016/j.measurement.2016.01.044

[3] JUREVICIUS, M., SKEIVALAS, J., KILIKEVICIUS, A., TURLA, V. Vibrational analysis of length comparator. *Measurement* [online]. 2017, **103**, p. 10-17. ISSN 0263-2241. Available from: https://doi.org/10.1016/j.measurement.2017.02.010

[4] KILIKEVICIUS, A. KASPARAITIS, A., LAZDINAS, R., KILIKEVICIENE, K. Kinematic and dynamic errors during the measurement of linear displacements using the angle transducers. *Mechanics* [online]. 2016, **22**(3), p. 212-216. ISSN 1392-1207. Available from: https://doi.org/10.5755/j01.mech.22.3.14180

[5] ZHANG, B., KONOMI, B. A., SANG, H., KARAGIANNIS, G., LIN, G. Full scale multi-output Gaussian process emulator with nonseparable auto-covariance functions. *Journal of computational physics* [online]. 2015, **300**, p. 623-642. ISSN 0021-9991. Available from: https://doi.org/10.1016/j.jcp.2015.08.006

[6] GUELLA, J. C., MENEGATTO, V. A., PORCU, E. Strictly positive definite multivariate covariance functions on spheres. *Journal of Multivariate Analysis* [online]. 2018, **166**, p. 150-159. ISSN 0047-259X. Available from: https://doi.org/10.1016/j.jmva.2018.03.001

[7] KITHULGODA, C. I., PEARS, R., NAEEM, M. A. The incremental Fourier classifier: leveraging the discrete Fourier transform for classifying high speed data streams. *Expert Systems with Applications* [online]. 2018, **97**, p. 1-17. ISSN 0957-4174. Available from: https://doi.org/10.1016/j.eswa.2017.12.023

[8] SINGH, C., SINGH, J. Quaternion generalized Chebyshev-Fourier and pseudo-Jacobi-Fourier moments for color object recognition. *Optics and Laser Technology* [online]. 2018, **106**, p. 234-250. ISSN 0030-3992. Available from: https://doi.org/10.1016/j.optlastec.2018.03.033

[9] ZHANG, L., WANG, J., CHEN, Z. Spectral density of mixtures of random density matrices for qubits. *Physics Letters A* [online]. 2018, **382**(23), p. 1516-1523. ISSN 0375-9601. Available from: https://doi.org/10.1016/j.physleta.2018.04.018

[10] KURT, A., ERYIGIT, R. Dynamics of a spin-boson model with structured spectral density. *Physics Letters A* [online]. 2018, **382**(19), p. 1262-1267. ISSN 0375-9601. Available from: https://doi.org/10.1016/j.physleta.2018.03.021

[11] ALOTTA, G., DI PAOLA, M., PINNOLA, F. P. Cross-correlation and cross-power spectral density representation by complex spectral moments. *International Journal of Non-Linear Mechanics* [online]. 2017, **94**, p. 20-27. ISSN 0020-7462. Available from: https://doi.org/10.1016/j.ijnonlinmec.2017.02.001

[12] BANDT, C., BARNSLEY, M. HEGLAND, M., VINCE, A. Old wine in fractal bottles I: Orthogonal expansions on self-referential spaces via fractal transformations. *Chaos Solitons and Fractals* [online]. 2016, **91**, p. 478-489. ISSN 0960-0779. Available from: https://doi.org/10.1016/j.chaos.2016.07.007

[13] ZHANG, Z., VANSWEARINGEN, J., BRACH, J. S., PERERA, S., SEJDIC, E. Most suitable mother wavelet for the analysis of fractal properties of stride interval time series via the average wavelet coefficient method. *Computers in Biology and Medicine* [online]. 2017, **80**, p. 175-184. ISSN 0010-4825. Available from: https://doi.org/10.1016/j.compbiomed.2016.11.009

[14] LIU, Y., WANG, Y., CHEN, X., ZHANG, C., TAN, Y. Two-stage method for fractal dimension calculation of the mechanical equipment rough surface profile based on fractal theory. *Chaos Solitons and Fractals* [online]. 2017, **104**, p. 495-502. ISSN 0960-0779. Available from: https://doi.org/10.1016/j.chaos.2017.09.012

[15] FLORINDO, J. B., BRUNO, O. M. Closed contour fractal dimension estimation by the Fourier transform. *Chaos Solitons and Fractals* [online]. 2011, **44**(10), p. 851-861. ISSN 0960-0779. Available from: https://doi.org/10.1016/j.chaos.2011.07.008

[16] XIE, S., YANG, C., YUAN, X. WANG, X., XIE, Y. Layered online data reconciliation strategy with multiple modes for industrial processes. *Control Engineering Practice* [online]. 2018, **77**, p. 63-72. ISSN 0967-0661. Available from: https://doi.org/10.1016/j.conengprac.2018.05.002

[17] ODRY, A., FULLER, R., RUDAS, I. J., ODRY, P. Kalman filter for mobile-robot attitude estimation: Novel optimized and adaptive solutions. *Mechanical Systems and Signal Processing* [online]. 2018, **110**, p. 569-589. ISSN 0888-3270. Available from: https://doi.org/10.1016/j.ymssp.2018.03.053

[18] SONG, C., XIAO, C., LI, X., LI, J., SUI, H. Structure-preserving texture filtering for adaptive image smoothing. *Journal of Visual Languages and Computing* [online]. 2018, **45**, p. 17-23. ISSN 1045-926X. Available from: https://doi.org/10.1016/j.jvlc.2018.02.002

[19] DU, F., LIU, P., ZHAO, W., TANG, X. Spatial-temporal adaptive feature weighted correlation filter for visual tracking. *Signal Processing: Image Communication* [online]. 2018, **67**, p. 58-70. ISSN 0923-5965. Available from: https://doi.org/10.1016/j.image.2018.05.013

[20] LI, Y., ZUO, M. J., LIN, J., LIU, J. Fault detection method for railway wheel flat using an adaptive multiscale morphological filter. *Mechanical Systems and Signal Processing* [online]. 2017, **84**, p. 642-658. ISSN 0888-3270, eISSN 1096-1216. Available from: https://doi.org/10.1016/j.ymssp.2016.07.009

[21] NATHEN, P., HAUSSMANN, M., KRAUSE, M. J., ADAMS, N. A. Adaptive filtering for the simulation of turbulent flows with lattice Boltzmann methods. *Computers and Fluids* [online]. 2018, **172**, p. 510-523. ISSN 0045-7930. Available from: https://doi.org/10.1016/j.compfluid.2018.03.042

[22] GUCHHAIT, S., BANERJEE, B. Constitutive error based parameter estimation technique for plate structures using free vibration signatures. *Journal of Sound and Vibration* [online]. 2018, **419**, p. 302-317. ISSN 0022-460X. Available from: https://doi.org/10.1016/j.jsv.2018.01.020

[23] SKEIVALAS, J., JUREVICIUS, M., KILIKEVICIUS, A., TURLA, V. An analysis of footbridge vibration parameters. *Measurement* [online]. 2015, **66**, p. 222-228. ISSN 0263-2241. Available from: https://doi.org/10.1016/j.measurement.2015.02.034

[24] LI, Y., MULANI, S. B., SCOTT, K. M. L., KAPANIA, R. K., WU, S., FEI, Q. Non-stationary random vibration analysis of multi degree systems using auto-covariance orthogonal decomposition. *Journal of Sound and Vibration* [online]. 2016, **372**, p. 147-167. ISSN 0022-460X. Available from: https://doi.org/10.1016/j.jsv.2016.02.018

[25] BAI, Z., WANG, C. A note on the limiting spectral distribution of a symmetrized auto-cross covariance matrix. *Statistics and Probability Letters* [online]. 2015, **96**, p. 333-340. ISSN 0167-7152. Available from: https://doi.org/10.1016/j.spl.2014.10.002

Salavat Mudarisov - Ilshat Gainullin - Ildar Gabitov - Eduard Khasanov

IMPROVEMENT OF TRACTION INDICATORS OF A TRACK-CHAIN TRACTOR

The aim of this research was to conduct the comparative traction tests for T-170M1.03-55 tractor with a flat and elliptical rim. Structurally rational geometry of the crawler's support surface for tractors with semi-rigid suspension is realized by placing the support rollers at different heights relative to the cart. The results of traction tests showed that elliptical track rim has increased the maximum traction power by 10.4%, conditional traction propulsive efficiency coefficient to 7.43% and the specific traction effort by 8%. The increase in indicators is provided by a lower rolling resistance of a tractor with an elliptical rim. Reduction of the resistance power to rolling of the tractor with an ellipse track rim occurs due to alignment of support rollers vertical load and reduction of resistance to rollers movement on internal contours of tracks and in hinges of track chain links. The results of the research indicate a significant improvement in traction performance of T-170M1.03.55 tractors with elliptical track-chain rim.

Keywords: *track-chain tractor, traction, power balance, drawbar tests; rolling coefficient*

1 Introduction

When performing technological operations in agriculture, a track-chain tractor is an effective pulling device [1]. Track-chain tractors in comparison to wheeled tractors have better traction, better off-road capabilities, low soil compaction indicators, 8-20% lower fuel consumption per unit of work performed [2-4]. Along with the advantages, a chain-track tractor has a number of disadvantages that worsen its performance [5].

At present, much attention is paid to improving the technical level of tractors and track assembly traction indicators, as well as to the reduction of metal content and soil compaction [4-6]. Based on results of the in-depth analysis of the mover's impact on soil, new methods for determining the maximum pressure of wheeled and track-chain movers on soil were developed, as well as calculation methods for defining the indicators of wheel and metal track-chain movers, which guarantee permissible machinery impact on soil, and stress-strain state of soils [2].

One of the ways to increase traction and reduce soil compaction is the tractor movers' improvement [7]. In tractor construction, more than twelve mover designs are used. Most of these movers have semi-rigid suspension [1, 4-5]. The use of rubber-reinforced tracks is a promising direction for improving the track-chain travel system. Rubber-reinforced tracks have both, advantages and disadvantages compared to metal track-chain movers [8-11].

Models of interaction between wheels and deformable gravel developed for wheeled vehicles are often used [12]. Influence of air pressure in the front and rear wheels on the rolling coefficient of the wheeled tractor and fuel consumption have been determined [13]. An improved model of tire-ground interaction has been proposed based on FEA-SPH modeling. [14]. A method for estimating a three-dimensional (3D) footprint of pneumatic agricultural tires has been developed based on the tire footprint molding with liquid plaster. Using a 3D scanner, the molds were then converted to three-dimensional models [15]. A model based on particle filtration was proposed to estimate thermomechanical parameters of wheel-ground interaction [16].

The models of wheel interaction with the ground [12, 15-16], which were proposed, can be used to model the interaction of rubber-reinforced tracks with soil [12-16]. The issue of improving traction properties is of a great importance for all types of track-type vehicles. An empirical model of traction characteristics of rubber tracks on agricultural soils has been developed [7]. A method for calculating crawler track traction on soft ground has been presented [17]. A simple general method has been proposed for calculating soil deformations by the track of track-type vehicles [18]. The traction characteristics of seafloor-tracked vehicles were evaluated based on the mechanical laboratory tests [19].

Interaction of steel and rubber crawler movers with the soil, the uneven distribution of soil reactions on the support surface of the mover, traction characteristics and calculations are considered in different studies [7, 17-19]. The impact of the mover on the soil increases together with the rise in uneven load distribution along the support part, which results in crop yields reduction [20-29]. Uneven distribution of loads between the lower support rollers of a track-chain tractor is accompanied by uneven wear

Salavat Mudarisov*, Ilshat Gainullin, Ildar Gabitov, Eduard Khasanov
Federal State Budgetary Educational Establishment of Higher Education Bashkir State Agrarian University, Ufa, Russian Federation
*E-mail of corresponding author: salamuda12@yahoo.com

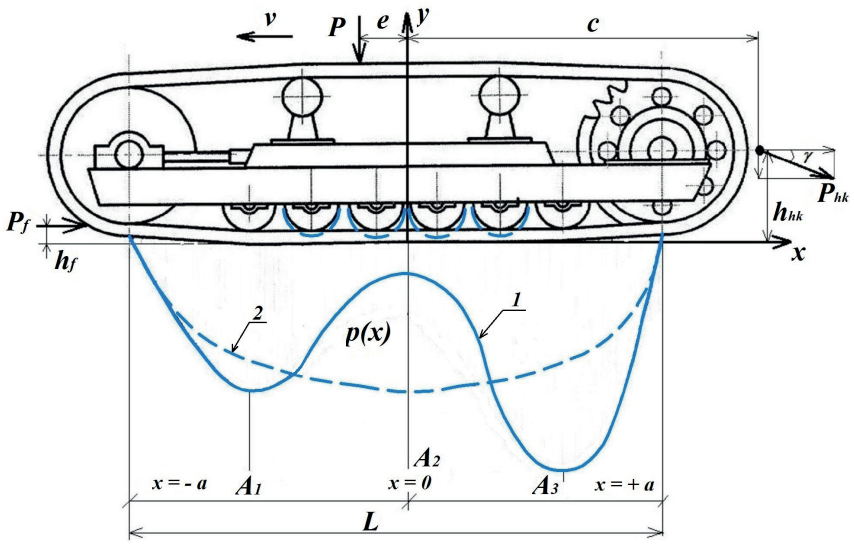


Figure 1 *Scheme of a track-chain mover and the stress distribution diagram in the longitudinal section of the support part of T-170.M1.03-55 tractor (1 - flat track rim; 2 - ellipse track rim): P [N] - load on a single mover; P_{hk} [N] – traction effort on the tractor hook; e-eccentricity [m]; L [m] - length of the support part; |±a| [m] - half-length of the support part; h_f [m] - shift of the longitudinal component of the rolling power P_f from relative to soil reaction; c [m] - distance from the contact center to the vertical component P_{hk}; h_{hk} [m] - the height of the trailer relative to the support surface*

of support rollers, tracks and a decrease in the mover durability [30].

During the experimental studies of the impact of T-170M1.03-55 tractor movers on soil it was found that the maximum values of its pressures reached 0.166 MPa [3]. At the same time, the pressure profile along the length of the support surface has two local extrema in the zone of the first and the sixth support rollers (Figure 1), which causes a decrease in the traction performance and increased soil compaction. It is more preferable to reduce the uneven load distribution along the support part, the maximum pressure and multiplicity of influences, than its uniform distribution on the site of contact of the basic surface. The pressure profile can be significantly decreased by changing the geometry of the tractor support part [31-33]. The purpose of this research was to evaluate influence of the geometry of the support part of a track-chain mover on the traction characteristics of the T-170M1.03-55 tractor.

2 Methods

Traction characteristics of a tractor are the basis in correct choosing of parameters of a machine-tractor aggregate. Using the traction characteristics, a theoretical analysis was carried out and potential energy and technical and economic performance of the machine-tractor aggregate were estimated [34]. In accordance with the requirements of the State All-Union Standard (SAUS) 23734, 25836, 7057, comparative traction tests of T-170M1.03-55 tractor with flat and elliptical track-chain rim *were carried out, taking into account scientific works [34-36]. This type of tractor is widely used in* the *post-Soviet countries [37-38].*

3 Equation of the support part geometry of a track-chain tractor

As theoretical studies have shown, a significant reduction in the maximum pressure ensures its uniform distribution over the contact area of the tractor support surface with the soil. Based on the contact problem of the theory of elasticity, an equation of the geometry of the support surface of a track-chain tractor with a semi-rigid suspension, providing a uniform distribution of pressure along the support surface, was obtained [25, 38]:

$$f(x) = 0.5p_{av}\pi\beta\{x\arcsin(x/a) + A - \\ -(B[xA/a^2 + \arcsin(x/a)]\} + C,' \quad (1)$$

where:

p_{av} [Pa] is the average pressure of the tractor on soil;

$p_{av} = G_e/(2bL)$;

G_e [N] is the operating tractor weight;

L [m] is the length of the tractor support surface;

b [m] is the track width;

$\beta = v_1 + v_2$;

$v_1 = (2(1-\mu_1^2)/\pi E_1$;

$v_2 = (2(1-\mu_2^2)/\pi E_2$;

E_1 [Pa] is the soil elasticity modulus;

μ_1 is the soil Poisson's ratio;

E_2 [Pa] is the steel elasticity modulus of a track link;

μ_2 is the Poisson's ratio of the track link steel;

$A[m] = \sqrt{a^2 - x^2}$;

$a = L/2$ is the contact half-width;

x [m] - the horizontal coordinate of the support surface point;

$B[\text{kN}] = P[e + \varphi_{hk}(h_{hk}\cos\gamma + c\sin\gamma) + fh_f]$;

$P = G_e + P_{hk}\cos\gamma$ - the load to a single mover;

P_{hk} [N] - the power on the hook;

Table 1 *The soil and atmospheric conditions during the traction tests*

Indicators	Track-chain mover	Transmission gear						
		1	2	3	4	5	6	7
Intake air temperature inside the air cleaner [°C]	elliptical	28	26	30	26	29	27	27
	flat	36	36	40	39	29	27	34
Diesel fuel temperature [°C]	elliptical	33						
	flat	42						
Atmospheric pressure [kPa]	elliptical	98						
	flat	97						
Soil density number of strokes	elliptical	7 ... 10						
	flat	7 ... 10						
Soil moisture [%]	elliptical	at a depth of 8 cm - 10%, at a depth of 12 cm - 15%						
	flat							
Track slope (in the movement direction/ against movement direction) [%]	elliptical	0.4/2.5						
	flat							

γ [◦] is the angle between the power on the hook and the horizontal plane;

$\varphi_{hk} = P_{hk}/P$ - the coefficient of the hitch weight use;

e [m] is the longitudinal coordinate of the gravitational center of the tractor relative to the middle of the track length on ground;

h_{hk} [m] - the height of the trailer relative to the support surface;

f - the resistance factor to the tractor movement, f = 0.07-0.15;

h_f [m] - the shift of the longitudinal component of the rolling power relative to the soil reaction, h_f = 0.015-0.029;

C [m] is the coefficient equal to the initial soil deformation determined empirically, C = - 0.027±0.003.

The established dependence $f(x)$ shows how far the points of the tractor support part are separated from the OX horizontal axis. Structurally elliptical geometry of the support part is implemented by lowering the axis of the less loaded lower support rollers, that is, by setting plates of appropriate thickness under the axis of lower support rollers. Thickness of the plates is determined by Equation (1).

The elliptical geometry of the support part depends on the following parameters: tractor weight, support surface length, location of the center of mass, tractive power and its loading point, tracks material properties and the soil type. Therefore, the T-170M1. 03. 55 tractor parameters were determined and its typical working conditions were: traction power is 40-80 kN, the length of the support surface is 2.88 m, the trailer height is 0.4 m; eccentricity e = - 0.165 m; track material, which is carbon steel (steel 20, steel 45, etc.); loamy soil; soil moisture is 16-20%.

All the parameters have been structurally applied for the T-170M1.03.55 tractor in the following way: when position of the 1st and 6th lower support rollers remain unchanged, the 2nd and 5th support rollers are lowered by use of spacers for 5 mm, the 3rd and 4th for 10 mm relative to the support surface of the cart frame. The same tractor was used during the traction tests, but the geometry of the mover support part was changed.

4 Ground and atmospheric conditions during the traction tests

The traction tests were carried out at the test base of Chelyabinsk tractor factory. The length of the traction track was measured to be 600 m, and the width was 25 m. Density and moisture of the soil of the track were measured on testing day at least 12 times. The density was measured with DorNII ram tester. Control of slopes and deviations from the track flatness was carried out by theodolite 2T30, special racks and a metal ruler. Atmospheric conditions were checked daily at the beginning and end of the tests, as well as during the experiments for determining the maximum traction power and maximum traction effort. The soil and atmospheric conditions during the traction tests are presented in Table 1.

Figure 2 *T-170M1.03-55 tractor with a mobile dynamic laboratory SDL-30*

Figure 5 *Registration of track driving wheel speed rate*

Figure 3 *Tractive power registration*

Figure 6 *Path measurement device*

Figure 4 *Engine speed registration*

Figure 7 *Brake tests of D-160 diesel engine*

5 Pilot unit and measuring

In accordance with the instruction manual, the undertaken maintenance and the run-in process lasted for 150 Moto-hours. At the beginning of testing, the tractor was weighed on a weighing complex with limits of 5-100 tons.

Resistance temperature devices were installed to measure the air temperature at the inlet to the air filter and the temperature of the fuel entering the fuel filter. The tractor was equipped by sensors and devices, which were a part of the measuring systems for a movable dynamic laboratory SDL-30 (Figures 2-6).

Parameters to be registered, measuring instruments, and measurement errors are presented in Table 2.

Calibration of a traction link was performed before and after the tests. Adjustment of the measuring instruments, which are a part of the measuring system of laboratory SDL-30, was done.

Characteristics of the D-160 engine were determined on a hydraulic brake stand using the breaking method in accordance with the State All-Union standard 25836-83, 23734-79 and 18509-88. At the front of the tractor, the engine crankshaft was attached to the E-1500 brake stand as follows. The rotor shaft of the balancing machine was connected to the main shaft through a pin bush coupling. The other end of the main shaft was connected to the engine through the driveshaft and the mid-shaft, which was mounted on a bearing support on the tractor bumper and connected to the engine crankshaft grooved sleeve (Figure 7).

All experiments were carried out in the steady traction load mode with straight-line movement of the tractor on the track, created by the SDL-30 laboratory through the trailer E-900 with a straight-gage element. After the mode stabilization, loading was carried out in stages with fixed values of P_{hk} for at least 7.5 seconds. The traction characteristics was measured on forward movement gears

Table 2 *List of parameters to be registered and measuring instruments*

Measured parameter	Measurement instrument	Measurement range	Measurement error
Traction power [kN]	Strain-gauge-type segment E-1810-2 Electro-magnetic transducer EMT-P	1…150	not more than 0.66 %
Crankshaft speed [min-1]	E-753-1SB sensor Electro-magnetic transducer EMT-P	0…1500	not more than 0.5 %
Driven wheel speed [min-1]	contact sensor E-1344 Electro-magnetic transducer EMT-P	0...60	not more than 0.5 %
Tractor trip mileage [m]	contact sensor E-985-14 Electro-magnetic transducer EMT-P	0...200	not more than 0.5 %
Experiment time [s]	Electro-magnetic transducer EMT-P timer	0…7.5	0.2 s
Fuel temperature, ambient air temperature, intake air temperature inside the air cleaner, [°C]	Multipoint electro-thermometer 234.00.000 consisting of: - resistance copper thermometer; - measuring unit. Thermometer TL-2	-50…+40	±1°C
Fuel density [t·m-3]	Density hydrometer ANT-1	0.77...0.83	±0.0005 t/m3
Atmospheric pressure [kPa]	Meteorological aneroid barometer BAMM-1	80…106	±0.2 kPa
Soil density, the number of strikes	DorNII ram tester	-	± 1 strike
Track slope [°]	Theodolite 2T30	+60…-55	±30”

in the form of a series of experiments covering the entire range of traction effort s in each gear.

Based on the measured and decoded values, the following indicators were calculated for each gear:

V_T [m·s^{-1}] - theoretical speed; V_a [m·s^{-1}] - actual speed; δ [%] – sliding.

The theoretical speed of the tractor on the j-gear is determined by the formulas:

$$V_T[m \cdot s^{-1}] = 0.0167 \frac{n_d \cdot t \cdot z}{i_j}, \quad (2)$$

where:

n_d [min^{-1}] is the rotational speed of a diesel engine crankshaft;

t [m] - track pitch;

i_j - tractor gear ratio;

z - number of chain track links tractor number of links moving during one driven wheel turn, z=13.5 cm.

The actual speed of the tractor per hour (when using a master gear, a circle perimeter of which is C = 2.515 m);

$$V_A = 0.027944 \frac{Q_{mg}}{\tau} \text{ [m·s}^{-1}\text{]}, \quad (3)$$

where:

Q_{mg} is the total number of rotation impulses of a master gear during the experiment;

τ - time of the experiment [s].

$$\delta = \left(1 - \frac{m_{mg}}{m_{dw}} \cdot \frac{Q_{mg}}{Q_{dw}}\right) \cdot 100, \quad (4)$$

where:

m_{dw} [m/imp] is impulse scale of sensors measuring the number of driven wheels rotations;

m_{mg} [m/imp] - impulse scale of a sensor, which measures the number of a master gear rotations;

Q_{mg} is the total number of impulses of the rotation sensor of a master gear during the experiment;

Q_{dw} - the total number of impulses of a driven wheel rotation sensor during the experiment.

Based on the calculation results, the graph of dependence of the tractor sliding coefficient on the tractor traction $\delta = f(P_{hk})$ was obtained. All the measurements made on all the gears, were used to create the graph of sliding

The traction power N_{hk} at a given point was calculated by formula:

Table 3 *Main traction indicators of the T-170M1.03-55 tractor at maximum power N_{hkmax} and at maximum traction effort P_{hkmax}*

Gear	Indicators													
	Test mode													
	Nhkmax						Phkmax							
	N_{hk} [kW]	P_{hk} [kN]	V_a [km/h]	δ [%]	G_T [kg/h]	g_{hk} [g/kW/h]	η_{cond}	n_{eng} [min-1]	η_{hk} [kN]	V_a [km/h]	N_{hk} [kW]	δ [%]	φ_{hkmax} [°]	neng [min-1]
flat track-chain rim														
I	84.49	113.47	2.68	2.8	30.06	356	0.684	1096	129.50	1.69	70.37	5.3	0.820	824
II	82.79	92.67	3.22	2.6	30.27	366	0.67	1112	109.68	2.43	73.90	4.0	0.694	848
III	78.09	68.32	4.12	2.6	30.32	388	0.638	1120	84.14	2.83	66.40	4.5	0.533	784
IV	75.96	51.52	5.31	2.5	30.84	406	0.619	1208	67.23	3.57	66.62	2.5	0.426	800
V	70.44	40.25	6.3	1.9	30.74	436	0.560	1232	53.90	3.93	58.82	1.5	0.341	768
VI	70.22	37.19	6.8	2.4	29.65	422	0.546	1120	44.35	5.07	62.43	3.2	0.281	848
VII	59.04	31.03	6.85	3.4	29.33	497	0.475	1016	34.66	5.65	54.34	2.2	0.219	840
elliptical track-chain rim														
I	89.5	117.3	2.75	1.7	30.17	337	0.700	1128	140.1	1.88	73.0	8.6	0.887	824
II	89.2	97.0	3.31	2.8	30.17	338	0.695	1152	116.5	2.29	74.2	5.4	0.738	824
III	85.7	72.1	4.28	0.8	30.02	350	0.673	1136	86.7	2.91	70.1	4.1	0.549	800
IV	82.1	57.6	5.14	0.2	30.07	366	0.640	1144	72.1	3.41	68.2	2.2	0.457	784
V	79.8	51.5	5.58	1.1	29.76	373	0.626	1104	61.8	3.98	68.3	3.7	0.391	792
VI	74.7	39.3	6.84	2.2	30.02	402	0.584	1136	47.7	4.52	59.9	4.5	0.302	768
VII	71.9	35.6	7.28	3.1	29.46	410	0.562	1072	42.0	5.35	62.4	4.3	0.266	800

$$N_{hk} = P_{hk} \cdot V_a, kW . \quad (5)$$

Conditional traction propulsive efficiency coefficient of a tractor by gear: η_{cond} is:

$$\eta_{cond} = N_{hk\max} / N_{op} \cdot C_{PRI} , \quad (6)$$

where:
N_{op} is the engine operating power, according to the State All-Union Standard 18509;
C_{PRI} - the engine power coefficient according to the State All-Union Standard 18509, corresponding to atmospheric conditions and fuel parameters during traction tests on a specific (i-th) gear.

The specific fuel consumption g_{hk}, at a given point, is calculated by formula:

$$g_{hk}[\mathrm{kg/h}] = \frac{G_t}{N_{hk}} , \quad (7)$$

where:
G_t [kg/h] - fuel consumption per hour.

6 Results and discussions

The main indicators of the traction tests of the T-170M1.03-55 tractor, with flat and elliptical track-chain rims, are given in Table 3. A graphical analysis of the data is presented in Figures 8 and 9.

To evaluate influence of the mover support part geometry on traction characteristics of the T-170M1.03-55 tractor, several indicators of traction characteristics were used. Those indicators are: the maximum traction power, N_{hkmax} [kW]; conditional *traction* efficiency coefficient [η_{cond}]; traction effort corresponding to the maximum traction power, P_{hk} [kN], engine speed at maximum traction power, n_e [min^{-1}]; the specific fuel consumption at maximum traction power, g_{hk} [kg.kW^{-1}.h^{-1}]; operating speeds, V_a [km/h]; coefficient of resistance to the movement, *f*; sliding coefficient [δ].

After testing, calculations were made on the comparable indicators of the traction characteristics of the T-170M1.03-55 tractor with a flat and elliptical chain-track rim (Table 4).

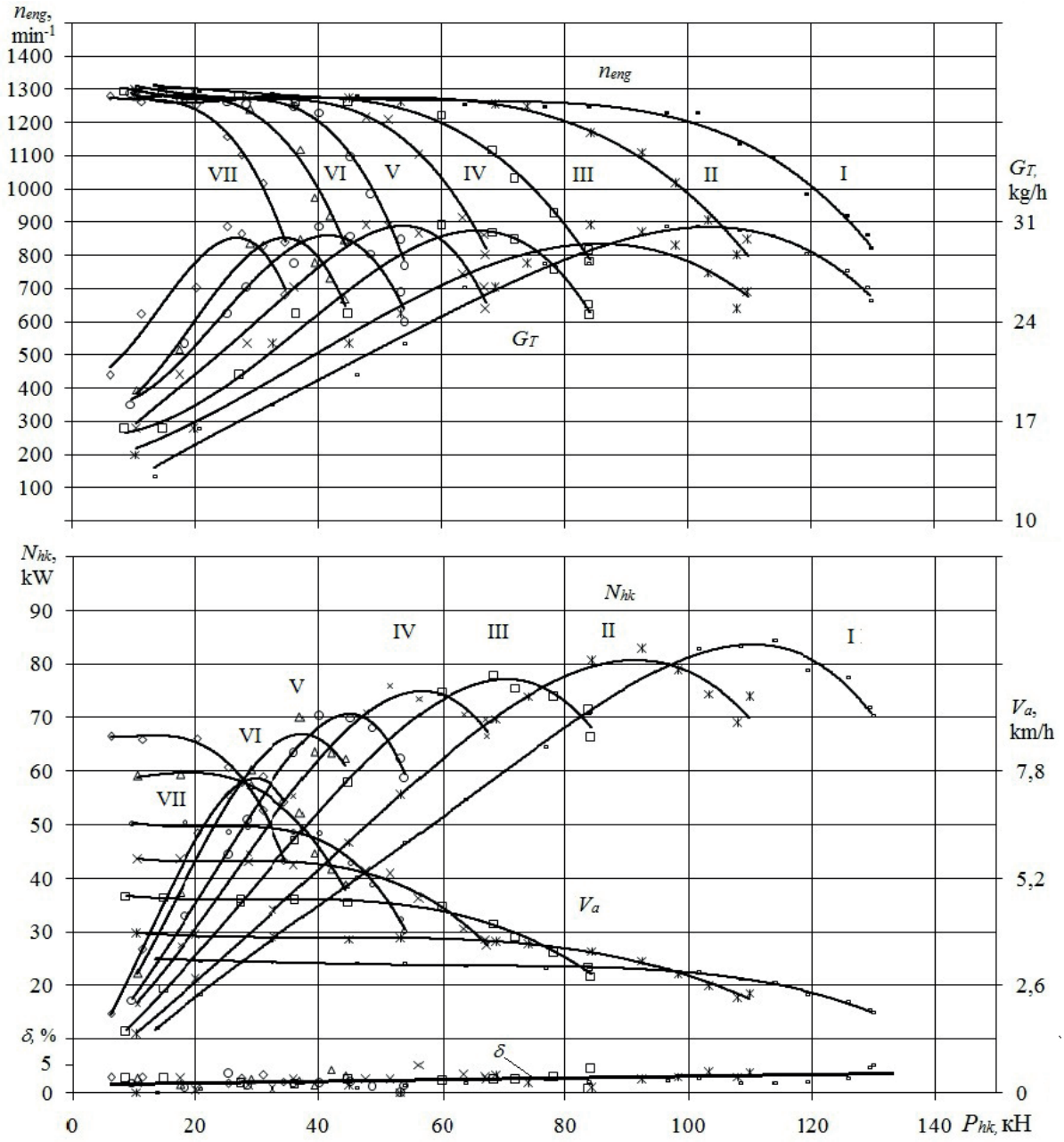


Figure 8 *The traction characteristics of the T-170M1.03-55 tractor with a flat track-chain rim*

The maximum traction power of the T-170M1.03-55 tractor with an elliptical track-chain rim is almost the same when driving in the first and second gears: when driving in the first gear the tractor's maximum traction power is 89.5 kW at 117.3 kN of thrust, speed of 2.75 km/h and the sliding of 1.7 %. In the second gear the tractor's maximum traction power is 89.2 kW at 97.0 kN of thrust, speed of 3.31 km/h and sliding of 2.88 %, which corresponds to a conditional maximum traction thrust coefficient of 0.7 (Figure 10).

Tractor T-170M1.03-55 with a flat track-chain rim does not have alignment of the maximum traction power in different gears. The maximum traction power is 5.9% lower and amounts to 84.49 kW in first gear with a traction effort of 113.5 kN, speed of 2.68 km/h, sliding of 2.8%. The maximum conditional traction efficiency coefficient is 0.684.

With increase in speed, the maximum traction power of the T-170M1.03-55 tractor, with an elliptical track-chain rim, increases from 5.9% in the first to 21.8% in the seventh gear compared to the corresponding indicators with a flat track-chain rim. The maximum traction power in all the gears increased on average by 7.39 kW, 10.4%. Due to increase of the maximum traction power, the potential capabilities of the T-170M1.03-55 tractor also increase.

The rolling resistance of the tractor with a flat track was between 16.48 kN to 22.8 kN with an average value of 19.64 kN. With an elliptical track, the rolling resistance was between 13.79 kN and 14.94 kN, with an average value of 14.36 kN. The rolling resistance of the tractor was determined at a speed of 2.15 to 4.25 km/h. The average difference in rolling resistance power values was 5.27 kN (Figure 11).

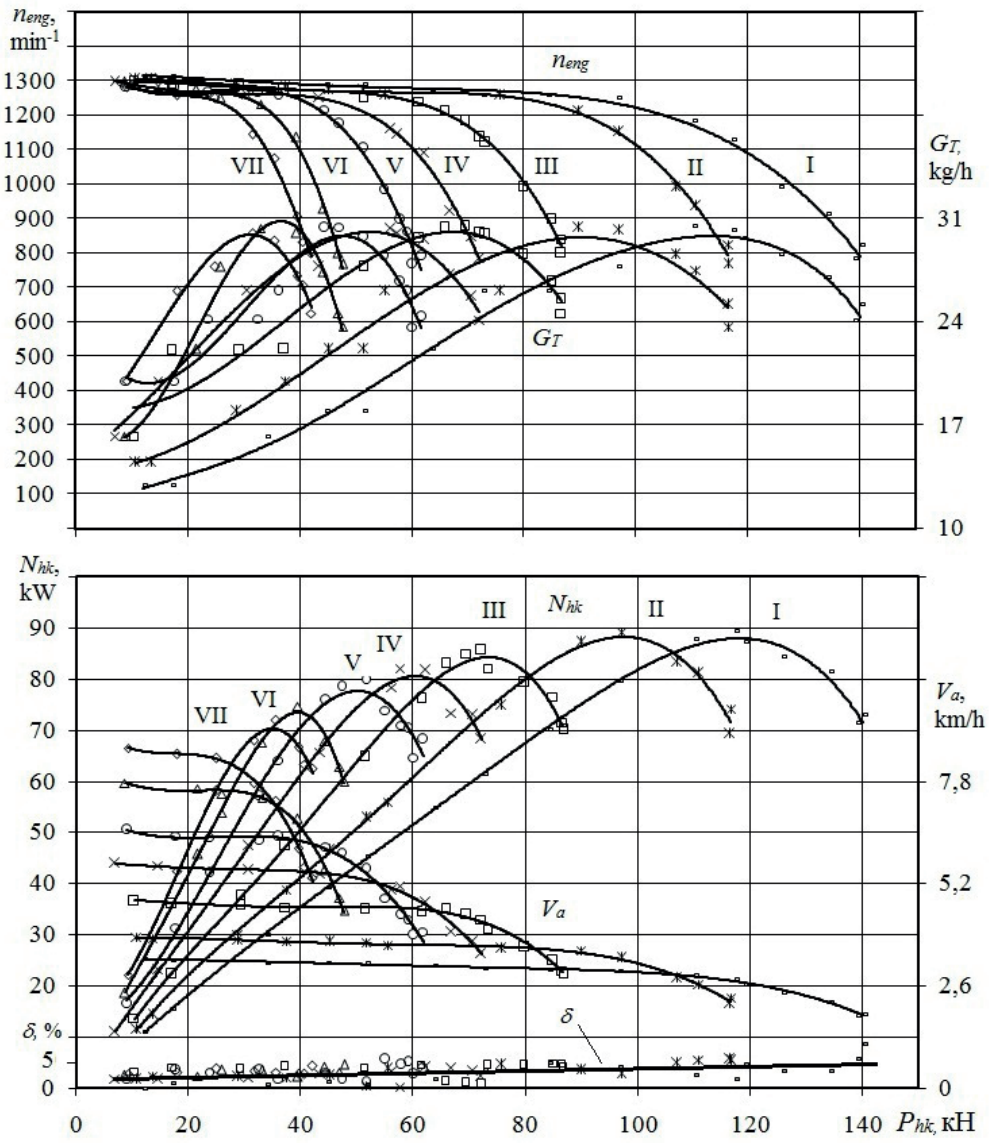


Figure 9 *The traction characteristic of the T-170M1.03-55 tractor with an elliptic track-chain rim*

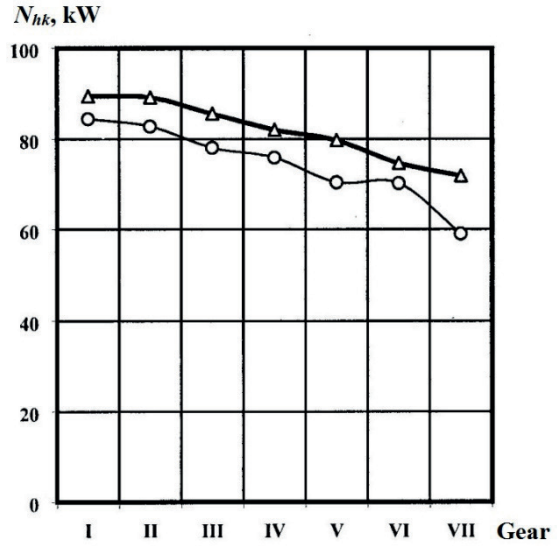


Figure 10 *The maximum traction power of the T-170M1.03-55 by gear: ° - flat track-chain rim; Δ - elliptical track-chain rim*

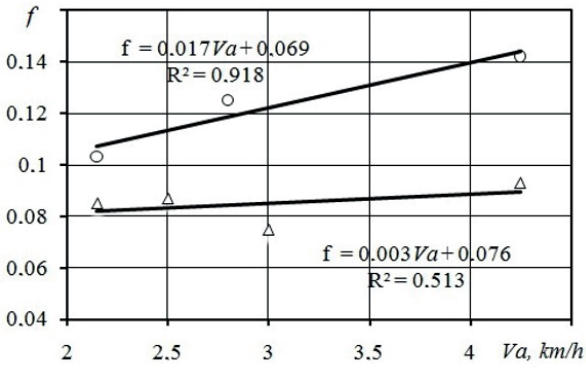


Figure 11 *Dependence of the rolling power coefficient of T-170M1.03-55 tractor on speed: - flat track-chain rim; Δ - elliptical track-chain rim*

Table 4 *Comparable indicators of the T-170M1.03-55 tractor's traction characteristics with different geometry of the mover support part*

Test mode		Mover		Gear ratios							Average values
				1	2	3	4	5	6	7	
N_{hkmax} [kW]	N_{hkmax} [kW]	elliptical		89.5	89.2	85.7	82.1	79.8	74.7	71.9	
		flat		84.49	82.79	78.09	75.96	70.44	70.22	59.04	
		ΔN_{hkmax}	[kW]	5.0	6.4	7.57	6.1	9.34	4.49	12.87	7.39
			[%]	5.9	7.7	9.7	8.0	13.3	6.4	21.8	10.40
	P_{hk} [kN]	elliptical		117.3	97.0	72.1	57.6	51.5	39.3	35.6	
		flat		113.5	92.67	68.32	51.52	40.25	37.19	31.03	
		ΔP_{hk}	[kN]	3.8	4.3	3.8	6.0	11.3	2.2	4.5	5.14
			[%]	3.3	4.7	5.6	11.7	28.0	5.8	14.6	10.54
		Δf		0.024	0.027	0.024	0.037	0.070	0.013	0.028	0.032
	V_a [km/h]	elliptical		2.75	3.31	4.28	5.14	5.58	6.84	7.28	
		flat		2.68	3.22	4.12	5.31	6.3	6.8	6.85	
		ΔVa	[km/h]	0.07	0.09	0.16	-0.17	-0.72	0.04	0.43	-0.01
			[%]	2.6	2.8	3.9	-3.2	-11.4	0.6	6.3	0.22
	δ, %	elliptical		1.7	2.8	0.8	0.2	1.1	2.2	3.1	
		flat		2.8	2.6	2.6	2.5	1.9	2.4	3.4	
		$\Delta\delta$ [%]		-1.1	0.2	-1.8	-2.3	-0.8	-0.2	-0.3	-0.9
	g_{hk} [g/kW·h]	elliptical		337	338	350	366	373	402	410	
		flat		356	366	388	406	436	422	497	
		Δghk	[g/kW·h]	-18.65	-27.34	-37.83	-39.58	-63.36	-20.40	-87.08	-42.04
			[%]	-5.2	-7.5	-9.7	-9.7	-14.5	-4.8	-17.5	-9.87
	η_{cond}	elliptical		0.700	0.695	0.673	0.640	0.626	0.584	0.562	
		flat		0.684	0.670	0.638	0.619	0.560	0.546	0.475	
		$\Delta\eta_{cond}$	-	0.016	0.025	0.035	0.021	0.066	0.038	0.087	0.04
			[%]	2.34	3.73	5.49	3.39	11.79	6.96	18.32	7.43
P_{hkmax} [kN]	P_{hkmax} [kN]	elliptical		140.1	116.5	86.68	72.14	61.75	47.73	42.04	
		flat		129.5	109.7	84.14	67.23	53.90	44.35	34.66	
		ΔP_{hk}	[kN]	10.6	6.86	2.54	4.91	7.86	3.37	7.38	6.22
			[%]	8.19	6.25	3.02	7.31	14.58	7.61	21.29	9.75
		Δf		0.066	0.043	0.016	0.031	0.049	0.021	0.046	0.039
	V_a [km/h]	elliptical		1.88	2.29	2.91	3.41	3.98	4.52	5.35	
		flat		1.69	2.43	2.83	3.57	3.93	5.07	5.65	
		ΔV_a	[km/h]	0.19	-0.14	0.08	-0.16	0.05	-0.55	-0.30	-0.12
			[%]	11.24	-5.76	2.83	-4.48	1.27	-10.8	-5.31	-1.58
	N_{hk} [kW]	elliptical		73.0	74.2	70.1	68.2	68.3	59.9	62.4	
		flat		70.37	73.90	66.40	66.62	58.82	62.43	54.34	
		ΔN_{hkmax}	[kW]	2.65	0.29	3.68	1.62	9.49	-2.57	8.09	3.32
			[%]	3.76	0.40	5.54	2.43	16.13	-4.12	14.88	5.57
	Δ [%]	elliptical		8.6	5.4	4.1	2.2	3.7	4.5	4.3	
		flat		5.3	4	4.5	2.5	1.5	3.2	2.2	
		$\Delta\delta$ [%]		3.3	1.4	-0.4	-0.3	2.2	1.3	2.1	1.37

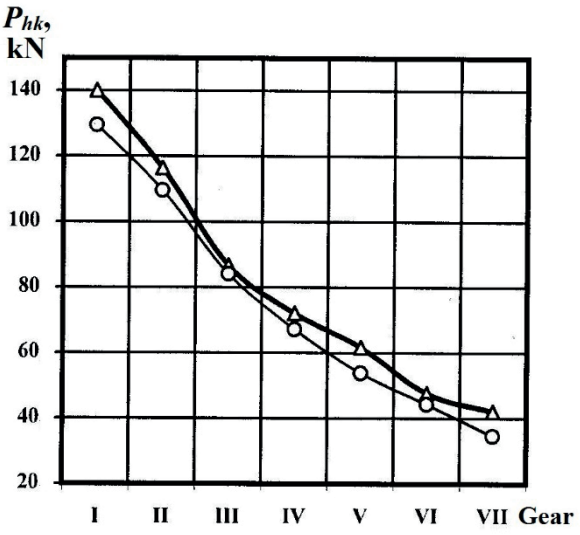


Figure 12 *The maximum traction effort of the T-170M1. 03-55 tractor in gears: - flat track-chain rim; Δ - elliptical track-chain rim*

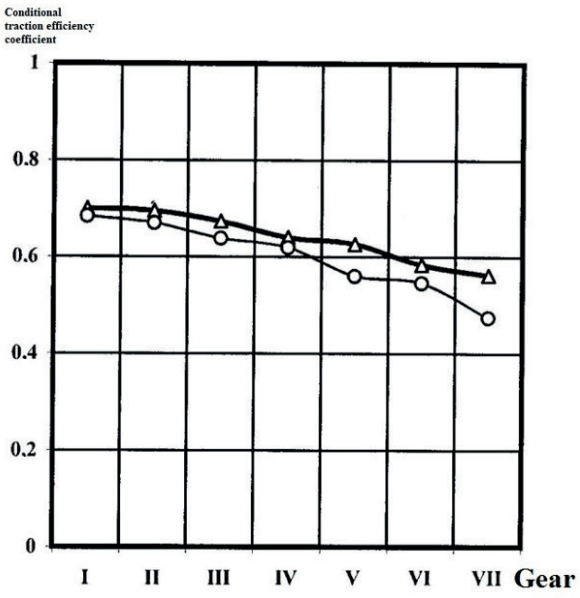


Figure 13 *The conditional traction efficiency coefficient (efficiency factor) of T-170M1.03-55 tractor according to gears: ° - flat track-chain rim; Δ - elliptical track-chain rim*

At the maximum traction power, the tractor develops traction effort: from 117.3 kN to 3.5 kN for a tractor with an elliptical track-chain rim and from 113.5 kN to 3.10 kN for a tractor with a flat track-chain rim (Table 4). A mount of increase in the traction effort ranges from 3.8 kN to 11.3 kN with an average value of 5.14 kN, absolute value of which is almost equal (5.27 kN) to the traction resistance to the rolling of the tractor. The increase in traction effort is observed in all the gears, while there is no certain pattern of their changes.

When analyzing the parameters at the maximum values of the traction effort, a similar pattern is observed. The maximum traction effort of the T-170M1.03-55 tractor is limited by the maximum torque of the engine, and is equal to 140.1 kN when sliding for elliptical track-chain rim is 8.6 %. When sliding for a flat track-chain rim is 5.3 %, the maximum traction effort is 129.5 kN. The maximum traction effort values of a tractor with elliptical rim do not depend on the speed and exceed the maximum traction effort values of a tractor with flat rim on average by 6.22 kN or 9.75% (Figure 12). The increase in the traction power in the maximum traction effort mode was approximately 3.32 kW or 5.57%.

The actual speed of a tractor with an elliptical rim exceeds that of a tractor with a flat rim on average by 3.24 % in all the gears. However, the speed of a tractor with an elliptical rim reduces by 7.3% in fourth and fifth gears. Such a speed difference can be explained by the fact that the maximum power given for the analysis, was obtained at different traction efforts and corresponding engine shaft speeds.

The value of sliding in the maximum traction power modes (in the entire range of the corresponding traction powers) is low for both variants (δ =0.2...3.1% - for an elliptical rim and δ =1.9...3.4% - for a flat rim). Sliding of a tractor with an elliptical rim is on average 0.9 % lower than that of a tractor with a flat track-chain rim (Table 4).

The conditional traction efficiency coefficient (efficiency factor) in gears of a tractor with an elliptical rim (η_{cond} = 0.7…0.562 - for an elliptical rim; $\eta_{cond\ d}$ = 0.684…0.475 - for a flat rim) is higher by 0.016...0.087, or by 2.3...18.3% than that of a flat rim. Same as for the maximum traction power, there is an increase in η_{cond} in all the gears of the tractor with an elliptical rim (Figure 13).

Increase in the traction power and in traction indicators of the T-170M1.03-55 tractor with an elliptical track-chain rim is accompanied by an increase in the tractor fuel efficiency. The specific fuel consumption per unit of the power output of a tractor with an elliptical track-chain rim is reduced from 5.2% in the first gear to 17.5% in the seventh gear (Figure 14). The specific fuel consumption is reduced by an average of 9.87 %.

There is an increase in the maximum traction power and conditional traction efficiency coefficient of a tractor with an elliptical rim in all the gears by an amount independent of gear with almost equal sped values in appropriate gears.

Analysis of results of the traction tests showed that the change in the geometry of the support part of a track-chain mover improves traction of the T-170M1.03.55 tractors and has a positive effect on the engine dynamics. The elliptical geometry of the support part of a track-chain tractor with semi-rigid suspension is realized in a simple way, i.e. installation of track rollers at different heights with the help of plates. Plates thickness is determined using Equation (1).

Reducing the rolling resistance leads to increase in the traction indicators of the T-170M1.03-55 tractor with elliptical rim when under the steady load.

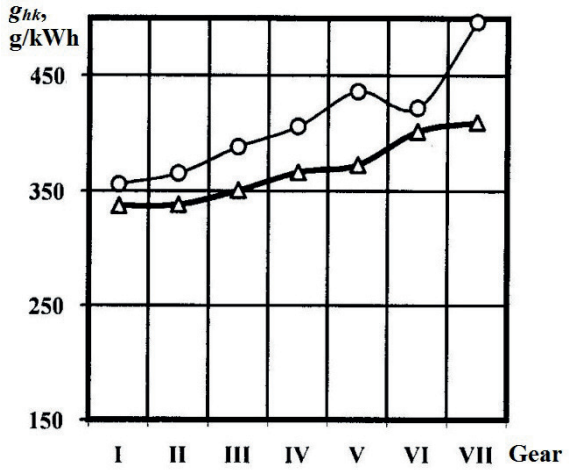


Figure 14 *The specific fuel consumption of the T-170M1.03-55 tractor in the maximum traction mode according to gears: - flat track-chain rim; Δ - elliptical track-chain rim*

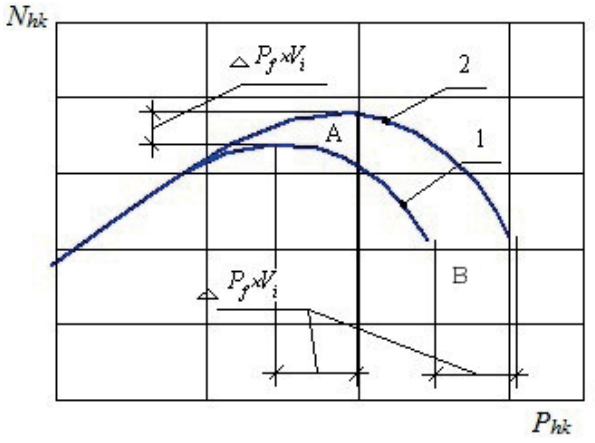


Figure 15 *Diagram of the traction power increase in gears for a tractor with an elliptical track rim: 1 - flat track-chain rim; 2-elliptical track-chain rim*

The tractor power balance equation can be written as:

$$N_{hk} = N_w - N_\delta - N_f = N_w - N_\delta - P_f V_a, \quad (8)$$

where:
N_{hk} [kW] is the traction power;
N_w [kW] - power on the driving wheel;
N_δ [kW] - power loss for sliding ;
N_f [kW] - power loss for movement;
P_f [kN] - movement resistance power;
V_a [km/h] - actual speed of the tractor.

It can be noted that an increase in the maximum traction power of a tractor with an elliptical track-chain rim in gears is determined by a decrease by an approximately constant amount in the movement resistance coefficient:

$$N_{hk\max}{}^{ell} = N_{hk\max}{}^{fl} + \Delta P_f V_i, \quad (9)$$

where:
$N_{hkmax}{}^{ell}$, $N_{hkmax}{}^{fl}$ is the maximum traction power in the i-th gear of the tractor with an elliptical and flat rim respectively;
ΔP_f ~ const = of 5.27 kN is a reduction of a drag power to rolling of the tractor with elliptical rim compared to flat track rim that is equal to increase in traction;
V_i- tractor speed with maximum power in the i-th gear.

Having analyzed the works of various authors, it can be noted that, generally, resistance to tractors' rolling depends on speed, traction power, soil conditions and a mover type [1, 4, 6, 8, 34, 39-45].

According to most studies, the increase in speed results in an increase in the rolling coefficient *f*, which leads to the power loss necessary for rolling. Such an intensive increase in power loss for rolling is explained by an increase in loss for friction and strikes in tracks, as well as vertical compaction of soil. In the traction balance, the tractor power losses per a share of these losses, account for about 80%, of which 40-60% are losses for friction. Generally, the growth of movement resistance powers and accordingly the growth of losses for movement with increasing speed is explained by the fact that there is an increase in power loss for friction in the bearings of track rollers, in idler wheels and track pins power loss for friction when track links slip on track rollers and driven wheels. Other factors have the same effect, like losses caused by the tractor weight when track rollers move and the power of tracks pre-tension when they roll; power loss for friction caused by the increasing irregularity of the pressure distribution along the length of the supporting part of the tracks; the power loss for friction caused by the uneven movement of the track.

Reduction in the rolling resistance of a tractor with an elliptical track rim by 27 % should be considered as a decrease in uneven distribution of normal loads between the track rollers, especially the first and the sixth support roller and a decrease in the maximum pressures under these rollers. The elliptical track rim can be represented as circle segment with a large radius. During the tractor movement the tracks elements move more smoothly relative to each other, which leads to a reduction in rolling resistance of track rollers on the inner track rims and in the track links joints. By reducing the maximum pressure and the number of impacts on soil to one, the loss for soil deformation also decreases.

Based on the data analysis of these studies, the linear dependence of the rolling resistance coefficient on speed can be determined by the following equations:

$$f = 0.017 \cdot V_a \text{ - for a flat track rim;} \quad (10)$$

$$f = 0.003 \cdot V_a + 0.076 \text{ - for an elliptical track rim,} \quad (11)$$

where V_a is tractor's actual speed, km/h.

Most authors have similar positions in Equations (10) and (11). They consider the dependence of the coefficient

of resistance to rolling on the speed to be linear or close to linear [1, 4, 33, 39-40].

The higher operation efficiency of a tractor with an elliptical rim, which isa part of a machine-tractor aggregate, compared to that one with a flat rim, is achieved by the presence of additional A zone of high traction powers and B zone, associated with an increase in the maximum traction effort in each gear (Figure 15). In this case, the speed of the aggregate, with an elliptical rim tractor, increases provided that the traction resistance is within the mentioned zones.

In further studies of a tractor with an elliptical track rim as a part of a plowing unit, the comparative operational field tests should be conducted in the unsteady load mode.

7 Conclusions

Based on the research results, the influence of the geometry of the support part of a track-chain mover on the traction characteristics of the T-170M1.03-55 tractor has been revealed. The results of traction tests showed that elliptical track rim has increased the maximum traction power by 10.4 %, conditional traction efficiency coefficient to 7.43% and specific traction effort by 8%. At the same time, the specific fuel consumption was reduced on average by 9.87%. The increase in indicators is provided by a lower rolling resistance of a tractor with an elliptical rim. Reduction of resistance power to rolling of a tractor with an elliptical track rim by 27% occurs due to alignment of support rollers vertical load and resistance reduction to rollers movement on the internal track rims and in the joints of track chain links.

The increase in the maximum traction powers and conditional traction efficiency coefficient of a tractor with an elliptical rim is observed in all the gears. The maximum values of the conditional traction efficiency coefficient are in the range of large traction efforts, where the power losses during the rolling are significantly reduced.

References

[1] ALSALEH, M., EVANS, W., HOFSTETTER, K. A Review for tracks systems. In: *Advances in soil dynamics*. UPADHYAYA, S. K., CHANCELLOR, W. J., PERUMPRAI, J. V. (eds.). 3. ed. Vol. 3. Chapt. 2. American Society of Agricultural and Biological Engineers, 2009, p. 130-161. ISBN 978-1892769732.

[2] KSENEVICH, I. P., RUSANOV, V. A. The problem of the effect of vehicles on the soil: some results of investigations. *Tractors and Agricultural Machines*. 2000, **1**, p. 15-20. ISSN 0321-4443.

[3] GAINULLIN, I. A. Reducing the compacting effect of a track-chain tractor on soil. *Tractors and Agricultural Machines*. 2001, **9**, p. 19-22. ISSN 0321-4443.

[4] KSENEVICH, I. P., GOBERMAN, V. A., GOBERMAN, L. A. *Ground traction and transport systems.* Moscow, 2003. ISBN 5-217-03161-1.

[5] KARPICHEV, V., SERGEEV, K., BOLOTINA, A. Modeling of technological processes of machine-building and repair manufacture. *Communications-Scientific Letters of the University of Zilina* [online]. 2019, **21**(4), p. 59-62. ISSN 1335-4205, eISSN 2585-7878. Available from: http://komunikacie.uniza.sk/index.php/communications/article/view/1518

[6] PERSSON, S. Basic traction mechanics. In: *Advances in soil dynamics*. UPADHYAYA, S. K., CHANCELLOR, W. J., PERUMPRAI, J. V. (eds.). 3. ed. Vol. 3. Chapt. 2. American Society of Agricultural and Biological Engineers, 2009, p. 25-28. ISBN 978-1892769732.

[7] GRISSO, R., PERUMPRAL, J., ZOZ, F. An empirical model for tractive performance of rubber-tracks in agricultural soils. *Journal of Terramechanics* [online]. 2006, **43**(2), p. 225-236. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2005.12.002

[8] ARVIDSSON, J., WESTLIN, H., KELLER, T., GILBERTSSON, M. Rubber track systems for conventional tractors - effects on soil compaction and traction. *Soil and Tillage Research* [online]. 2011, **117**, p. 103-109. ISSN 0167-1987. Available from: https://doi.org/10.1016/j.still.2011.09.004

[9] MOLARI, G., BELLENTANI, L., GUARNIERI, A., WALKER, M., SEDONI, E. Performance of an agricultural tractor fitted with rubber tracks. *Biosystems Engineering* [online]. 2012, **111**(1), p. 57-63. ISSN 1537-5110. Available from: https://doi.org/10.1016/j.biosystemseng.2011.10.008

[10] KELLER, T., ARVIDSSON, J. A model for prediction of vertical stress distribution near the soil surface below rubber-tracked undercarriage systems fitted on agricultural vehicles. *Soil and Tillage Research* [online]. 2016, **155**, p. 116-123. ISSN 0167-1987. Available from: https://doi.org/10.1016/j.still.2015.07.014

[11] RASOOL, S., RAHEMAN, H. Improving the tractive performance of walking tractors using rubber tracks. *Biosystems Engineering* [online]. 2018, **167**, p. 51-62. ISSN 1537-5110. Available from: https://doi.org/10.1016/j.biosystemseng.2017.12.013

[12] TAHERI, S., SANDU, C., TAHERI, S., PINTO, E., GORSICH, D. A technical survey on terramechanics models for tire-terrain interaction used in modeling and simulation of wheeled vehicles. *Journal of Terramechanics* [online]. 2015, **57**, p. 1-22. Available from: https://doi.org/10.1016/j.jterra.2014.08.003

[13] JANULEVICIUS, A., DAMANAUSKAS, V., PUPINIS, G. Effect of variations in front wheels driving lead on performance of a farm tractor with mechanical front-wheel-drive. *Journal of Terramechanics* [online]. 2018, **77**, p. 23-30. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2018.02.002

[14] EL-SAYEGH, Z., EL-GINDY, M., JOHANSSON, I., OIJER, F. Improved tire-soil interaction model using FEA-SPH simulation. *Journal of Terramechanics* [online]. 2018, **78**, p. 53-62. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2018.05.001

[15] FARHADI, P., GOLMOHAMMADI, A., SHARIFI, A., SHAHGHOLI, G. Potential of three-dimensional footprint mold in investigating the effect of tractor tire contact volume changes on rolling resistance. *Journal of Terramechanics* [online]. 2018, **78**, p. 63-72. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2018.05.003

[16] PADMANABHAN, C., GUPTA, S., MYLSWAMY, A. Estimation of terramechanics parameters of wheel-soil interaction model using particle filtering. *Journal of Terramechanics* [online]. 2018, **79**, p. 79-95. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2018.07.003

[17] YANG, C., CAI, L., LIU, Z., TIAN, Y., ZHANG, C. A calculation method of track shoe thrust on soft ground for splayed grouser. *Journal of Terramechanics* [online]. 2016, **65**, p. 38-48. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2016.02.001

[18] EDWIN, P., SHANKAR, K., KANNAN, K. Soft soil track interaction modeling in single rigid body tracked vehicle models. *Journal of Terramechanics* [online]. 2018, **77**, p. 1-14. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2018.01.001

[19] WANG, M., WANG, X., SUN, Y., GU, Z. Tractive performance evaluation of seafloor tracked trencher based on laboratory mechanical measurements. *International Journal of Naval Architecture and Ocean Engineering* [online]. 2016, **8**(2), p. 177-187. ISSN 2092-6782, eISSN 2092-6790. Available from: https://doi.org/10.1016/j.ijnaoe.2016.01.005

[20] HAMZA, M. A., ANDERSON, W. K. Soil compaction in cropping systems: A review of the nature, causes and possible solutions. *Soil and Tillage Research* [online]. 2005, **82**(2), p. 121-145. ISSN 0167-1987. Available from: https://doi.org/10.1016/j.still.2004.08.009

[21] HOLTKEMEYER, V. Measuring tire deformation from various wheel loads and inflation pressures. *Landtechnik* [online]. 2005, **60**(2), p. 76-77. ISSN 0023-8082. Available from: https://doi.org/10.15150/lt.2005.1167

[22] ELAOUD, A., CHEHAIBI, S. Soil compaction due to tractor traffic. *Journal of Failure Analysis and Prevention* [online]. 2011, **11**(5), p. 539-545. ISSN 1547-7029. Available from: https://doi.org/10.1007/s11668-011-9479-3

[23] NAWAZ, M. F., BOURRIE, G., TROLARD, F. Soil compaction impact and modelling. A review. *Agronomy for Sustainable Development* [online]. 2013, **33**(2), p. 291-309. ISSN 1774-0746, eISSN 1773-0155. Available from: https://doi.org/10.1007/s13593-011-0071-8

[24] CUETO, O. G., CORONEL, C. E. I., BRAVO, E. L., MORFA, C. A. R., SUAREZ, M. H. Modelling in FEM the soil pressures distribution caused by a tyre on a Rhodic Ferralsol soil. *Journal of Terramechanics* [online]. 2016, **63**, p. 61-67. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2015.09.003

[25] GAINULLIN, I. A., ZAINULLIN, A. R. Improvement of energy and environmental parameters of the track-chain mover of T-170M1.03-55 tractor (in Russian). *Achievements of Science and Technology in Agriculture* [online]. 2017, **31**(2), p. 69-72. ISSN 0235-2451. Available from: https://cyberleninka.ru/article/n/uluchshenie-energeticheskih-i-ekologicheskih-pokazateley-gusenichnogo-dvizhitelya-traktora-t-170m1-03-55/viewer

[26] GABITOV, I. I., MUDARISOV, S. G., GAFUROV, I. D., ABLEEVA, A. M., NEGOVORA, A. V., DAVLETSHIN, M. M., RAKHIMOV, Z., KHAMALETDINOV, R., MARTYNOV, V., YUKHIN, G. P. Evaluation of the efficiency of mechanized technological processes of agricultural production. *Journal of Engineering and Applied Sciences* [online]. 2018, **13**(S10), p. 8338-8345. ISSN 1816-949x, eISSN 1818-7803. Available from: https://doi.org/10.3923/jeasci.2018.8338.8345

[27] PAULSON, I. W., DOLOVICH, A. T., NOBLE, S. D. Development of a dynamic simulation model of a towed seeding implement. *Journal of Terramechanics* [online]. 2018, **75**, p. 25-35. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2017.10.006

[28] RAKHIMOV, Z. S., MUDARISOV, S. G., GABITOV, I. I., RAKHIMOV, I., RAKHIMOV, R., FARKHUTDINOV, I. M., TANYLBAEV, M., VALIULLIN, I., YAMALETDINOV, M., AMINOV, R. Mathematical description of the mechanical erosion process in sloping fields. *Journal of Engineering and Applied Sciences* [online]. 2018, **13**(S8), p. 6505-6511. ISSN 1816-949x, eISSN 1818-7803. Available from: https://doi.org/10.3923/jeasci.2018.6505.6511

[29] PULIDO-MONCADA, M., MUNKHOLM, L. J., SCHJONNING, P. Wheel load, repeated wheeling, and traction effects on subsoil compaction in northern Europe. *Soil and Tillage Research* [online]. 2019, **186**, p. 300-309. ISSN 0167-1987. Available from: https://doi.org/10.1016/j.still.2018.11.005

[30] GAINULLIN, I. A. Influence of operating conditions on the intensity of tracks wear of T-170M1. 03-53 tractor. *Bulletin of the Bashkir State Agrarian University.* 2019, **1**(49), p. 121-126. ISSN 10-8900B. Available from: 10.31563/1684-7628-2019-49-1-121-126.

[31] GAINULLIN, I. A. Substantiation of the geometry of the support surface of a track-chain mover and gravity center of a tractor with semi-rigid suspension. *Bulletin of the Chelyabinsk State Agricultural Engineering University.* 2001, **34**, p. 42-47. ISSN 629.114.2.032.1

[32] ZAKHMATOV, I. P., SHATALOV, V. T., RAZUVAEV, F. S. *Geometry of the supporting part of a track-chain tractor with a semi-rigid suspension*. Chelyabinsk, 1982, p. 25-32.

[33] KYCHEV, V. N., KARLOV, A. G. *The relationship of "mover-soil-tool" functioning and the efficiency of the chassis of a track-chain tractor*. Scientific works of Chelyabinsk Institute of Mechanization and Electrification of Agriculture. Chelyabinsk, 1985, p. 34-35.

[34] WULFSOHN, D., WAY, T. R. Factors that influence tractive performance of weels, tracks, and vehicles. In: *Advances in soil dynamics*. UPADHYAYA, S. K., CHANCELLOR, W. J., PERUMPRAI, J. V. (eds.). 3. ed. Vol. 3. Chapt. 2. American Society of Agricultural and Biological Engineers, 2009, p. 209-252. ISBN 978-1892769732.

[35] WAY, T. R. Single wheel testers, single-track testers, and instrumented tractors. In: *Advances in soil dynamics* [online]. UPADHYAYA, S. K., CHANCELLOR, W. J., PERUMPRAI, J. V. (eds.). 3. ed. Vol. 3. Chapt. 2. American Society of Agricultural and Biological Engineers, 2009, p. 253-271. ISBN 978-1892769732. Available from: https://www.ars.usda.gov/research/publications/publication/?seqNo115=232549

[36] ELAOUD, A., CHEHAIBI, S., ABROUGUI, K. Simulation of soil behavior following the passage of tractors. *International Journal of Current Engineering and Technology*. 2015, **5**(1), p. 534-538. ISSN 2347-5161, eISSN 2277-4106.

[37] LEE, J. W., KIM, J. S., KIM, K. U. Computer simulations to maximise fuel efficiency and work performance of agricultural tractors in rotovating and ploughing operations. *Biosystems Engineering* [online]. 2016, **142**, p. 1-11. ISSN 1537-5110. Available from: https://doi.org/10.1016/j.biosystemseng.2015.11.012

[38] GAINULLIN, I. A., ZAINULLIN, A. R. Influence of design parameters of movers and loading modes of tractors on soil. *Fundamental Research* [online]. 2017, **2**, p. 31-36. ISSN 1812-7329. Available from: https://doi.org/10.17513/fr.41352

[39] GUSKOV, V. The effect of drawbar pulls on the rolling resistance of track-laying tractors. *Journal of Terramechanics*. 1968, **5**(4), p. 27-32. ISSN 0022-4898.

[40] BOGATYREV, A. P., BERDOV, E. I., GAINULLIN, I. A. Determination of the coefficient of resistance to the movement of track-chain tractors. *Bulletin of the Cheliabinsk Agricultural Engineering University*. 2000, **32**, p. 33-39. ISSN 631.372.004.15:530.152.1.(045).

[41] KOSTIUCHENKO, V. I. Direct evaluation of the specific traction of a track-chain tractor which is optimal according to traction efficiency coefficient. *Bulletin of the South Ural State University. Series: Engineering*. 2005, **14**(54), p. 90-92. ISSN 629.114.2.

[42] KURJENLUOMA, J., ALAKUKKU, L., AHOKAS, J. Rolling resistance and rut formation by implement tires on tilled clay soil. *Journal of Terramechanics* [online]. 2009, **46**(6), p. 267-275. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2009.07.002

[43] BOTTA, G. F., TOLON-BECERRA, A., TOURN, M., LASTRA-BRAVO, X., RIVERO, D. Agricultural traffic: Motion resistance and soil compaction in relation to tractor design and different soil conditions. *Soil and Tillage Research* [online]. 2012, **120**, p. 92-98. ISSN 0167-1987. Available from: https://doi.org/10.1016/j.still.2011.11.008

[44] TAGHAVIFAR, H., MARDANI, A. Investigating the effect of velocity, inflation pressure, and vertical load on rolling resistance of a radial ply tire. *Journal of Terramechanics* [online]. 2013, **50**(2), p. 99-106. ISSN 0022-4898. Available from: https://doi.org/10.1016/j.jterra.2013.01.005

[45] BATTIATO, A., DISERENS, E. Tractor traction performance simulation on differently textured soils and validation: A basic study to make traction and energy requirements accessible to the practice. *Soil and Tillage Research* [online]. 2017, **166**, p. 18-32. ISSN 0167-1987. Available from: https://doi.org/10.1016/j.still.2016.09.005

Serhiy Buriakovskyi - Borys Liubarskyi - Artem Maslii - Danylo Pomazan - Tatyana Tavrina

RESEARCH OF A HYBRID DIESEL LOCOMOTIVE POWER PLANT BASED ON A FREE-PISTON ENGINE

This article describes one of the possible ways for improving the energy efficiency of shunting diesel locomotives. It means a replacing a traditional traction electric transmission with a diesel generator set with a hybrid transmission with a free-piston internal combustion engine and a linear generator. The absence of a crankshaft in an internal combustion engine makes it possible to reduce thermal and mechanical losses, which, in turn, leads to an increase in the efficiency of traction electric transmission of the diesel locomotive.

Keywords: *internal combustion engine, power plant, free piston engine, linear generator, crankshaft*

1 Introduction

On modern diesel locomotives, a generator-motor system that converts the thermal energy of fuel combustion into the kinetic energy of the piston motion is employed. The progressive motion of the piston is converted into rotational motion of the crankshaft and the generator shaft, which in the end results in the generation of an EMF generator. As a result of repeated energy conversion in the internal combustion engine (ICE), significant energy losses occur.

As mentioned in [1], in the ICE with a crankshaft, the combustion process occurs near its top dead centre (TDC); therefore the highest temperature and pressure are maintained for a relatively long time. It leads to high heat loss, which degrades the efficiency of the engine. Because of the complex design, the problem for internal combustion engines with a crankshaft is friction losses. The loss of power for friction is the major part of all mechanical losses. Mainly these losses fall on the following pairs: the piston and piston rings - cylinder walls, crankshaft and camshaft journals - plain bearings, the piston pin - piston bosses and the upper connecting rod head, the valve stem - the bushing. Losses on friction increase with increasing load on the engine, increasing crankshaft speed, rough processing of the surface of mating parts, unjustified increase in their size, the use of poor-quality oils, disruptions of the lubrication system and the cooling system and deterioration of the technical state of the engine. Losses in the mechanism of the crankshaft account for 16-19% of all mechanical losses, and the frictional losses between the piston and the cylinder liner that result from the formation of lateral forces of the piston are 42-50% [2].

Overcoming these drawbacks is possible with the development of new types of the ICE, namely engines with a free piston. In a free-piston engine, the movement of the piston is not limited to the presence of a rigid connection with the crankshaft.

The free piston engine consists of two main components: a free piston ICE and linear generator.

As opposed to common ICEs with crankshaft, pistons of the free piston engine freely move in the cylinder, which allows you to change the compression ratio and optimize the combustion process. Changing the compression ratio, this type of engine can work with a homogeneous charge ignition, which allows increasing its thermal efficiency and reducing the emissions of harmful substances into the atmosphere.

The free piston engine is simpler in design, which reduces production costs, compared to the ICE with the crankshaft. The absence of a crankshaft reduces friction losses, and the movement of a free piston can be carried out with greater acceleration. According to [3], the peak piston acceleration in a free-piston engine is about 60% higher than in a conventional engine and the free-piston engine runs less in the TDC zone, where the gas pressure and temperature are the highest. Consequently, the loss of heat in the cylinder of the free piston engine is less than that of a conventional engine.

In the context of the need to reduce the consumption of fuel and energy resources for the locomotive traction, the development and investigation of new types of power plants for rolling stock is quite urgent. As it was said in [4], at present, hybrid traction electric transmissions are increasingly used. The use of such transmission in a diesel locomotive implies the presence of an on-board energy

Serhiy Buriakovskyi[1], Borys Liubarskyi[2], Artem Maslii[3], Danylo Pomazan[3,*], Tatyana Tavrina[4]
[1]Scientific Research and Design Institute Molniya, National Technical University Kharkiv Polytechnic Institute, Ukraine
[2]Department of Electrical Transport and Diesel Locomotive, National Technical University Kharkiv Polytechnic Institute, Ukraine
[3]Department of Electroenergy, Electrical Equipment and Electromecanics, Ukrainian State University of Railway Transport, Kharkiv, Ukraine
[4]Department of Physics, National Technical University Kharkiv Polytechnic Institute, Kharkiv, Ukraine
*E-mail of corresponding author: danil.pomazan@ukr.net

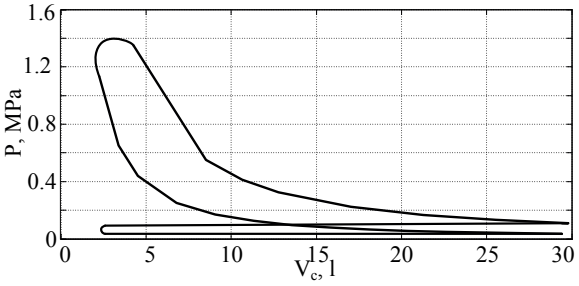


Figure 1 *An indicator diagram of K6S310DR diesel engine operation*

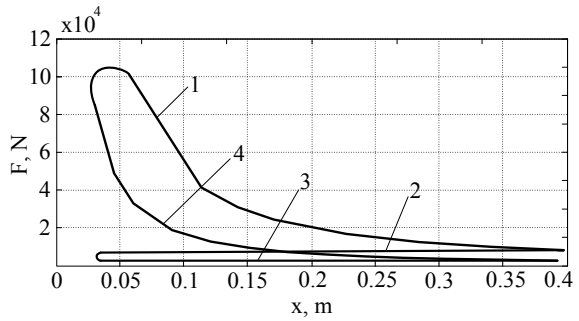


Figure 2 *A force-displacement diagram of the K6S310DR diesel engine*

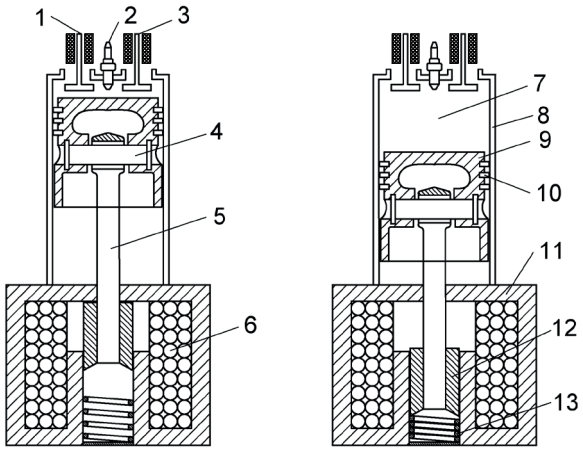


Figure 3 *Free piston engine structure: 1 - inlet electromagnetic valves, 2 - heating plug, 3 - exhaust electromagnetic valves, 4 - piston pin, 5 - piston rod, 6 - coil, 7 - combustion chamber, 8 - cylinder, 9 - piston, 10 - piston rings, 11 - stator, 12 - armature, 13 - spring*

Table 1 *K6S310DR diesel engine specification*

Parameter	Value
Cylinder diameter, mm	310
Piston stroke, mm	360
Power, kW	993
Shaft speed, rpm	750
Boost pressure, kPa	98.1
Total cylinder volume, l	177
Compression ratio	13

storage device in the form of an accumulator, capacitor, or other types of batteries [5]. This makes it possible to reduce the capacity of the power plant and use the energy which is stored in the batteries [6]. Despite the higher efficiency of such a transmission compared to traditional, the issue of low efficiency of the power plant remains unresolved. Based on this, one of the ways to increase the efficiency of such systems is using an engine with a free piston. Because such a power unit consists of independent generator modules, it is characterized by a jog operation. This drawback is offset by the shift of the work cycles of the modules relative to each other and their work on common energy storage. Based on the foregoing, the integration of a free-piston ICE with a linear generator into a hybrid traction electric transmission will not cause additional complication of the system but will make it possible to increase the efficiency of the locomotive's power plant.

The purpose of this work is to study the working properties of a free-piston engine of a shunting diesel locomotive. To solve this problem, the design of the engine and generator was selected, and simulation models of their work, that allows to study the features of this system, were compiled.

2 The choice of design of the internal combustion engine and generator

The development of the design of a free-piston engine includes determining the types of an ICE and a linear generator. To unify the newly developed power plant with an ICE that currently exists on the railways, the K6S310DR diesel engine of the ChME3 shunting diesel locomotive, the technical characteristics of which are shown in Table 1, was chosen as a prototype.

The design of the new engine uses a pair of piston-cylinder, as well as the intake and exhaust systems of the existing diesel.

The choice of a linear generator is justified by the condition that it creates forces equal to those produced by the ICE's piston. Based on this, the calculation of the diesel engine working processes was carried out according to the technique given in [7]. By the calculation results,

the indicator diagram of the diesel engine operation was plotted, which is shown in Figure 1.

For the subsequent work, the use of a diagram in the P and V axes is highly inconvenient, so the diagram shown in Figure 1 is transformed into a diagram in the axes F and x, where F is the force acting on the piston, x is the displacement of the piston. The diagram of the dependence of the force acting on the piston against displacement is shown in Figure 2.

In order to reach the maximum coefficient of efficiency, the mechanical characteristic of the generator should be as close as possible to the work diagram of the diesel engine. In the operation of the diesel engine, four main modes can be distinguished: 1) working stroke, 2) exhaust, 3) cylinder filling, 4) compression.

As a linear generator, an electromagnet was chosen as its characteristic has a shape similar to the shape of the compression and expansion curves of a diesel engine. The use of this design of an electromechanical converter (EMC) does not allow for a four-stroke mode of operation of the ICE. The use of electromagnetic inlet and outlet valves makes it possible to put the ICE into the two-stroke mode that will be considered further.

The result of the analysis done is the choice of the design of a free piston engine for a hybrid diesel locomotive, which is shown in Figure 3.

The linear generator (Figure 3) in the form of electromagnet contains coil 6, armature 12 and stator 11. The generator has two basic positions: coincided one and unmatched one, the connection of the armature with the rod is designed so that in the position of the piston at the top dead center, the armature was in the coincided position. The use of a two-stroke internal combustion engine involves two strokes of the piston, in which the cylinder is filled and the fuel-air mixture is compressed, the piston is stroked and gas is exhausted. The main is the stroke of the piston. During the stroke of the piston, the generator generates electrical energy, which is accumulated in the storage system. As a storage system, a storage battery or a supercapacitor battery and their combination, can be used. During compression of the fuel-air mixture, the generator operates in engine mode and moves the piston. The design provides for the use of a spring 13, it serves to increase the force developed by the electromagnet at the beginning of the compression stroke of the fuel-air mixture.

3 Development of a simulation model of the engine work

The work of an EMC of electromagnetic type is described by the following system of equations:

$$\begin{cases} \dfrac{di}{dt} = \dfrac{1}{\dfrac{\partial \Psi(i,x)}{\partial i}} \cdot \left[E - ri - \dfrac{\partial \Psi(i,x)}{\partial x} \cdot v \right]; \\ \dfrac{dv}{dt} = \dfrac{F_{el}(i,x) + F_s - F_r}{m}; \\ \dfrac{dx}{dt} = v. \end{cases} \quad (1)$$

where:

i - stator current,

Ψ - the armature flux linkage,

E - the power supply voltage,

r - the coil resistance,

x - the translation of generator armature,

v - the speed,

F_{el} - the electromagnetic force,

F_s - the spring pressure,

F_r - the resistance force.

In the linear EMF of electromagnetic type, the developed force and flux linkage of the winding depend on armature floating; therefore, these parameters were calculated with the aid of the FEMM software package [8]. In this package, the calculation is done by the finite element method, and the calculated zone is divided into elementary triangular zones (Figure 4.1). The distribution of the magnetic field in the stator and the armature of the electromagnet is shown in Figure 4.2.

Using a subroutine written in a text file in the Lua programming language, the calculation process was computerised. Whereupon the boundary conditions are the movement of the armature and the stator from the coincided position to the unmatched one that comprises 315 mm, and the permissible value of MMS at the level of 40000 A.

The results of the calculation are the obtained values of the flux linkage and the force put into action by the magnet at different values of displacement and MMS of the armature winding, which are presented in Figures 5 and 6.

To obtain continuous dependencies of flux linkage, the results of digital modelling were approximated by continuous functions. When choosing the type of function, the following tasks were solved:

- the function and its derivatives for moving the armature at the beginning and end of the interval on which the approximation is carried out should be the same [9];

- the proposed function must have a fairly simple form of analytic partial derivatives with respect to all coordinates [10].

Taking into consideration the said above, it is proposed the function of the following form:

$$f(i,x) = \sum_{l=1}^{m} \left[\left(\sum_{h=0}^{n} aa_h \cdot i^h \right) \times \cos(lx) + \left(\sum_{h=0}^{n} bb_h \cdot i^h \right) \sin(lx) \right] + \sum_{h=0}^{n} cc_h \cdot i^h, \quad (2)$$

where:

aa_h, bb_h, cc_h, - polynomial coefficients,

i - the stator current,

x - the armature translation,

l - the harmonic number.

To determine the coefficients of the polynomial, we used the method based on Chebyshev polynomials on the set of equidistant points [11].

Reduction of the order of the mathematical model without limitation can be obtained if the force is determined from the results of the calculation of the magnetic field by

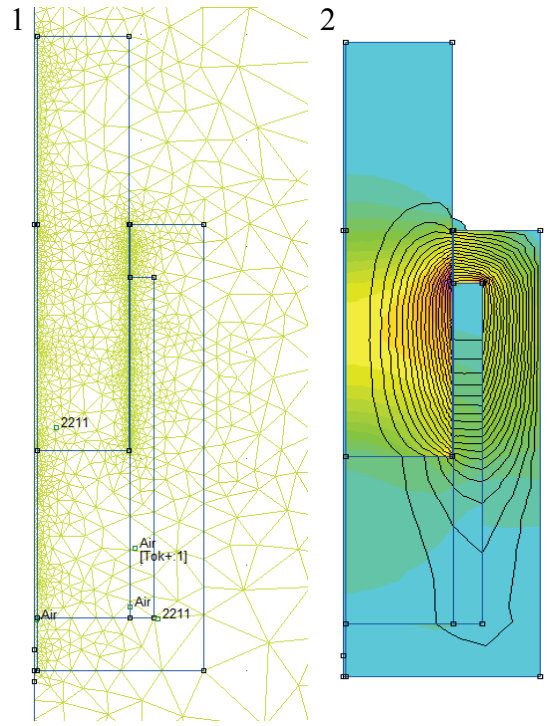


Figure 4 *Estimated area (1) and magnetic flux density distribution in the generator (2)*

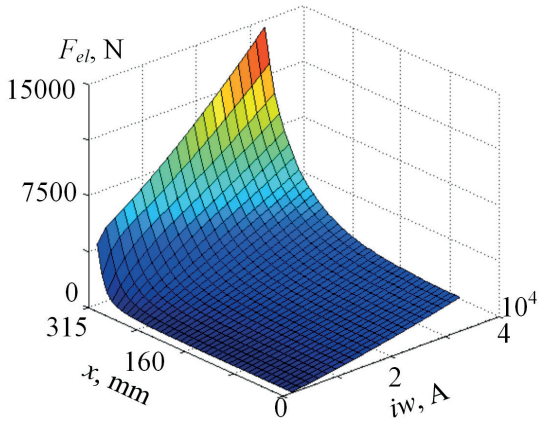


Figure 5 *Values of flux linkage obtained as a result of calculation*

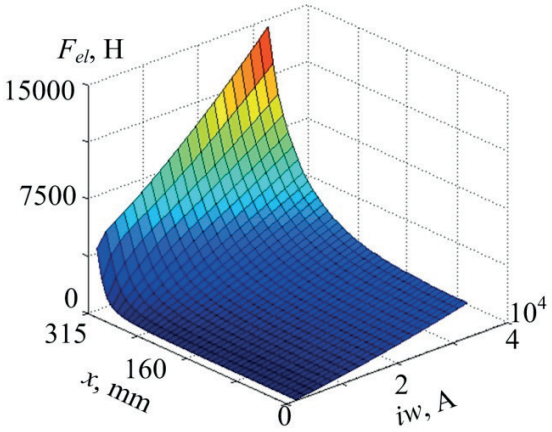


Figure 6 *Values of the force obtained as a result of calculation*

the finite element method (using the FEMM program [8]), and then approximated in the polynomial of the following form:

$$F_e(i,x) = \sum_{j=0}^{JF}\sum_{k=0}^{KF}\left(ma_{jk}(MJF\cdot i + ZJF)^j \times \atop \times (MKF\cdot x + ZKF)^k\right), \quad (3)$$

where:

ma_{jk} - polynomial coefficients that are determined by the Chebyshev method,

MJF, *MKF* - scale current and clearance coefficients, respectively,

ZJF, *ZKF* - the bias of current and clearance, respectively.

The polynomial describing the function of stator winding flux linkage has the following form:

$$\Psi(i,x) = w\sum_{j=0}^{J}\sum_{k=0}^{K}\left(aa_{jk}(MJ\cdot i + ZJ)^j \times \atop \times (MK\cdot x + ZK)^k\right), \quad (4)$$

where:

aa_{jk} - polynomial coefficients for winding that are determined by the Chebyshev method,

w - number of winding turns,

MJ, *MK* - scale current and clearance coefficients, respectively,

ZJ, *ZK* - the bias of current and clearance, respectively,

J, *K* - the degrees of the approximating polynomial of current and clearance, respectively.

To create a simulation model, functional diagrams of the free piston engine, which are shown in Figures 5 and 6, were drawn up.

On the basis of the functional diagrams (Figures 7 and 8) and the system of Equations (1), a structural diagram of the model of the operation of a free piston engine with linear EMC was drawn up (Figure 9).

On the basis of the structural scheme, an imitation model of the engine operation was developed. The simulation was performed in the MatLab environment [12-13].

The ICE model is based on the diagram of its operation (Figure 2) approximated by polynomials. The linear EMC model is constructed on the basis of the system of Equations (1). The electronic commutator includes 2 transistors and 2 reverse diode, the models of which were taken from the SimPowerSystems section [14-16]. In the motor mode, the winding is powered through transistors, and in the generator mode current flows through the diodes. A capacitor, that is charged by EMC in generation mode and

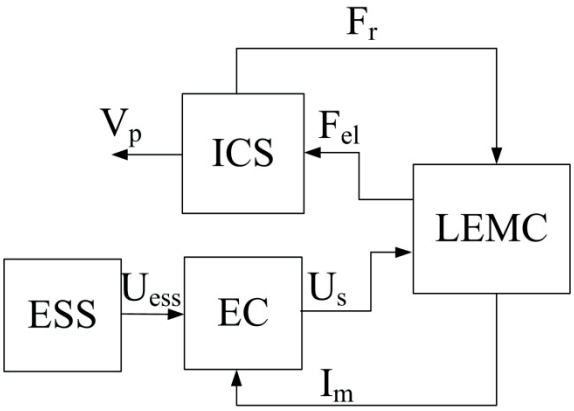


Figure 7 *Flow diagram of free piston engine operation in compression mode: ESS - energy storage system, ICS - internal combustion system, EC - electronic commutator, LEMC - linear electromechanical converter, U_{ess} - supply voltage of EC from ESS, U_s - armature supply voltage, I_m - engine current, F_{el} - electromagnetic force, F_r - ICE resistance force, V_p - piston speed*

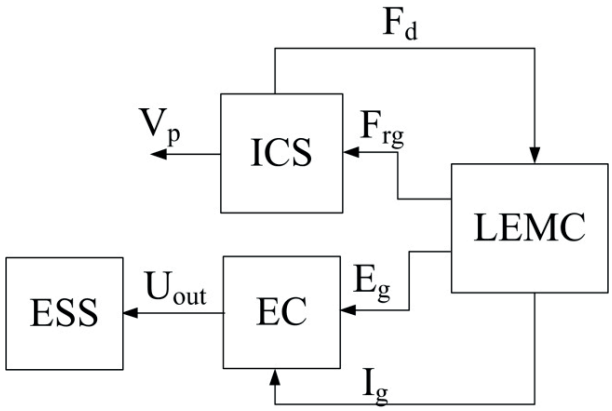


Figure 8 *Flow diagram of free piston engine operation in operating motion mode: U_{out} - output voltage of EC for charging ESS, E_g - generator voltage, I_g - generator current, F_{rg} - generator resistance force, F_d - force generated by the working gases of a diesel engine, V_p - piston speed*

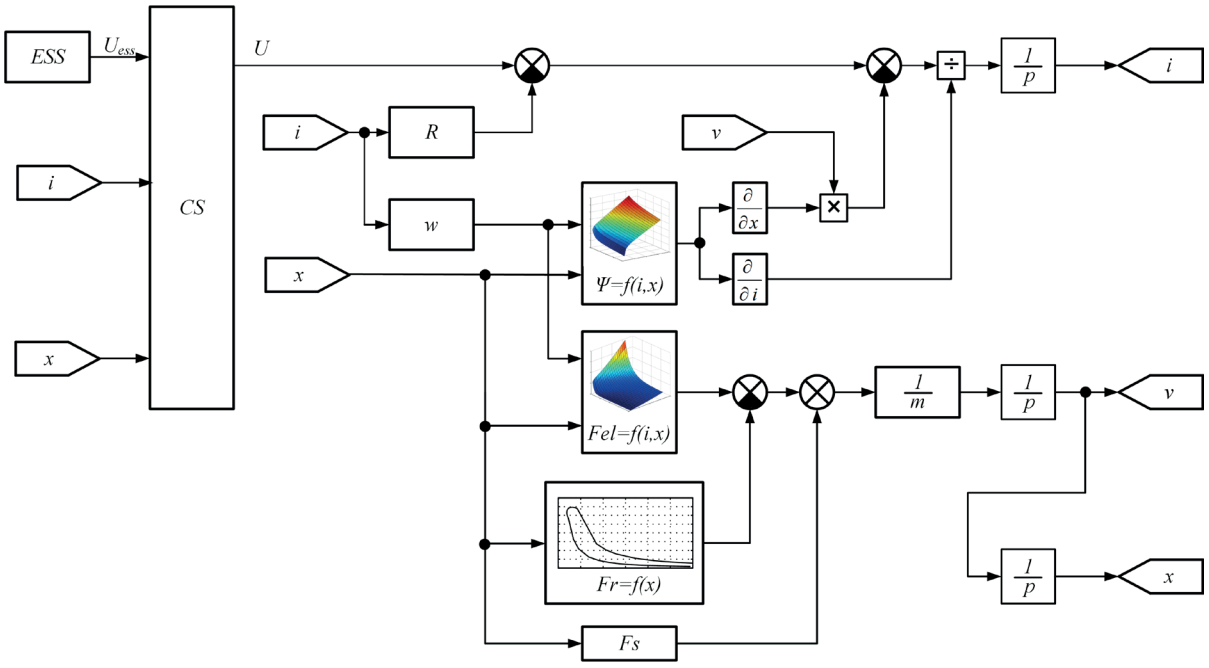


Figure 9 *Free piston engine operation flow chart*

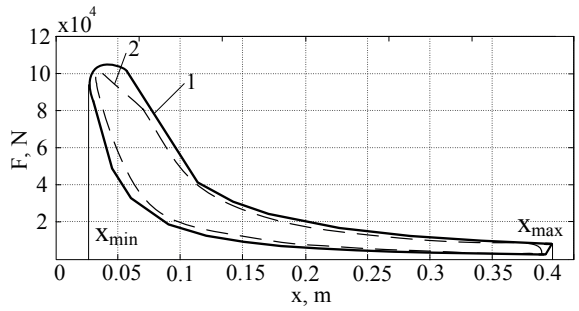


Figure 10 *Free piston engine operation simulation results: 1 - force of ICE, 2 - force of linear EMC*

supply it in an engine mode, was used as an energy storage device.

As a result of the simulation, oscillograms of the forces of one module generator, that includes one cylinder of K6S310DR diesel engine and one linear generator, are obtained (Figure 10).

The difference between the voltage values on the capacitor at the beginning of the cycle and at its end was 34 V; the average speed of the piston was 10 m/s. Based on the obtained diagram (Figure 9), according to Equations (5) and (6), the useful powers of the ICE cylinder and linear generator were calculated:

$$P_{dies} = \int_{x_{\min}}^{x_{\max}} F_{\exp}(x)dx - \int_{x_{\min}}^{x_{\max}} F_{comp}(x)dx, \quad (5)$$

where:
x_{min} - minimum piston movement,
x_{max} - maximum piston movement,
F_{exp} - ICE force during expansion stroke,
F_{comp} - force required to compress the fuel-air mixture.

$$P_{gen} = \int_{x_{\min}}^{x_{\max}} F_{gen}(x)dx - \int_{x_{\min}}^{x_{\max}} F_{eng}(x)dx, \quad (6)$$

where:
x_{min} - minimum piston movement,
x_{max} - maximum piston movement,
F_{gen} - EMC force during a power generation mode,
F_{eng} - EMC force in engine mode.

For the diesel engine, the useful power was 207.6 kW, and for the generator 152.3 kW.

The coefficient of the full utilization of the diesel engine power, which is equal to the ratio of the generator power to the diesel engine power, was 73.36%. At the same time, the fuel explosion energy not used by the generator is accumulated in the spring, which allows increasing the starting force of the EMC in the engine mode.

4 Conclusions

The use of a free piston engine with a linear generator of electromagnetic type as a diesel locomotive power plant is proposed. The suggested design has fewer parts and allows excluding the appearance of lateral forces acting on the piston. As a prototype and for the purpose of unifying the engines, the K6S310DR diesel engine from the ChME3 diesel locomotive is taken as a basis. Under the conditions of the greatest approximation of the forces of the internal combustion engine and the strength of the linear generator, the configuration of the generator is chosen. As a generator a linear electromagnetic generator was employed. For this type of generator the change in magnetic induction depending on the position of the armature is inherent; for this reason, the values of force and flux linkage of the winding depending on the movement of the armature were calculated. The values obtained as a result of this calculation were approximated by Chebyshev polynomials to obtain continuous dependencies throughout the calculation section. Functional and structural schemes of the simulation model of the engine operation were drawn. In the MatLab package environment, the suggested system was simulated and the graphs of the realization of the forces of the ICE and LEMC were obtained. The useful power of the diesel engine was 207.6 kW, and for the generator 152.3 kW, while the coefficient of the full utilization of the diesel power was 73.36%.

The suggested design of the power plant is a promising direction in the development of transport engineering. The use of different types of linear EMCs, in the future, allows us to work not only on two-stroke cycles of a diesel engine but also on other, more efficient thermodynamic cycles.

References

[1] BAPI RAJU, V., PHANINDRA, Y. B. S. Optimization of internal combustion engine piston. *International Journal for Research in Applied Science and Engineering Technology* [online]. 2018, **6**(VI), p. 1699-1705 [accessed 2019-10-12]. ISSN 2321-9653. Available from: https://doi.org/10.22214/ijraset.2018.6250

[2] ISERMANN, R. On the control and diagnosis of internal combustion engines. In: *Combustion Engine Diagnosis* [online]. Berlin, Heidelberg: Springer Vieweg, 2017, ATZ/MTZ-Fachbuch book series. ISBN 978-3-662-49466-0. eISBN 978-3-662-49467-7, p. 51-73 [accessed 2019-10-13]. Available from: https://doi.org/10.1007/978-3-662-49467-7_3

[3] YUAN, C., XU, J., HE, Y. Parametric study on the starting of a free-piston engine alternator. *International Journal of Engine Research* [online]. 2018, **19**(4), p. 411-422 [accessed 2019-10-11]. ISSN 1468-0874, eISSN 2041-3149. Available from: https://doi.org/10.1177/1468087417712161

[4] CIPEK, M., PAVKOVIC, D., KLJAIC, Z., MLINARIC, T. Assessment of battery-hybrid diesel-electric locomotive fuel savings and emission reduction potentials based on a realistic mountainous rail route. *Energy* [online]. 2019, **173**, p. 1154-1171 [accessed 2019-10-11]. ISSN 1468-0874, eISSN 2041-3149. Available from: https://doi.org/10.1016/j.energy.2019.02.144

[5] YATSKO, S., SIDORENKO, A., VASHCHENKO, Y., LYUBARSKYI, B., YERITSYAN, B. Method to improve the efficiency of the traction rolling stock with onboard energy storage. *International Journal of Renewable Energy Research* [online]. 2019, **9(2)**, p. 848-858 [accessed 2019-07-07]. eISSN 1309-0127. Available from: https://www.ijrer.org/ijrer/index.php/ijrer/article/view/9143

[6] BURIAKOVSKYI, S., BABAIEV, M., LIUBARSKYI, B., MASLII, A., KARPENKO, N., POMAZAN, D., MASLII, A., DENYS, I. Quality assessment of control over the traction valve-inductor drive of a hybrid diesel locomotive. *Eastern-*

European Journal of Enterprise Technologies [online]. 2018, **2(91)**, p. 68-75 [accessed 2019-07-07]. ISSN 1729-3774, eISSN 1729-4061. Available from: https://doi.org/10.15587/1729-4061.2018.122422

[7] BARI, S. *Diesel engine. Combustion, emissions and condition monitoring*. London: IntechOpen, 2013. ISBN 978-953-51-6335-0.

[8] Finite element method magnetics: documentation [online]. [Viewed 2019-04-18]. Available from: http://www.femm.info/wiki/Documentation/

[9] KUZNETSOV, B., TURENKO, A., NIKITINA, T., VOLOSHKO, A., KOLOMIETS, V. Method of synthesis of closed-loop systems of active shielding magnetic field of power transmission lines. *Technical Electrodynamics* [online]. 2016, **4**, p. 8-10 [accessed 2019-09-10]. ISSN 1607-7970, eISSN 2218-1903. Available from: https://doi.org/10.15407/techned2016.04.008

[10] SEZEN, S., KARAKAS, E., YILMAZ, K., AYAZ, M. Finite element modeling and control of a high-power SRM for hybrid electric vehicle. *Simulation Modelling Practice and Theory* [online]. 2016, **62**, p. 49-67 [accessed 2019-09-10]. ISSN 1569-190X, eISSN 1878-1462. Available from: https://doi.org/10.1016/j.simpat.2016.01.006

[11] HELMUT, P. Representing derivatives of Chebyshev polynomials by Chebyshev polynomials and related questions. *Open Mathematics* [online]. 2017, **15(1)**, p. 1156-1160 [accessed 2019-08-07]. ISSN 2391-5455. Available from: https://doi.org/10.1515/math-2017-0096

[12] Modeling - MATLAB & Simulink - MathWorks [online] [Viewed 2018-09-18]. Available from: https://www.mathworks.com/help/simulink/modeling.html

[13] PERELMUTER, V. *Electrotechnical systems simulation with Simulink® and SimPowerSystems™*. Boca Raton: CRC Press, 2013. ISBN 9781315216430.

[14] KANUCH, J., GIROVSKY, P. Analysis of the PM motor with external rotor for direct drive of electric wheelchair. *Communications - Scientific Letters of the University of Zilina* [online]. 2019, **21(3)**, p. 66-71 [accessed 2019-10-07]. ISSN 1335-4205, eISSN 2585-7878. Available from: http://komunikacie.uniza.sk/index.php/communications/article/view/1507

[15] BURIAKOVSKYI, S., LIUBARSKYI, B., MASLII, A., POMAZAN, D., PANCHENKO, V., MASLII, A. Mathematical modelling of prospective transport systems electromechanical energy transducers on basis of the generalized model. In: 9th International Conference on Advanced Computer Information Technologies ACIT 2019 : proceedings. 2019. ISBN 978-1- 7281-0449-2, p. 76-79.

[16] KALAIVANI, L., SUBBURAJ, P., WILLJUICE IRUTHAYARAJAN, M. Speed control of switched reluctance motor with torque ripple reduction using non-dominated sorting genetic algorithm. *International Journal of Electrical Power and Energy Systems* [online]. 2013, **53**, p. 69-77 [accessed 2019-09-10]. ISSN 0142-0615. Available from: https://doi.org/10.1016/j.ijepes.2013.04.005

Jorge Rafael Gonzalez-Teodoro - Enrique Romero-Cadaval - Rafael Asensi - Vladimir Kindl

EQUIVALENT ELECTRICAL MODEL OF AN INDUCTOR EXCITED BY A TRIANGULAR CURRENT INCLUDING SATURATION

A model for an equivalent electrical circuit designed for a ferrite (3C90) inductor usually used in power converters excited by a non-sinusoidal current appropriate for use in power electronics is proposed. This study, based on 3D finite element analysis, leads to significant precision advantages over 2D analysis for non-symmetric inductors. The frequency range of the analysis for the toroidal core was between 15 kHz and 1 GHz, with different levels of excitation in non-saturation and saturation status focusing on the power loss.

Keywords: *ferrite inductors, 3d finite element analysis, power losses, saturation, hysteresis, eddy current losses, magnetization*

1 Introduction

Ferrites (3C90) are common used due to their losses data and permeability characteristics [1-6] in power converters. These components have non-linear behaviour that needs to be added in electromagnetic analysis to develop a transient simulation of these power converters [7-11]. Models of these magnetic components can be found in the bibliography [12-14], However, non-linear models that represent non-linear behaviour is an an absence of power converterfield. An electrical-magnetic model of a ferrite inductor valid for triangular current excitations is presented for a 15 kHz to 1 GHz frequency range (the range of switching frequencies used by power electronic converters based on Si, SiC or GaN semiconductors) with different signals to include the saturation status in the analysis.

The core for the inductor component analysed in this work was toroidal because they are common in transformers and inductive components, they do not have symmetry and they cannot be solved using Maxwell's equations in either 1D or 2D finite element analysis (FEA).

A comparison between sinusoidal and triangular excitation currents for the inductor component is included in the analysis as well.

Section II explains the FEM procedure, Section describes the signals used in this study and Section IV is the definition of core loss used for the scripts in the FEM software. At the end, the conclusions from the FEM results are presented.

2 Procedure for parameter extraction

This work aims to achieve inductance, L, and resistance, R, it means, the parameters of the the equivalent circuit for the analysed inductor to obtain the output voltage and current waveform for an input triangular current excitation utilised in the inductor component. *The inducatance and the resistance* connect in series. L depends on the excitation current, $I=f(L,I)$, and R depends on the I_{rms} and frequency, $R=f(I_{rms}$, *freq.) being non linear components.*

To develop the parameter extraction, a 3D component without simplifications was analysed using FEM analysis (Transient Solver with 2% energy error in Maxwell Ansys to obtain the convergence), which involves three different steps: pre-modelling, simulation and post-modelling phases. If the computational limitations do not allow convergence in the FEM simulation, a simplification model is described in [15].

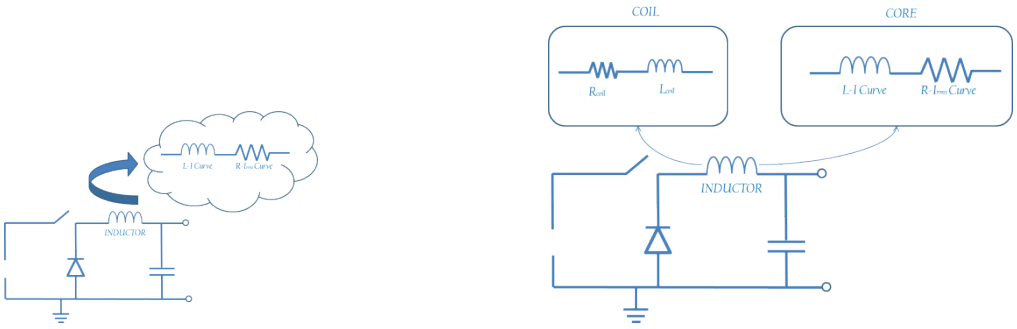
L was estimated applying a triangular current and the BH data [16] in the pre-modelling that characterizes the core material. After the simulation, the L-I curve described in [17] can be defined the following parameter (ϕ-I curve):

$$L(I) = \frac{d\phi}{dI}, \tag{1}$$

where ϕ is the magnetic flux.

R was derived during the post- modelling step R-I_{rms} curve as described in [17]:

Jorge Rafael Gonzalez-Teodoro[1,*], Enrique Romero-Cadaval[1], Rafael Asensi[2], Vladimir Kindl[3]
[1]University of Extremadura, Badajoz, Spain
[2]Politecnica de Madrid, Spain
[3]University of West Bohemia, Pilsen, Czech Republic
*E-mail of corresponding author: jordirgt@hotmail.com



a) *b)*

Figure 1 *Equivalent circuit in Saturation and non-saturation: a) Model 1, b) Model 2*

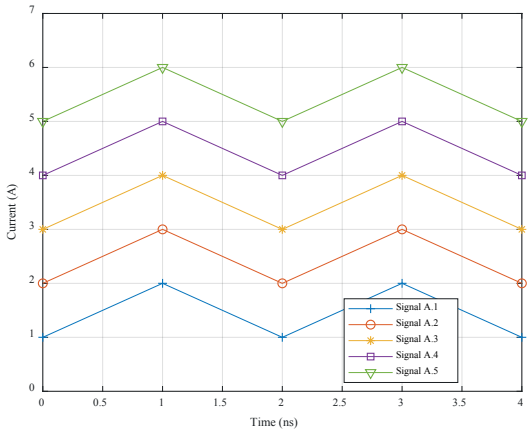


Figure 2 *Triangular signals, Type A*

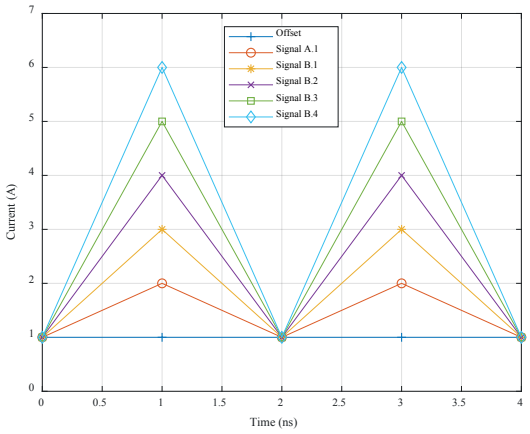


Figure 3 *Triangular signals, Type B*

$$R = \frac{\overline{P}}{I_{rms}^2}, \tag{2}$$

where $\overline{P}$ is the average value of the power and I_{rms} is the *rms* value of the current.

The equivalent resistance of a magnetic component is normally defined as the resistance for a sinusoidal signal instead of a triangular current signal. Nevertheless, the resistance discussed in this work is the average resistance for a triangular signal in a transient analysis developed by FEM.

The inductor component studied is shown in Figure 1(a) (Model 1). This model is valid for saturation and non-saturation status.

In the case of non-saturation and core linear behaviour, it is possible to add the R and L for the winding in the model. See Figure 1(b) (Model 2). These parameters were obtained during the 3D FEM post-modelling, using the procedure in [18] and Equations (3)-(4) to obtain the coefficients for any frequency range for the winding parameters considering a linear system because the superposition theorem was used for deriving the values of the parameters.

In summation, in saturation status, Model 1(a) is used because the core power loss is dominant over other parameters. Model 1(b) is selected for non-saturation including the winding parameters in the equivalent circuit of the inductor component.

$$L_{ij} = \frac{1}{I_0^2} \iiint_V \mathrm{Re}(\vec{B}_{io} \cdot \vec{H}_{jo}^{*}) dv\,, \tag{3}$$

$$R_{ij} = \frac{1}{I_0^2} \iiint_V \frac{1}{\sigma} \cdot \mathrm{Re}(\vec{J}_{io} \cdot \vec{J}_{jo}^{*}) dv\,, \tag{4}$$

3 Signals description

3.1 Triangular signals

Different triangular signals (type A and B) were analysed.

Signal type A is a triangular signal with different offsets and type B is a triangular signal with the same offset, varying the peak value. See Figures 2-3. In total, 9 different triangular signals were analysed.

The I_{rms} for each analysed signal are indicated in Table 1, where A.3, A.4 and A.5 produce core saturation according to the manufacturer's datasheet.

3.2 Sinusoidal signal

The sinusoidal signal selected for comparison with the triangular signal is shown in Figure 4 and Table 2 for the I_{rms} of this signal and the corresponding triangular signal. Signal C.1 was set to have the most similar *rms* value with the triangular signal chosen (signal A.1). All the signals shown are repeated from the frequency range studied.

Table 1 *RMS value for different signals*

Signal Name	I_{rms} (A)	B(T)
A.1	1.286	0.244
B.1	1.570	0.292
B.2	1.856	0.341
B.3	2.141	0.365
A.2	2.286	0.333
B.4	2.426	0.382
A.3	3.286	0.4
A.4	4.286	0.444
A.5	5.286	0.462

Table 2 *RMS value for triangular/sinusoidal signal*

	A.1	C.1
I_{rms} (A)	1.286	1.354

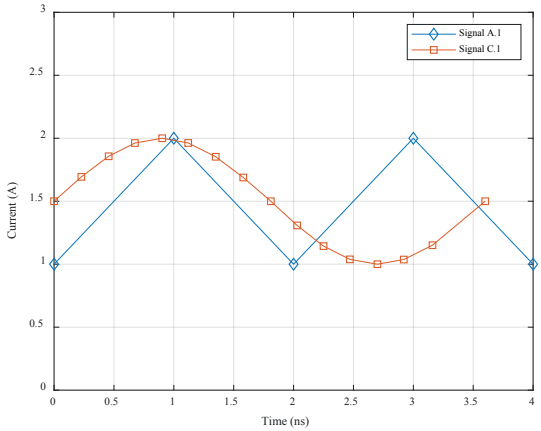


Figure 4 *Triangular signals, Type A, and corresponding sinusoidal signal, Type C*

4 Determination of the core power loss

The core loss calculated by Maxwell Ansys uses the Steinmetz equation with parameters defined by Ansys or modified by the user. This formula presents several issues, the method is only for static, does not predict important frequency/rate dependency and, in addition to linear dynamics, does not capture non-linearity in excess loss [19].

Nevertheless, Ansys Maxwell permits calculating the power loss separately (5), including the hysteresis loop data in the pre-modelling by transient analysis solver.

$$P_t = P_{hysteresis} + P_{eddy}\,. \tag{5}$$

There is another loss, $P_{anomalous}$ (loss due to the material properties modification due to the eddy current); however, for this particular case, it is negligible for the eddy current produced in the ferrite core component.

4.1 Hysteresis loss

During each AC cycle, current flowing in the forward and reverse directions alternatively magnetizes and demagnetizes the core. Energy is lost in each hysteresis cycle within the magnetic core. Energy loss is dependent on the properties (e.g. coercivity) of the core material and is proportional to the area of the hysteresis loop [16]. The calculation during post-modelling is defined by [20]:

$$P_{hysteresis} = \int_{vol} \frac{1}{2} \cdot \omega \cdot \mathrm{Im}(B \cdot H^*) dV\,, \tag{6}$$

where $\vec{B}$ is the magnetic flux density, $\vec{H}^*$ is the complex conjugate of the magnetic field and ω is the angular frequency.

4.2 Eddy current loss

An eddy current is an electric current set up by an alternating magnetic field. Thus, if the core is manufactured

Table 3 *Dissipated energy vs Irms*

B(T)	0.244	0.292	0.341	0.365	0.333	0.382	0.4	0.444	0.462
I_{rms} (A)	1.286	1.570	1.856	2.141	2.286	2.426	3.286	4.286	5.286
E_{surf} (J)	0.063	0.122	0.163	0.173	0.37	0.265	0.71	1.497	1.723
E_{core} (J)	2E-12	3E-12	5E-12	8E-12	4E-12	1E-11	8E-12	1E-11	2E-11
E_{hys} (J)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0027	0.0035	0.0041

Table 4 *Values of coefficients for k vs Irms*

	k_S	k_C	k_H
Non-saturation	$0.036 \cdot I_{rms}^{2.5}$	$2 \cdot 10^{-12} \cdot I_{rms}^{1.4}$	0.0001
Saturation	$0.036 \cdot I_{rms}^{2.5}$	$2 \cdot 10^{-12} \cdot I_{rms}^{1.4}$	$0.0008 \cdot I_{rms}$

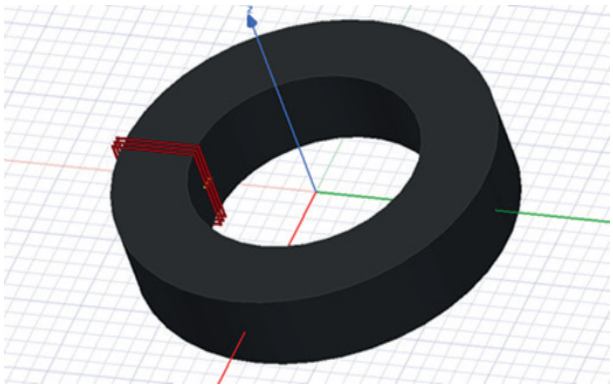


Figure 5 *Inductor analyzed*

with a conductor material, the eddy current losses arise modyfing the flux and producing circulatin current into the core. Eddy current loss depends upon the rate of change of flux as well as the resistance of the path. According to the theory, it expects that the loss changes with the square of both the maximum flux density and frequency if the core has been mmanufactured withferromagnetic materials.. These currents, circulating in the core material cause resistive heating in this material [20]:

$$P_{core} = \int_{vol} \frac{1}{2\sigma} \cdot \mathrm{Re}(J \cdot J^*) dV \,, \tag{7}$$

where J is the current density, J^* is the complex conjugate of the current density and σ is the material conductivity for the core. For the core surfaces, the eddy loss is given by [20]:

$$P_{surface} = \sqrt{\frac{\omega \mu_0 \mu_r}{8\sigma}} \cdot \int H_t \cdot H_t^* ds \,, \tag{8}$$

where $\vec{H}_t$ is the tangential component of $\vec{H}$ on the boundary and $\vec{H}_t^*$ is the complex conjugate tangential component of $\vec{H}$ on the boundary.

Consequently, the total eddy current loss is defined as:

$$P_{eddy} = P_{core} + P_{surface} \,. \tag{9}$$

5 Analysis using 3D FEM

The selected Toroidal component core, C107.65.25 with 1 winding (4 turns with 1° lateral distance) using a 3C90 material, has been chosen because it is a non-symmetric component [15] and 3D FEM analysis has more precision than 2D FEM analysis, see Figure 5.

The FEM analyses were performed with the magnetization choice selected for the core in the FEA tool solver to apply Equations (6)-(9) for calculating the losses.

The results for the 9 different triangular signals are shown in *Appendix I*, where the losses (hysteresis and eddy) and the magnetic field density peak for the different cases are presented. The hysteresis losses from signal Type *A* become linear at high frequencies when the I_{rms} increases due to the magnetization impact. The hysteresis losses reach larger peaks than Type *B* for cases where the peak current value is the same.

The eddy losses (surface) suffer an inrush effect to get to a constant value for Type *A* signals. Type *B* signals do not reach the saturation state on the core surface.

The eddy losses (core) have a sinusoidal behaviour according to the current for both types.

Since the analyses were developed from 15 kHz to 1 GHz, they can be used to calculate the dissipated energy. See *Appendix II*, where the energy from hysteresis losses and eddy current losses are shown. It is evidence that

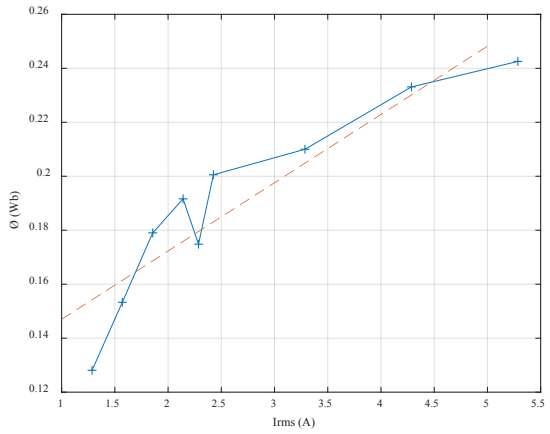


Figure 6 *Relationship between ϕ and I_{rms}*

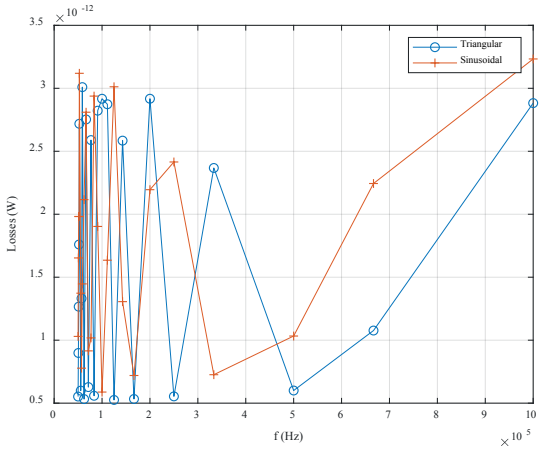


Figure 8 *Comparison between A.1 vs C.1 (Eddy Loss at the core)*

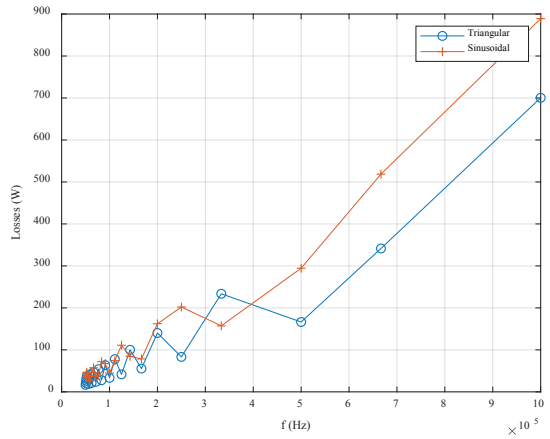


Figure 7 *Comparison between A.1 and C.1 (Hysteresis Loss)*

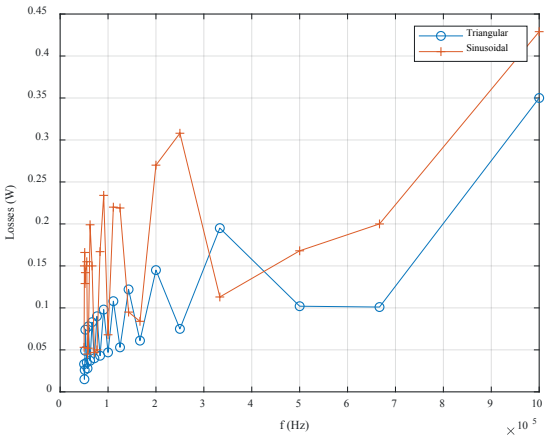


Figure 9 *Comparison between A.1 and C.1 (Eddy Loss at the surface)*

dissipated energy from eddy currents has a linear tendency but the energy from hysteresis currents has a quadratic one with the frequency.

See Table 3 for the dissipated energy by nature and corresponding core peak magnetic field density produced. At this point, the dissipated energy for the inductive component versus frequency can be predicted. Mathematical regressions from the energies results from the FEA tool were calculated; thus, the eddy current is divided into the energy at the boundary (10) and the core, (11) and the energy for hysteresis is defined in (12).

$$E_{surface} = k_S \cdot f \,. \tag{10}$$

$$E_{core} = k_C \cdot f \,. \tag{11}$$

$$E_{hys} = k_H \cdot f^2 \,. \tag{12}$$

The dissipated energy for the component built with ferrite can be calculated by summing the energies (10)-(12). The specific coefficients, k_S, k_c and k_H, are indicated in Table 4 for saturation and non-saturation status for the core used in this analysis and f is the frequency. Based on the results, the dissipated energies from the eddy current are equal independent of the saturation; however, the hysteresis energy suffers a tendency to change when the core is in saturation.

The magnetic density peak for each signal according to the I_{rms} has been plotted in Figure 6 to demonstrate that there is a linear relationship between ϕ and I_{rms} for the analysed inductor.

The comparison of the hysteresis and eddy losses between the triangular signal (A.1) and the sinusoidal (C.1) is shown in the Figures 7-9. The tendency of the hysteresis loss of Signal C.1. is similar to the loss of Signal A.1. The values of the losses are in agreement with the I_{rms} of the signals. The values for the eddy losses from the core have similar values and are negligible from the eddy losses from the surface. The loss due to the eddy current for the core surface is larger in Signal C.1.

Since electronics engineers design inductors for working at a defined operation point in terms of frequency, it is necessary to have a dedicated analysis for this frequency operation point to obtain the difference for the output voltage for different signals.

In saturation status, the inductance (1) and resistance (2) for the core can be introduced in the simulator PSIM to

obtain the voltage and current waveforms. If the core is not saturated, the winding inductance (3) and resistance (4) needs to be added into the model inductor in the simulator for a defined frequency range.

The original contribution of this work that it can modify the electrical parameters from one operating frequency point and I_{rms} to another without performing any FEM simulation using:

$$P = \frac{E}{T}, \tag{13}$$

where E *is* $E_{surface}+E_{core}+E_{hys}$, calculated previously, and T is *1/f* to obtain the resistance. The inductance modification according to the frequency and I_{rms} is shown in Figure 6 with a linear relationship. With these results, it comes back to Equations (1)-(2) to obtain R and L at the desired frequency and signal for the core.

6 Conclusions

A model of an equivalent electrical circuit y designed for inductors made with 3C90 core used with triangular waveforms is presented in this manuscript.

The electrical parameters for the core and the coil, depending on the saturation status, have been estimated using a 3D FEM-model from 15 kHz to 1 GHz. The 3D model was used to involve all the high-frequency effects in the analysis that cannot be calculated in 2D.

This work focused on the behaviour of the core power loss for toroidal components excited by triangular signals (9 different triangular signals were selected).

Determining the core power loss was divided by nature and they were not calculated for the Steinmez Equation.

The first conclusion of this paper is that the tendency of the energy dissipated is linear with the frequency for eddy losses and quadratic for hysteresis losses. The specific coefficients for the case shown in the paper are dependent on the I_{rms} and independent on the triangular shape excitation.

A comparison between the triangular and sinusoidal waveforms is presented along with the long-term effect that is not dependent on the signal nature and depends on the value of I_{rms}.

The original contribution of this study is the capacity to modify the electrical parameters in the simulator PSIM according to the frequency point and I_{rms} using the conclusions in this paper, as indicated in Equation (13) and Figure 6 where the evolution of R and L for the Toroidal inductor depend on the frequency and current, without developing a new 3D FEM analysis to calculate the output waveforms.

Acknowledgement

This research has been supported by the Junta de Extremadura, under project IB18067 (with support of European Regional Development Fund) and by the Ministry of Education, Youth and Sports of the Czech Republic under the project OP VVV Electrical Engineering Technologies with High-Level of Embedded Intelligence CZ.02.1.01/0.0/0.0/18_069/0009855 and by funding program of the University of West Bohemia number SGS-2018-009.

Appendix

Appendixes I and II where the evolution of the core loss and magnetic field density are plotted depending on the input current used for the FEM analysis are at the end of the paper.

References

[1] FISH, G. E. Soft magnetic materials. *Proceedings of the IEEE* [online]. 1990, **78**(6), p. 967-972. ISSN 0018-9219. Available from: https://doi.org/10.1109/5.56909

[2] SNELING, E.C. *Soft ferrites: properties and applications*. 2. Ed. Oxford, U.K.: Butterworths, 1998. ISBN 978-0408027601,

[3] MOHAN, N.. UNDELAND, T. M.. ROBBINS, W. P. *Power electronics: converters, applications and design*. 3. ed. Hoboken, NJ: Wiley, 2002. ISBN 978-0-471-22693-2.

[4] KOLEDINTSEVA, M. Y. Soft ferrites for EMC applications. In: IEEE Symposium on Electromagnetic Compatibility, Signal Integrity and Power Integrity EMC, SI & PI 2018 : proceedings. IEEE. 2018. p. 1-31.

[5] DAKOVA, L., FUZER, J., DOBAK, S., KOLLAR, P., OSADCHUK, Y. G., STRECKOVA, M., FABEROVA, M., BURES, R., KUREK, P., VOJTKO, M. Analysis of magnetic losses and complex permeability in novel soft magnetic composite with ferrite nanofibers. *IEEE Transactions on Magnetics* [online]. 2018, **54**(12), p. 1-6. ISSN 0018-9464, eISSN 1941-0069. Available from: https://doi.org/10.1109/tmag.2018.2866814

[6] RIMAL, H. P. Dynamic model of soft ferrites for avionic applications. In: IEEE 4th International Forum on Research and Technology for Society and Industry RTSI 2018 : proceedings. IEEE. 2018. ISBN 978-1-5386-6282-3, p. 1-5.

[7] TRELA, K., GAWRYLCZYK, K. M. Frequency response modeling of power transformer windings considering the attributes of ferromagnetic core. In: International Interdisciplinary PhD Workshop IIPhDW 2018 : proceedings [online]. IEEE. 2018. p. 71-73. Available from: https://doi.org/10.1109/IIPHDW.2018.8388328

[8] MATSUMORI, H., SHIMIZU, T., WANG, X., BLAABJERG, F. A practical core loss model for filter inductors of power electronic converters. *IEEE Journal of Emerging and Selected Topics in Power Electronics* [online]. 2018, **6**(1), p. 29-39. ISSN 2168-6777. Available from: https://doi.org/ 10.1109/JESTPE.2017.2761127

[9] HE, R., ZHANG, Y., ZHANG, D., XIE, D. An improvement of core losses estimation model in power electronic transformer. In: IEEE Student Conference on Electric Machines and Systems 2018 : proceedings. IEEE. 2018. p. 1-5.

[10] ZEGADI, L., ROUSSEAU, J. J., ALLARD, B., TENANT, P., RENAULT, D. Model of power soft MnZn ferrites, including temperature effects. *IEEE Transactions on Magnetics* [online]. 2000, **36**(4), p. 2022-2032. ISSN 0018-9464, eISSN 1941-0069. Available from: https://doi.org/10.1109/20.875308

[11] WILSON, P. R., ROSS, J. N., BROWN, A. D. Modeling frequency-dependent losses in ferrite cores. *IEEE Transactions on Magnetics* [online]. 2004, **40**(3), p. 1537-1541. ISSN 0018-9464, eISSN 1941-0069. Available from: https://doi.org/10.1109/TMAG.2004.826910

[12] SHIXIA, M., HONGFU, W., JUNXIAN, H., TIEZHU, W. Fast electromagnetic transient simulation model of photovoltaic power system. In: International Conference on Power System Technology POWERCON 2018 : proceedings. 2018. p. 286-291.

[13] COSSART, Q., COLAS, F., KESTELYN, X. Simplified converters models for the analysis and simulation of large transmission systems using 100% power electronics. In: 20th European Conference on Power Electronics and Applications EPE'18 ECCE Europe 2018 : proceedings. IEEE. 2018. p. P.1-P.10.

[14] KULMANOV, V., ANUCHIN, A., OSTRIROV, V., RUSAKOV, A., VAGAPOV, Y. Simulation and testing of a power converter for aircraft AC electric power generation system using software elimination of higher harmonics. In: 53rd International Universities Power Engineering Conference UPEC 2018 : proceedings [online]. IEEE. 2018. p. 1-6. Available from: https://doi.org/10.1109/UPEC.2018.8541922

[15] GONZALEZ-TEODORO, J. R., PRIETO, R., ASENSI, R. Simplifications in 3D high-low frequency models of multiwinding magnetic components (EE & toroidal cores). *International Journal of Magnetics and Electromagnetism* [online]. 2015, **1**(1), p. 1-7. ISSN 2631-5068. Available from: https://doi.org/10.35840/2631-5068/6503

[16] Data handbook of ferroxcube - Ferroxcube International Holding B.V. [online] [Viewed 2017-11-09]. Available from: https://www.ferroxcube.com/en-global/download/index

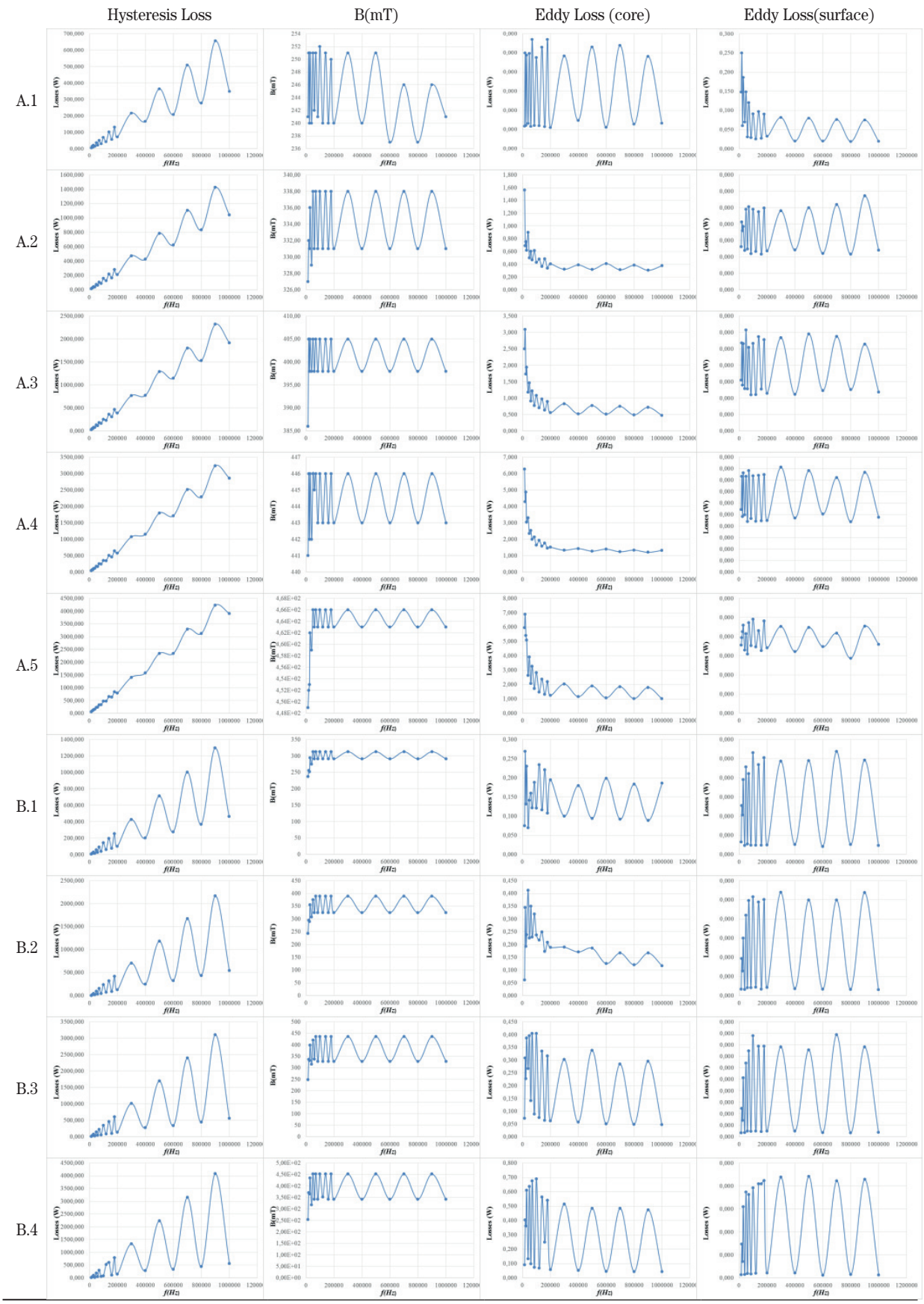
[17] SALAS, R. A., PLEITE, J. Equivalent electrical model of a ferrite core inductor excited by a square waveform saturation and power losses for circuit simulation. *IEEE Transactions on Magnetics* [online]. 2013, **49**(7) p. 4257-4260. ISSN 0018-9464, eISSN 1941-0069. Available from: https://doi.org/10.1109/TMAG.2013.2238223

[18] ASENSI, R., PRIETO, R., COBOS, J. A., UCEDA, J. Modeling high-frequency multiwinding magnetic component using finite-element analysis. *IEEE Transactions on Magnetics* [online]. 2007, **43**(10), p. 3840-3850. ISSN 0018-9464, eISSN 1941-0069. Available from: https://doi.org/10.1109/TMAG.2007.903162

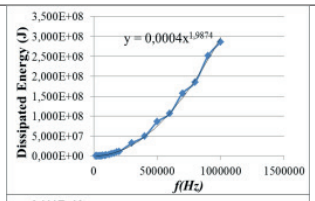
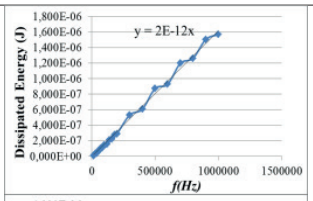
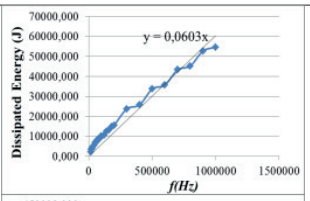
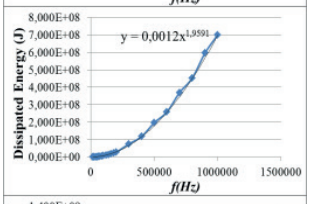
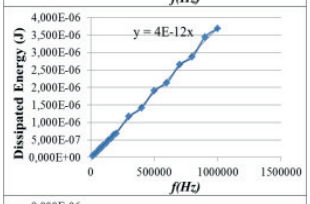
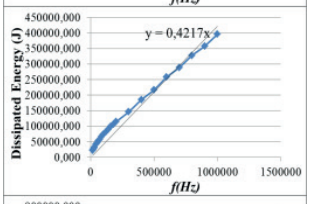
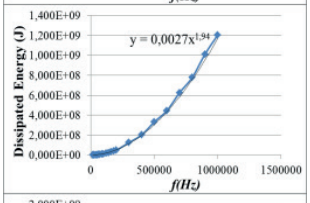
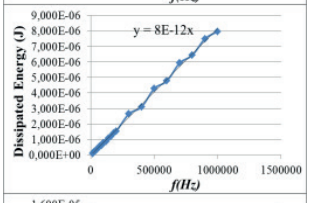
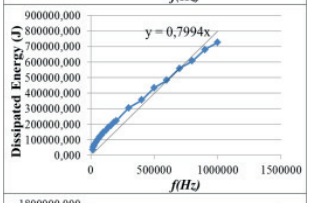
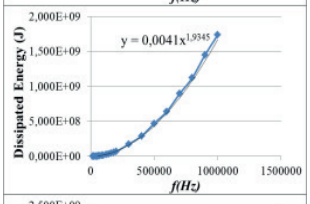
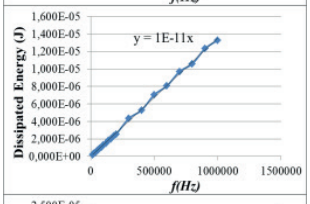
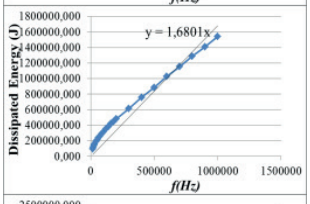
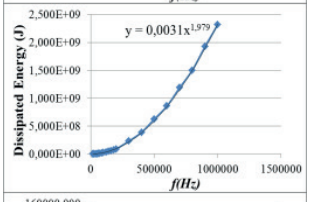
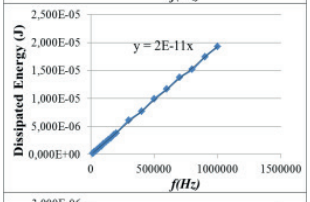
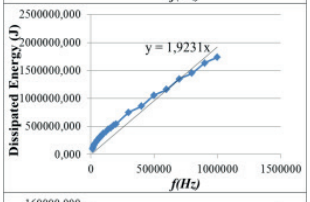
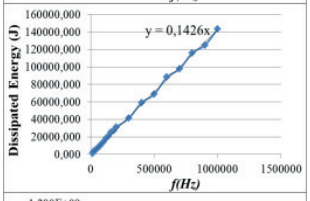
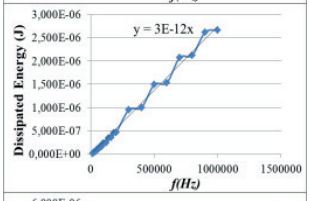
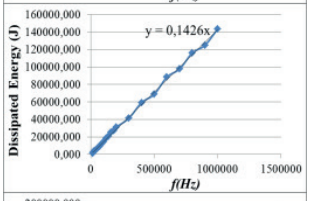
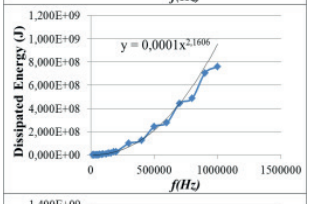
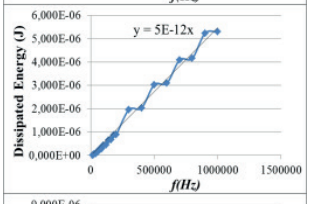
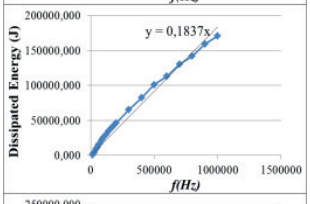
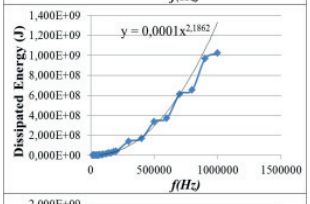
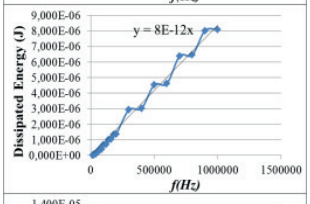
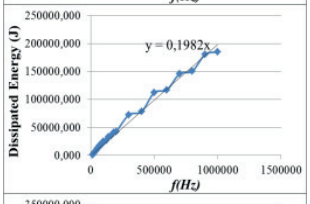
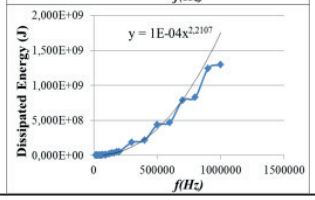
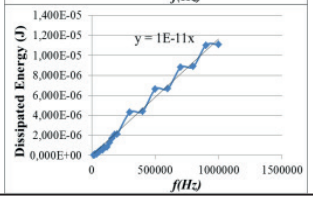
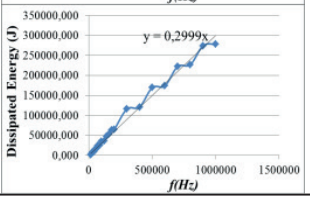
[19] ALBACH, M.. DURBAUM, T., BROCKMEYER, A. Calculating core losses in transformers for arbitrary magnetizing currents - a comparison of different approaches. In: IEEE Power Electronics Specialists Conference : proceedings. 1996. p. 1463-1468.

[20] ANSYS Maxwell - ANSYS, Inc. [online]. Available from: https://www.ansys.com/es-es/products/electronics/ansys-maxwell

Appendix I


Hysteresis Loss
B(mT)
Eddy Loss (core)
Eddy Loss(surface)
A.1
A.2
A.3
A.4
A.5
B.1
B.2
B.3
B.4
Losses (W)
B(mT)
f(Hz)

Appendix II

	Hysteresis	Eddy Loss (core)	Eddy Loss(surface)
A.1			
A.2			
A.3			
A.4			
A.5			
B.1			
B.2			
B.3			
B.4			

Jan Nedoma - Michal Kostelansky - Michael Fridrich - Jaroslav Frnda - Miroslav Pinka - Radek Martinek
Martin Novak - Stanislav Zabka

FIBER OPTIC PHASE-BASED SENSOR FOR DETECTION OF AXLES AND WHEELS OF TRAM VEHICLES

This paper presents a novelty approach to usage of the fiber-optic phase-based sensor in railway transportation. We designed and tested the real deployment of this sensor working on the principle of light interferences within optical fibers. The proposed construction of the sensor allowed to increase the sensitivity and thanks to this can be detected and calculated individual axles and wheels of tram vehicles. We performed long-time period measurements (April to September 2019) in diverse climatic conditions, including measurements of 642 tram passages (several different construction types) in real urban traffic. The detection accuracy level was slightly above 99.4 %.

Keywords: *vehicle transport, fiber-optic, sensor, interferometric sensor, Mach-Zehnder*

1 Introduction

In order to maintain the safety of railway operation (tram or train), it's necessary to know the exact position and number of carriages within a rail vehicle. Nowadays, wheel detectors or axle counters are used for this purpose. They are characterized by relatively old technology methods with gradual descending reliability. Since these methods do not meet the contemporary criteria, many research teams try to find alternative approaches. One of them - fiber-optic sensors - seems to be a potential solution.

Research cooperation between the Faculty of Electrical Engineering and Computer Science (Czech Republic) and Faculty of Operation and Economics of Transport and Communications (Slovakia) has brought several interesting research outputs in this field [1-4]. This paper directly follows and extends our previously published study [2] in which we present interferometric sensor primarily used for tram vehicle detection, as well as for detection of frequencies generated during trams passage. We point out the ability to detect individual tram axles. Since the sensor sensitivity was originally not created for the detection of wheels and axles, the detection veracity was low (less than 50 %). Figure 1 shows the comparison of the tram (the same type of tram) passage detection obtained through the sensor described in the paper [2] and via the same sensor [2] in another day (individual axles cannot be identified).

It can be seen that the sensitivity of the sensor is not sufficient, and influences like the weather (the stronger wind is sufficient) can cause those individual axles cannot be detected with a high success rate.

In the traffic industry, fiber-optic sensors represent an alternative monitoring technique used to analyze basic parameters such as vehicle detection, traffic density, speed measuring or even vehicle weighing. The main advantage of these sensor types lies in their small size and weight. If suitable materials are used, the sensors are resistant to electromagnetic interference (EMI). They also offer a possibility of remote measuring evaluation (place of measurement is separated from the place of evaluation) with regards to the power of radiation source and attenuation of a connected optical fiber (between the sensor and evaluating unit). The telecommunication sector has brought significant development of optical fiber components and items of which most can be used for sensor applications. Therefore, the final price of the proposed sensor system is admissible.

2 State-of-the-art

In this section, we provide an overview of sensor types oriented to railway transportation. Monitoring of the track occupation was the primary task of sensor deployment in this sector. The classic approach uses track circuits and operates on the principle of separated parts of track division (so-called blocks). When a train is passing via the block, a circuit is created between the

Jan Nedoma[1], Michal Kostelansky[1], Michael Fridrich[1], Jaroslav Frnda[2,*], Miroslav Pinka[3], Radek Martinek[4], Martin Novak[1], Stanislav Zabka[1]
[1]Department of Telecommunications, Faculty of Electrical Engineering and Computer Science, VSB - Technical University of Ostrava, Czech Republic
[2]Department of Quantitative Methods and Economic Informatics, Faculty of Operation and Economics of Transport and Communications, University of Zilina, Slovakia
[3]Department of Geotechnics and Underground Engineering, Faculty of Civil Engineering, VSB - Technical University of Ostrava, Czech Republic
[4]Department of Cybernetics and Biomedical Engineering, Faculty of Electrical Engineering and Computer Science, VSB – Technical University of Ostrava, Czech Republic
*E-mail of corresponding author: jaroslav.frnda@fpedas.uniza.sk

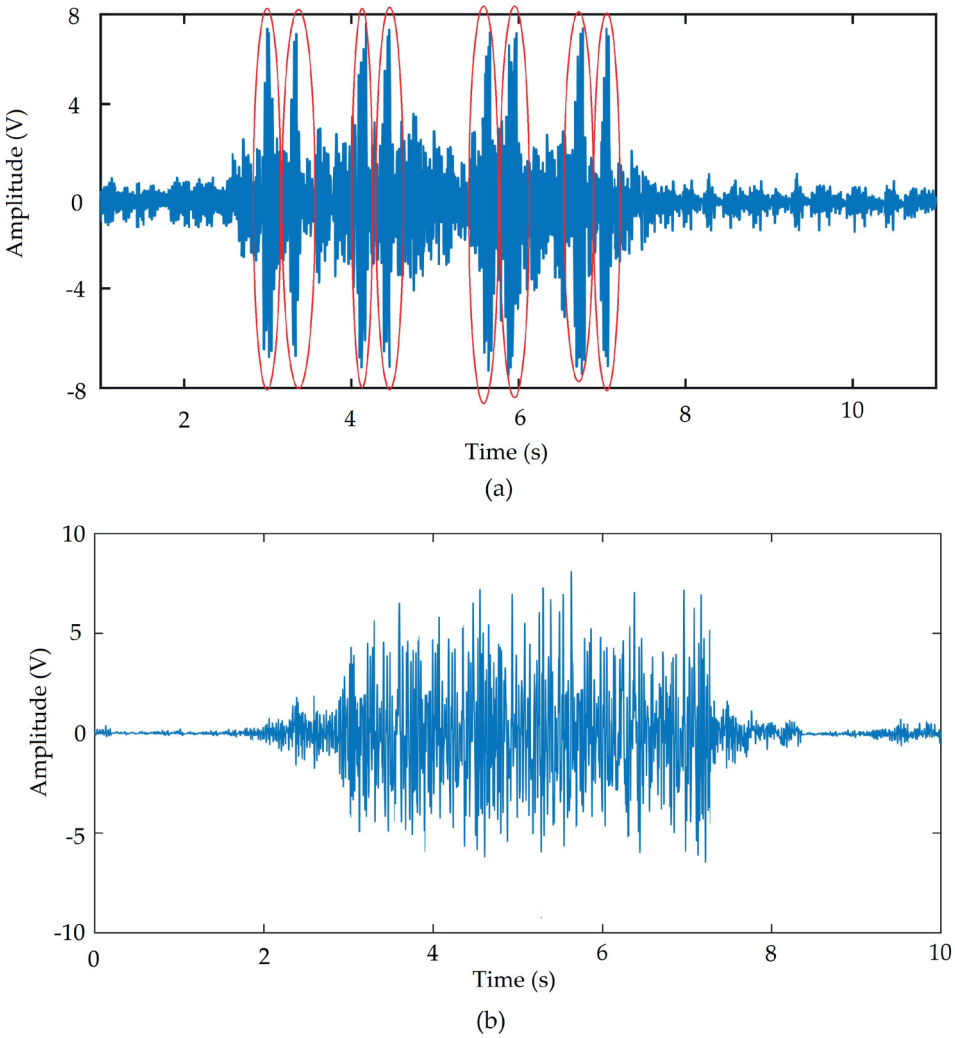


Figure 1 *Comparison of tram vehicle detection; (a) successful tram axles detection described and presented in paper [2], (b) results from another day recorded by the same sensor presented in paper [2] (individual axles cannot be identified)*

rails, which leads to recognition of the train within the block [5-7].

Another solution is based on wheel detectors or axle counters that require the flange for contact connection. This system is typically used not only for train identification, but also for detection of speed and direction [8-10]. The Light Detection and Ranging (LIDAR) instrument (installed close to the rails) emits rapid laser signal (thousands of pulses per second) and gathers signals bouncing back from obstacles. This tool can measure the train speed, the curvature of the track or the train position [11-14]. The European Train Control System (ETCS) is a current signaling and controlling system. Transceiver Eurobalise is installed between rails of each track. This element transmits information to the train, and it is energized by the power from the train antenna [12-14].

Only a few published papers have focused on the usage of fiber-optic interferometric sensors for traffic applications so far. An interesting study is presented in [15], in which the authors show measurements from a fiber-optic interferometer to detect trains in the Prague subway system. The measuring arm of the interferometer was fixed on the glass pad, and the authors proclaimed that they reached a higher sensitivity level. The research papers [16-17] proposed a system based on 5three-armed Mach-Zehnder interferometer consisting of one or more passive fiber trackside sensors and x86 family of instruction set architecture microprocessor. The system is able to measure traffic density within the rail transportation. In a study [18], a research team led by professor Li used fiber-optic Michelson interferometer for tram detection. With a total of 1435 passages in a selected tram, they obtained a 100 % success ratio. Paper [19] deals with an acoustic fiber-optic sensor for monitoring of the railway network from an extensive distance grounded on an interferometric connection. The paper [20] pointed out the use of a fiber-optic interferometer for the perimetric applications. The authors created a sensor based on the Mach-Zehnder interferometer to detect vibration response generated by people moving around the sensor. According to the authors, sensor can be useful for traffic density monitoring, but it has not been tested in real situations yet.

A notable contribution was presented by the paper [21]. Its authors used the interferometric sensor for the

Table 1 *Summary of the basic parameters of the mentioned fiber-optic sensors*

Type of sensor	Frequency range (Hz)	Size (mm)	Weight (kg)	Price ($) Sensor
Actual	4-160	400 x 350 x 95	1.4	350
from paper [2]	2-100	500 x 500 x 130	3	500

utilization of road information, such as vehicle type or its speed. Optical fiber had been fixed on the road surface, and in order to increase the sensitivity, an optical Fabry-Perot (F-P) fiber interference was selected. None of the above-mentioned publications is primarily focused on the detection of individual axles or wheels of tram vehicles with regard to fiber-optic technology.

Our paper proposes innovative interferometric sensor based on Mach-Zehnder two-armed interferometer, when by the specific construction of measuring and reference arms, storing both couplers, and innovative design changes of the sensor measuring part led to increased sensibility in such a level that individual axles of tram vehicles could be detected with high accuracy of more than 99 %. Sensor functionality was performed by a long measurement period lasted for 17 days in various climate conditions, observing 642 tram vehicle passages in real urban traffic of Ostrava city (Radvanice street, Czech Republic). As shown in the results below, the success rate of the detection of individual axles was 99.46 % regardless of the type, weight, and length of the tram vehicle.

Below we present the basic parameters of the proposed interferometric sensor in Table 1. This table shows the comparison of construction parameters and total prices. Our novel proposed sensor has a higher frequency range while having a smaller size and less weight than the sensor mentioned in paper [2]. Also, the price is lower due to changes in sensor design (the price is calculated based on the choice of individual components that were purchased individually).

3. Methods

3.1 Fiber-optic interferometry background

The interferometric sensors are based on the well-known physical phenomenon called interference. The practical output of this paper lies in a two-arm interferometer that allows detecting superpositions of two waves that have traversed various distances while having different phases. We distinguish three basic parameters which can influence the result phase changes according to the Equation (1):

$$\Delta\varphi = 2\pi\frac{n}{\lambda}\delta L + 2\pi\frac{L}{\lambda}\delta n - 2\pi nL\left(\frac{1}{\lambda^2}\right)\delta\lambda\,. \tag{1}$$

Based on the above, wave phase $\Delta\varphi$ change depends on the length of L path, as well as on refractive index n and wavelength λ. As for the Equation 1, its first and second parts describe phase changes in measuring arm, while the third part explains phase changes caused by the source of radiation.

The following equations are related to the Mach-Zehnder type interferometer (hereinafter referred to as M-Z). The input intensity of the M-Z type interferometer that has been used is specified by the Equation 2, and it is related to an electrical current that uses the optical sensor also known as a photodetector.

$$I = 2I_0\left\{1 + \cos\left[\frac{2\pi}{\lambda}n(L_1 - L_2)\right]\right\}, \tag{2}$$

where L_1 and L_2 represent the length of measuring and reference arm of the proposed sensor. The expected signal (caused by the passing tram vehicles generating low vibration frequency response ω) on the output of the photodetector can be expressed by Equation 3:

$$i = \varepsilon \times I_0\alpha\cos(\varphi_d + \varphi_s \times \sin\omega t)\,, \tag{3}$$

where ε represents the sensitivity of the photodetector, I_0 represents the medium signal value, α represents losses (primarily caused by the instability of the light polarization) on the interferometer, φ_d represents the changing phase shift, φ_s is the duration of the amplitude, and ω represents a low vibration frequency response applied to the measurement arm of interferometer [22-24].

3.2 Proposed interferometric sensor and evaluation part

Figure 2 shows a diagram scheme of the proposed interferometric system. Used 3 Hz high-pass filter (HP) was used to filter out the DC (direct current) and low frequencies (temperature influences). Diagram scheme further consist of photodetector PbSe (photoconductive lead selenide photodetector), DFB (Distributed Feedback) laser (LD) with central wavelength 1550 nm and output power of 10 mW, and an A/D (Analog/Digital) converter (type NI-USB 6210 measuring device by National Instruments with sampling frequency 500 S/s).

The published results derived from our papers [1-4] served as a basis to design a part of the sensor. In these papers were tested various covering and damping materials for reference arm, as well as different fiber protection, and suitable forms of photodetectors and lasers. Innovative aspects in this paper related to the proposed sensor include:

1. Storing of the measuring (spiral) and reference fiber (spiral).
2. Storing of both couplers.

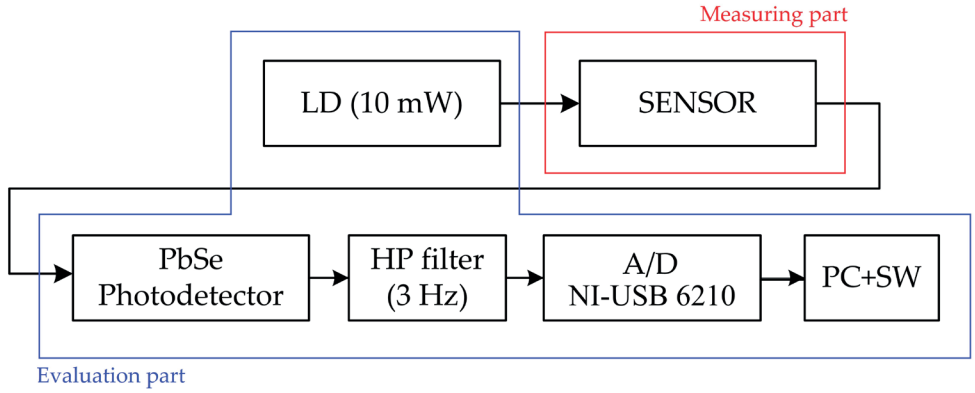
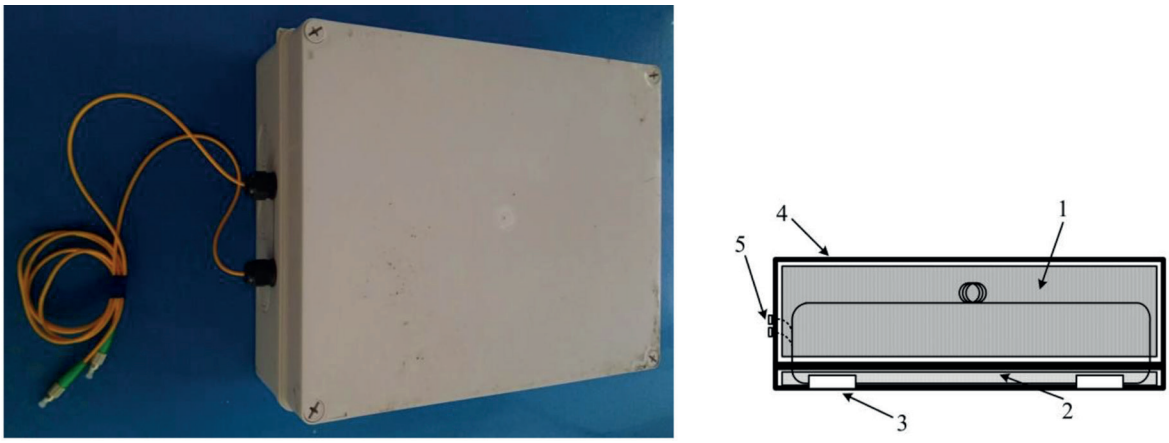


Figure 2 *A diagram scheme of the proposed interferometric system*



a) *b)*

Figure 3 *a) A photo of the proposed interferometric sensor; b) diagram scheme of the proposed interferometric sensor Numeric signs explanation: 1. Reference part (height 8.5 cm), 2. Measuring part (height 1 cm), 3. Both couplers, 4. Protective PVC waterproof box, 5. I/O interface (FC/APC connectors).*

3. Modified length of measuring and reference arm to 3.5 m.
4. Complete encapsulation of the measuring part into one cm high epoxy resin layer.
5. Removing the glass pad (resonator) that was part of the measuring section [2] - now the measuring arm is stored on a 1 mm thin PVC (polyvinylchloride) layer (based on the laboratory measurements), from which the waterproof protective box is made.

Applying of this construction design solution allowed us to reach such a level of sensor sensitivity increase, that even simple two-arm interferometer could detect individual tram vehicle axles (including tram wheels). In our design model, we chose classic telecommunication fibers of G.653 type. The input and output interface was created by screwing connectors FC/APC that secured fixed connection and minimal attenuation. Based on the previous research [4] and [16], we decided to use optical couplers with a coupling ratio of 1:1 (tolerance +/- 5 %). The prototype was covered with a waterproof box. A software application has been developed and enhanced [2] for processing and visualization of the measured data within the LabVIEW environment (National Instruments, Austin, TX, USA). The application loads raw signal from the measuring card, the next step is processing part (filtering of signal to eliminate the noise is included), the final output is a tram passage recording as depicted in Figure 6 and 7 in time and frequency domain. A prototype of the created interferometric sensor is shown in Figure 3.

The most significant construction change was the creation of separate segments for both measuring and reference parts, with both couplers being moved into the measuring part. This solution helped us to achieve higher phase differences between the measuring and reference arms of the interferometer. Measuring part was completely encapsulated into a 1 cm high layer made from epoxy resin, which resulted in maintaining the optimal ratio of Young's modulus of elasticity and density, as well as effective transmission of vibration-acoustic response induced by tram vehicle passage via the measurement fiber.

4 Experimental measurement

The experimental measurement took place in the peripheral part of Ostrava city (Radvanice), where we had official approval to conducted real measurements. The measurements were performed under the various climatic conditions (measurement period lasted from April to September 2019). Figure 4 shows the measurement

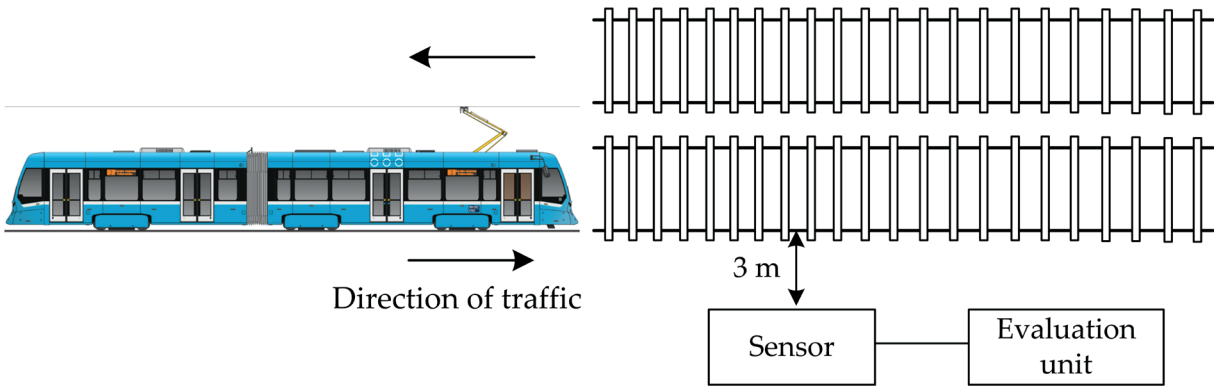
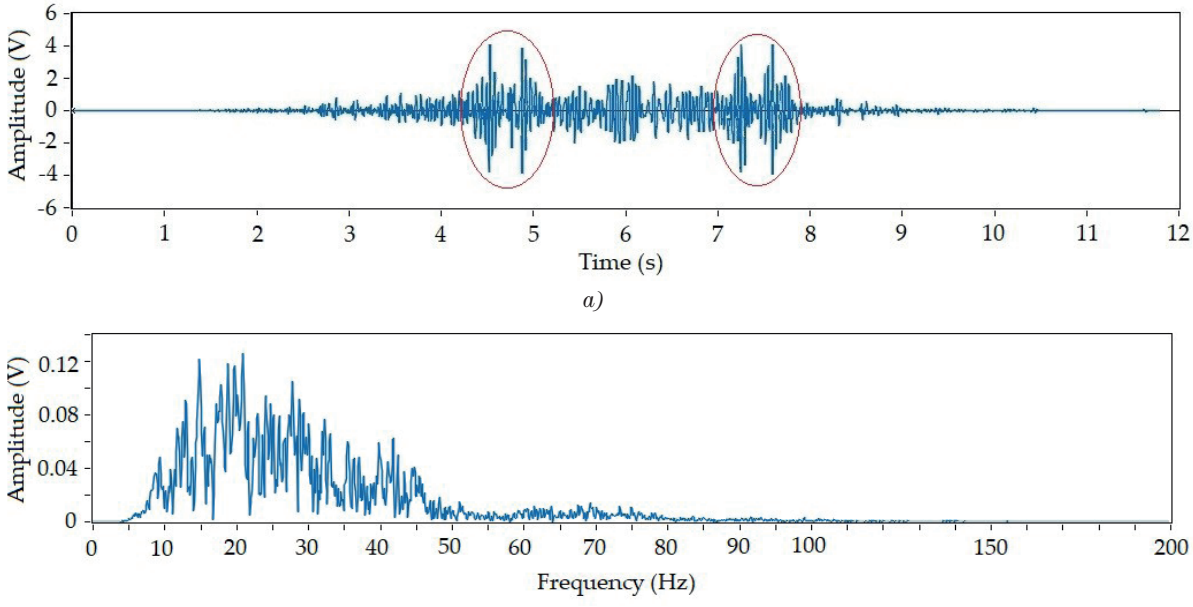


Figure 4 *A diagram scheme of measuring situation*

a) *b)* *c)*

Figure 5 *Tram vehicles photos were taken during the measuring process (the red circle labels the position of the sensor)*



a)

b)

Figure 6 *Recording of passage for tram vehicle with two axles, respective 4 wheels (type Vario LFR); a) time record b) frequency record*

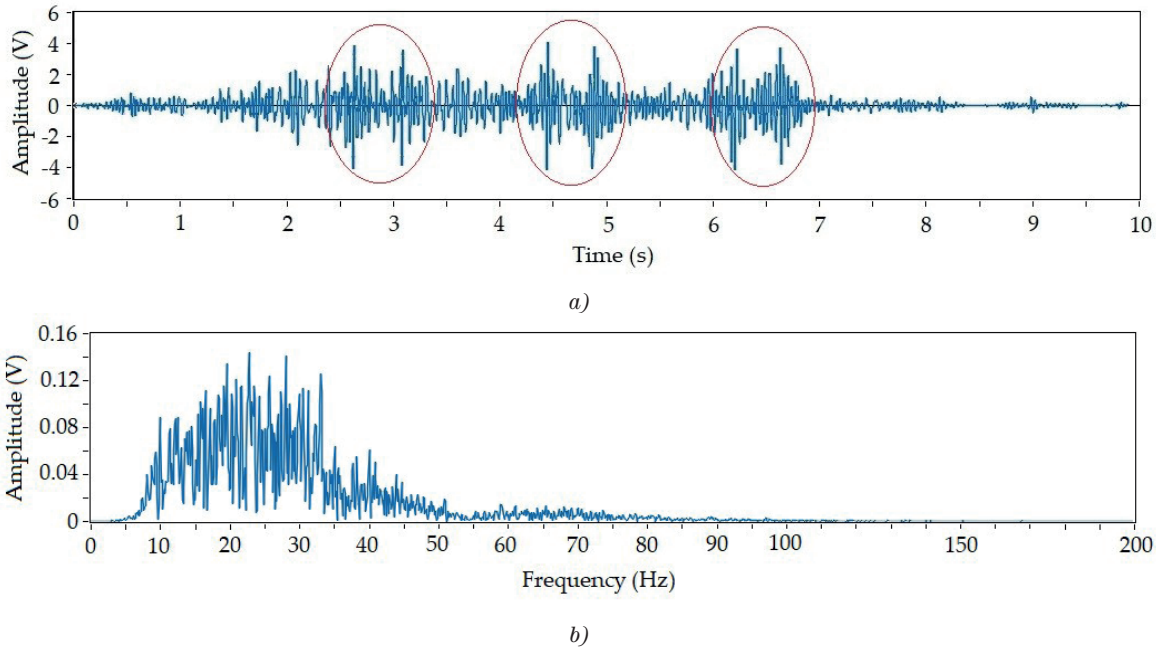


Figure 7 *Passage recording for tram vehicle with 3 axles / 6 wheels (type Stadler Tango NF2); a) time record b) frequency record*

scheme. Tram vehicles speed was reaching 40 - 75 kph (in accordance with the infrastructure limits). Distance between the sensor and the rail foot was set to 3 m. This parameter had been defined as a minimum value for safety operation by Czech technical standard (see CSN 31 1500, second edition) [25]. Three types of tram vehicles operated on this testing path, namely Vario LFR (a), Stadler Tango NF2(b) and Inekon 2001 TRIO (c) with 2 or 3 axles (4 or 6 wheels). Figure 5 shows real photos of these tram vehicles taken during the measurement period. Measurements were performed on the rail closer to the sensor, as depicted in Figure 4.

The interferometric sensor is placed on the left bottom (red labelled), 3 m from the rail foot (Figure 4). During all measurements, the individual types of tram and the number of wagons were calculated manually (reference), detailed information about the tram axles and wheels is freely accessible in [26]. The detection was considered successful when the sensor detected not only axles but also the number of wheels for each axle.

Figure 6 (a) shows an example of the time recording of passage for tram vehicle with 2 axles (4 wheels). The X-axis represents the time (s), Y-axis reflects the voltage amplitude (V), and it is directly proportional to light intensity caught by the used photodetector. Individual maxima (red labelled) indicate the axles (number of wheels) of the individual tram vehicles. Figure 6 (b) shows the measured frequency spectrum for individual tram vehicles, and its values correspond with the standard published in [27].

This standard defines the frequency range which can be generated by rail vehicles during their passage, as presented in paper [2]. In this case, X-axis represents the frequency (Hz) and Y-axis represents the voltage amplitude (V).

Figure 7 (a) shows an example of the time recording of passage for tram vehicle with 3 axles (6 wheels). The graph shows 3 groups with 6 individual maxima, each one corresponding to 1 wheel. Figure 7(b) depicts the measured frequency spectrum that corresponds with the values presented by the technical standard in [27].

It is obvious that our proposed sensor is able to detect individual axles and wheels with sufficient accuracy (please see summary in Table 2). The individual maxima (Figure 6 and Figure 7) correspond with the number of wheels. Table 2 summarizes the measurement results, measurement dates, as well as the weather conditions (we defined 4 basic types: sunny, cloudy, windy and rainy), further Number of passes (-) of trams, Wrong detection (-) and Detection success (%).

5 Discussion

While performing the measurements, we identified certain limitations. If tram vehicles meet at the measurement point (passing in the opposite direction), sensor does not have to distinguish individual axles successfully. This case happened only two times during the whole measurement period. In case of bad weather (heavy rain and windy), from the obtained signal, we were not able to recognize the individual axles. This also happened two times. Since the testing was done from April to September, we did not have a chance to test the sensor in winter conditions. We plan to continue with the data collection and prepare our system for adaptation in low temperatures and snow cover, which is going to be the primary task of our next study.

Given the fact that the proposed sensor consists of conventional fibers and FC/APC connectors, it could be

Table 2 *Summary of realized measurements*

Day / Month / Weather	Number of Passes (-)	Wrong Detection (-)	Detection Success (%)
1 / April / cloudy	38	0	100
2 / April / windy	41	0	100
3 / April / sunny	36	0	100
4 / May / rainy	44	2	99.12
5 / May / rainy	28	0	100
6 / May / sunny	17	0	100
7 / June / cloudy	29	0	100
8 / June / sunny	47	1	99.53
9 / July / sunny	53	0	100
10 / July / sunny	25	0	100
11 / July / sunny	36	0	100
12/August / cloudy	39	0	100
13 / August / sunny	43	0	100
14 / August / sunny	52	0	100
15 / September / rainy	56	0	100
16 / September / windy	31	1	99.69
17 / September / windy	27	0	100
Summary	642	4	99.37

connected to dark telecommunication fibers along the rail infrastructures. Measurements evaluation could be performed remotely (laboratory tested distance was up to 2 km), the main limitation here would primarily be related to the fiber type for interconnection between the sensor and evaluation unit, and also to the power of radiation source. The improved sensitivity to axle detection is significantly higher in comparison to the sensitivity presented in our previously published paper [2]. Innovative construction led to an increase in the frequency range, but also to weight, size and price reduction. Another important advantage of our solution is that the system does not require installation into the rails (in comparison to other conventional systems [5-7]). During the measurements, we had to observe the minimum distance of 3 m from the rail according to [25], but a higher sensitivity of the sensor can be expected if the sensor will be placed closer to the rail.

Our future plans include the detection of flat wheels because this is the current and discussed topic. Flat wheels can cause travellers' discomfort and damages to tram axles [28-29].

The sensor has been tested in tram traffic; we assume that it will be possible to use it in classic train traffic (it is problematic to obtain a permission for measurement). Also, this will be one of the primary tasks of our next study.

6 Conclusion

This paper deals with the alternative usage of fiber-optic technology in the public transportation sector, especially for the detection of axles and individual wheels of tram vehicles. We have proposed an interferometric phase-based Mach-Zehnder sensor which works on the principle of light interferences within optical fibers. We performed real deployment based on the measuring of 642 tram vehicles during the long-time period of April - September 2019 in different climatic conditions. The successful detection rate of axles and wheels reached above 99 %. Characteristic features of the proposed solution are low price, practical construction, as well as the possibility to evaluate remotely information thanks to the connection via dark optical fibers located alongside the rail infrastructures.

Acknowledgement

The paper has been funded with support of conceptual development of science, research and innovation in 2019, assigned to VSB-Technical University of Ostrava, The Ministry of Education, Youth and Sports in the Czech Republic. This article was supported by the Ministry of Education of the Czech Republic (Projects Nos. SP2020/165, SP2020/38 and SP2020/166). This article was also supported by the European Regional Development Fund in the Research Centre of Advanced Mechatronic Systems project, project number CZ.02.1.01/0.0/0.0/16_019/0000867 within the Operational Programme Research, Development and Education. This work was supported by the European Regional Development Fund in Research Platform focused on Industry 4.0 and Robotics in Ostrava project CZ.02.1.01/0.0/0.0/17_049/0008425 within the Operational Programme Research, Development and Education. This research has been partially supported

by the Ministry of Education, Youth and Sports of the Czech Republic through the grant project No. CZ.1.07/2.3.00/20.0217 within the frame of the operation program Education for Competitiveness financed by the European Structural Funds and from the state budget of the Czech Republic and by institutional research of Faculty of Operation and Economics of Transport and Communications - University of Zilina no. 11/PEDAS/2019.

References

[1] NEDOMA, J., FAJKUS, M., MARTINEK, R., WITAS, K., MEC, P., JARGUS, J., HEJDUK, S., ZAVODNY, P., VASINEK, V. Sensor system based on the Mach-Zehnder interferometer for the rail transport. In: SPIE - The International Society for Optical Engineering : proceedings [online]. 10654, art. no. 106541G, 2018. Available from: https://doi.org/10.1117/12.2304582

[2] NEDOMA, J., STOLARIK, M., FAJKUS, M., PINKA, M., HEJDUK, S. Use of fiber-optic sensors for the detection of the rail vehicles and monitoring of the rock mass dynamic response due to railway rolling stock for the civil engineering needs. *Applied Sciences* [online]. 2019, **9**(1), p. 134. eISSN 2076-3417. Available from: https://doi.org/10.3390/app9010134

[3] NEDOMA, J., KEPAK, S., CUBIK, J., FRNDA, J., DURICA, M., FAJKUS, M., MARTINEK, R. Vital sign monitoring: a practical solution by a MR compatible phonocardiography interferometric probe. *Journal of Optoelectronics and Advanced Materials* [online]. 2019, **21**(11-12), p. 656-662. ISSN 1454-4164, eISSN 1841-7132. Available from: https://joam.inoe.ro/articles/vital-sign-monitoring-a-practical-solution-by-a-mr-compatible-phonocardiography-interferometric-probe/

[4] NEDOMA, J., FAJKUS, M., KAHANKOVA, R., MARTINEK, R., DVORSKY, M., VANUS, J., VASINEK, V., CVEJN, D. Fiber-optic interferometric sensor for monitoring automobile and rail traffic. *Turkish Journal of Electrical Engineering and Computer Sciences* [online]. 2018, **26**(6), p. 2986-2995. eISSN 1303-6203. Available from: https://doi.org/10.3906/elk-1712-166

[5] MARSHALL, J. *A biographical dictionary of railway engineers*. North Pomfret, 1978. ISBN 0715374893.

[6] CHEN, J., ROBERTS, C., WESTON.P. Fault detection and diagnosis for railway track circuits using neuro-fuzzy systems. *Control Engineering Practice* [online]. 2008, **16**(5), p. 585-596. ISSN 0967-0661. Available from: https://doi.org/10.1016/j.conengprac.2007.06.007

[7] RSSB. Dealing with a train accident or train evacuation [online]. [Viewed 2019-11-19]. Available from: https://www.jonroma.net/media/rail/opdocs/world/uk/GERM8000-master-module%20Iss%206.pdf

[8] NIELSEN, J. C. O., JOHANSSON, A. Out-of-round railway wheels-a literature survey. *Proceedings of the Institution of Mechanical Engineers. Part F: Journal of Rail and Rapid Transit* [online]. 2000, **214**(2), p. 79-91. ISSN 0954-4097, eISSN 2041-3017. Available from: https://doi.org/10.1243/0954409001531351

[9] JARDINE, A. K. S., LIN, D., BANJEVIC. D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing* [online]. 2006, **20**(7), p. 1483-1510. ISSN 0888-3270. Available from: https://doi.org/10.1016/j.ymssp.2005.09.012

[10] NGIGI, R. W., PISLARU, C., BALL, A., GU, F. Modern techniques for condition monitoring of railway vehicle dynamics. *Journal of Physics: Conference Series*. 2012, **364**, con. 1. ISSN 1742-6588, eISSN 1742-6596. Available from: https://doi.org/ 10.1088/1742-6596/364/1/012016

[11] SMUTNY, J., NOHAL, V., VUKUSICOVA, D., SEELMANN, H. Vibration analysis by the Wigner-Ville transformation method. *Communications - Scientific Letters of the University of Zilina* [online]. 2018, **20**(4), p. 24-28. ISSN 1335-4205, eISSN 2585-7878. Available from :/komunikacie.uniza.sk/index.php/communications/article/view/635

[12] BIAGI, M., CARNEVALI, L., PAOLIERI, M., VICARIO. E. Performability evaluation of the ERTMS/ETCS - level 3. *Transportation Research Part C: Emerging Technologies* [online]. 2017, **82**, p. 314-336. ISSN 0968-090X. Available from: https://doi.org/10.1016/j.trc.2017.07.002

[13] BABCZYNSKI, T., MAGOTT, J. Dependability and safety analysis of ETCS communication for ERTMS level 3 using performance statecharts and analytic estimation. *Advances in Intelligent Systems and Computing* [online]. 2014, **286**, p. 37- 46. ISBN 978-3-319-07012-4, eISBN 978-3-319-07013-1. Available from: https://doi.org/10.1007/978-3-319-07013-1_4

[14] KEPAK, S., CUBIK, J., ZAVODNY, P., SISKA P., DAVIDSON, A., GLESK, I., VASINEK, V. Fibre optic track vibration monitoring system. *Optical and Quantum Electronics* [online]. 2016, **48**, p. 354. ISSN 0306-8919, eISSN 1572-817X. Available from: https://doi.org/10.1007/s11082-016-0616-9

[15] KEPAK, S., CUBIK, J., ZAVODNY, P., HEJDUK, S., NEDOMA, J., DAVIDSON, A., VASINEK, V. Fibre optic portable rail vehicle detector. In: SPIE - The International Society for Optical Engineering : proceedings. 10142. 2016, p. 402-408. Available from: https://doi.org/10.1117/12.2257061

[16] KEPAK, S., CUBIK, J., NEDOMA, J., HRUBY, D., HEJDUK, S., ZAVODNY, P., FAJKUS, M., VASINEK, V. Compact fiber optic trackside sensor for rail vehicle detection and analysis. *IFAC-PapersOnLine* [online]. 2018, **51**(6), p. 220-224. ISSN 2405-8963. Available from: https://doi.org/10.1016/j.ifacol.2018.07.157

[17] NEDOMA, J., FAJKUS, M., ZABKA, S., MARTINEK, R. Fiber optic tram vehicle detector. *Optoelectronics and Advanced Materials - Rapid Communications* [online]. 2019, **13**(1-2), p. 37-43. ISSN 1842-6573, eISSN 2065-3824. Available from: https://oam-rc.inoe.ro/articles/fiber-optic-tram-vehicle-detector/

[18] PAPP, B., DONNO, D., MARTIN, J. E., HARTOG, A. H. A study of the geophysical response of distributed fibre optic acoustic sensors through laboratory-scale experiments. *Geophysical Prospecting* [online]. 2017, **65**, p. 1186-1204. eISSN 1365-2478. Available from: https://doi.org/10.1111/1365-2478.12471

[19] NEDOMA, J., FAJKUS, M., MARTINEK, R., MEC, P., NOVAK, M., BEDNAREK, L., VASINEK, V. Interferometer for securing entrance areas of buildings. In: SPIE - The International Society for Optical Engineering : proceedings. 10440. 2017, p. 176-181. Available from: https://doi.org/10.1117/12.2277108

[20] FENG, L. L., WANG, Y. T., RUAN, C., TAO, S. Road vehicle information collection system based on distributed fiber optics sensor. *Advanced Materials Research* [online]. 2014, **1030-1032**, p. 2105-2109. ISSN 1662-8985. Available from: https://doi.org/10.4028/www.scientific.net/AMR.1030-1032.2105

[21] GOODWIN, E. P., WYANT J. C. *Field guide to interferometric optical testing*. Field Guide Series, SPIE, 2006. ISBN 978-081-9465-108.

[22] KOLAR, V., HRBAC, R., MLCAK, T., STYSKALA, V. Regulated electric drainage and its interference with track circuits. *Advances in Electrical and Electronic Engineering* [online]. 2018, **16**(2), p. 205-210. ISSN 1336-1376, eISSN 1804-3119. Available from: https:/doi.org/ 10.15598/aeee.v16i2.2478

[23] BORN, M., WOLF, E. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. 7th ed. New York, NY, USA: Cambridge University Press, 1999. ISBN 05-216-4222-1.

[24] ZENG, M., ZHAO, H., WU, D., CHEN, H., CAI, J. A vibration-based traffic monitoring system using distributed optical sensing technology. *Journal of Testing and Evaluation* [online]. 2020, **48**. ISSN 0090-3973. Available from: https://doi.org/10.1520/JTE20190184

[25] CSN 34 1500 ED.2 (341500) - Czech technical standard [online] [Viewed 2019-11-19]. Available from: http://www.technicke-normy-csn.cz/341500-csn-34-1500-ed-2_4_84547.html

[26] DPO - city public transport company from Ostrava [online] [Viewed 2019-11-19]. Available from: https://www.dpo.cz/o-spolecnosti/vozy/tramvaje/3448-vario-lfr.html

[27] ISO standard 4866:2010. Mechanical vibration and shock, vibration of fixed structures - guidelines for the measurement of vibrations and evaluation of their effects on structures [online] [Viewed 2019-11-19]. Avaliable from: https://www.iso.org/standard/38967.html

[28] FANG, X., LIN, S., YANG, Z., LIN, F., SUN, H., HU, L. Adhesion control strategy based on the wheel-rail adhesion state observation for high-speed trains. *Electronics* [online]. 2018, **7**(5), p. 70. eISSN 2079-9292. Available from: https://doi.org/10.3390/electronics7050070

[29] NOWAKOWSKI, T., KOMORSKI, P., SZYMANSKI, G. M., TOMASZEWSKI, F. Wheel-flat detection on trams using envelope analysis with Hilbert transform. *Latin American Journal of Solids and Structures* [online]. 2019, **16**(1). ISSN 1679-7817, eISSN 1679-7825. Available from: https://dx.doi.org/10.1590/1679-78255010

Esra'a Alkafaween - Ahmad B. A. Hassanat

IMPROVING TSP SOLUTIONS USING GA WITH A NEW HYBRID MUTATION BASED ON KNOWLEDGE AND RANDOMNESS

Genetic algorithm (GA) is an efficient tool for solving optimization problems by evolving solutions, as it mimics the Darwinian theory of natural evolution. The mutation operator is one of the key success factors in GA, as it is considered the exploration operator of GA.

Various mutation operators exist to solve hard combinatorial problems such as the TSP. In this paper, we propose a hybrid mutation operator called "IRGIBNNM", this mutation is a combination of two existing mutations; a knowledge-based mutation, and a random-based mutation. We also improve the existing "select best mutation" strategy using the proposed mutation.

We conducted several experiments on twelve benchmark Symmetric traveling salesman problem (STSP) instances. The results of our experiments show the efficiency of the proposed mutation, particularly when we use it with some other mutations.

Keyword: *knowledge-based mutation, inversion mutation, slide mutation, RGIBNNM, SBM*

1 Introduction

1.1 Travelling salesman problem (TSP)

TSP is considered as one of the combinatorial optimization problems [1], that is easy to describe but difficult to solve, and it is classified among the problems that are not solved in polynomial time; i.e. it belongs to the NP-hard problem [2].

A solution of TSP aims at finding the shortest path (tour) through a set of nodes (starting from a node N and finishing at the same node) so that each node is visited only once [3].

The classic problem of a traveling salesman is an active and attractive field of research because of its simple formulation [2], and it was proved to be NP-complete problem, since no one found any effective way to solve an NP problem of a large size, in addition, many problems in the world can be modeled by TSP [4].

The TSP is classified into:

1. Symmetric traveling salesman problem (STSP): The cost (distance) between any two cities in both directions is the same (undirected graph), i.e. the distance from *city1* to *city2* is the same as the distance from *city2* to *city1*. There are (N-1)! /2 possible solutions for N cities.
2. Asymmetric travelling salesman problem (ATSP): The cost between any two cities in both directions is not the same. There are (N-1)! possible solutions for N cities [5].

TSPs are used in various applications, including: job sequencing, computer wiring, crystallography, wallpaper cutting, dartboard design, hole punching, overhauling gas turbine, etc. [6].

Over the years various techniques have been suggested to solve the TSP, such as genetic algorithm (GA) [7-8], hill climbing [9], nearest neighbor and minimum spanning tree algorithms [10], simulated annealing [11], ant colony [9], tabu search [12], particle swarm [13], elastic nets [14], neural networks [15], etc. Genetic algorithms are one of the algorithms that extensively applied to solve the TSP [16].

1.2 Genetic algorithm (GA)

GA is an optimization algorithm [17] that is classified as global search heuristic; it is one of the categories that form the family of the evolutionary algorithms, which mimics the principles of natural evolution [18]. GA is a population-based search algorithm, as in each generation, a new population is generated by repeating three basic operations on the population, namely, selection, crossover, and mutation [19]. GA has been used extensively in many fields, such as computer networks [20], speech recognition [21], image processing [22], software engineering [23], etc.

A simple GA algorithm is described as follows [16]:

Step1: Create a random population of potential solutions [24] consisting of n individuals (initial populations): The first phase of any GA is initial population seeding. It generates a first population randomly or by heuristic initialization as

Esra'a Alkafaween[1,*], Ahmad B. A. Hassanat[1,2,3]
[1]IT Department, Mutah University, Karak, Jordan
[2]Computer Science Department, Community College, University of Tabuk, Saudi Arabia
[3]Industrial Innovation and Robotics Center, University of Tabuk, Saudi Arabia
*E-mail of corresponding author: esrakafaween86@gmail.com

input for the GA [25], such as: random, nearest neighbor, k-means clustering and initialization mechanism based on regression techniques [26].

Step2: Evaluate the fitness value f(x) of each individual, x, in the population.

Step3: Repeat the following three steps to create a new population until completion of the new population.

Step4: Selection phase: This is the process of choosing the best parents of the current generation in the community for mating, such as: roulette wheel, elitism and tournament.

Step5: Apply crossover with a certain ratio to create offspring: This process takes two parents (chromosomes) to create a new offspring by switching segments of the parent genes. It is more likely that the new offspring (children) will contain good parts of their parents, and consequently perform better as compared to their ancestors. There are many types of crossover, such as: modified crossover, uniform crossover and arithmetic crossover.

Step6: Apply mutation with a certain ratio: This is where there is a change or a switch between specific genes within a single chromosome to create chromosomes that provide new solutions for the next generation, with the aim of obtaining the best possible solutions, and thus introduce a certain level of diversity to the population, and as a result this also does not fall into the local optimum [27]. There are many types of mutation such as: exchange mutation, scramble mutation, insert mutation and inversion mutation.

Step7: The previous operations are repeated until the completion criterion is met.

The performance of the GA is affected by several key factors, such as the population size, the selection's strategy, the mutation operator used, the crossover operator used and the coding scheme [28-30]. In this paper, we focus on the mutation operator.

Mutation operator plays an important role in the GA, where it helps to stimulate the GA to explore new areas in the search space [19]. It is an effective mechanism for preserving the diversity of individuals [28], where mutation provides variation in the population through random changes of individuals [29]. And therefore, overcoming the so-called premature convergence [31], also preventing the loss of genetic material [32].

In this paper, we propose a hybrid mutation operator called inversion RGIBNNM (IRGIBNNM) to provide an efficient solution to TSP, we use simple GA with mutations only; there is no other variable/parameter that controls the workflow of such a simple GA, as we want to examine the strength of the proposed mutation apart from the effect of other parameters; we compare the performance of this mutation with the performances of three existing mutations, and we used it with two other mutations to form a multi-mutations GA. The comparisons are made on symmetric TSP instances.

The organization of this paper is as follows. In Section 2 we present some of the related work. In section 3, we present the proposed mutation, the existing two mutations and the mutation strategy. In Section 4 we describe the experiments conducted, and discuss them. The conclusion is written in Section 5.

2 Background

Over the years, researchers have suggested several types of mutations to be used in various types of encoding, including: flip mutation, creep mutation and insert mutation [33], gaussian mutation, exchange mutation [34], displacement mutation [35], uniform mutation [1], inversion mutation [36] and some other types.

Louis and Tang proposed a new mutation called greedy-swap mutation, where two cities are chosen randomly from the same chromosome, and switching between them if the length of the new tour obtained is "better" (shorter) than the previous ones [37].

Potvin [2] and Larranaga et al. [8] presented a review of representing the TSP, explaining the advantages and disadvantages of different mutation operators. Soni and Kumar studied many types of mutations that solve the problem of travelling salesmen, including interchanging mutation, reversing mutation and scramble mutation [1]. Otman and Jaafar used reverse sequence mutation (RSM) and several types of crossover to solve the TSP [32]. Korejo et al, introduced a directed mutation (DM), this method used the statistical information provided by the current population to explore the promising areas in the search space [19].

Having such a large number of mutations, the problem becomes which mutation to use? As the problem lies in choosing the appropriate mutation. To answer this question, several researchers have developed new types of GAs that use more than one mutation at the same time [38-41].

Katayama et al. presented a promising GA for TSP, called a hybrid mutation genetic algorithm (HMGA), which employed a local search algorithm called stochastic hill climbing (SHC), in order to avoid falling into the local optima [42].

Hong et al. proposed a new GA, called dynamic genetic algorithm (DGA) in order to choose the appropriate mutation and crossover operators and their ratios automatically, this algorithm use more than one mutation at the same time, such as: uniform crossover, (0,1) change, inversion, bit-change and swapping [43].

Hassanat et al. proposed 10 types of knowledge-based mutations; the most promising one is called "random gene inserted beside nearest neighbor mutation" (RGIBNNM). In addition, they proposed two selection strategies for the mutation operators called: "select the best mutation" (SBM) and "select any mutation" (SAM). They applied all mutations and strategies on several TSP instances [38].

Regardless the extensive research in this domain, there is no one mutation ideally suited for all TSP instances. Since no one method found in the literature that guarantees an optimal solution for any TSP instance. Therefore, there is still room for improvement in this domain.

3 The proposed method

In this section, we explain some of the existing mutation operators that are proposed for the permutation coded GA; these include slide mutation [44], inversion mutation [36] and RGIBNNM [38]. Moreover, we explore the strategy of choosing the best mutation; the SBM [38]. We also present the proposed hybrid mutation, which is nothing but a combination of the inversion mutation and the RGIBNNM, we call it IRGIBNNM.

3.1 Slide mutation

This mutation chooses two genes randomly, and then conveys the first to follow the second, and then shift the rest of the city, as depicted by *Example 1*.
Example 1: Consider the following TSP *tour C*:

*C=(**5 3 10 2 1 8 9 7 4 6**)*.

If the third *gene 10* and the eighth *gene 7* are randomly selected, then the sub tour is:

*(**2 1 8 9 7**)*.

The mutated tour will be:
*(Offspring)=(**5 3 10 1 8 9 7 2 4 6**)*.

3.2 Inversion mutation

This mutation chooses two random genes, and then reverses the subset between them, as depicted by *Example 2*.
Example 2: Consider the following *tour C*:

*C=(**5 3 10 2 1 8 9 7 4 6**)*.

If the third and eighth positions are randomly selected, then the sub tour is:

*(**2 1 8 9 7**)*,

and then reversed to be:

*(**7 9 8 1 2**)*.

The mutated tour will be:

*(Offspring)=(**5 3 10 7 9 8 1 2 4 6**)*.

3.3 RGIBNNM mutation

This mutation is a knowledge-based operator designed especially for the TSP problem. However, it can be customized to fit some other problems. This operator uses the idea of the nearest neighbor cities, where this mutation selects a random gene (city), and finds its nearest city (Ncity), then swap the random city with one of the neighbors of the nearest city.
Example 3: suppose that the chromosome chosen for mutation is:

$B \rightarrow E \rightarrow C \rightarrow A \rightarrow D$,

as depicted in Figure 1. By applying RGIBNN:
Step 1: Suppose that the city which has been selected at random is *E*.
Step 2: Find the nearest city to the random city, which is *D* according to the distance table. This city is called Ncity.
Step 3: Now *E* is moved prior to *D*, and (*B* and *C*) are shifted to get a new chromosome

$B \rightarrow C \rightarrow A \rightarrow E \rightarrow D$.

3.4 The proposed IRGIBNNM

We propose a hybrid mutation called: IRGIBNNM.

In this mutation we combine two mutation operators, the inversion mutation and RGIBNNM.

The IRGIBNNM initially applies the inversion mutation on an individual, and then the RGIBNNM is applied to the resulting individual. Thus, the new offspring benefit from both mutations' characteristics, attempting to enhance the performance of both mutations, by increasing diversity in the search space, and therefore to provide better results. The IRGIBNNM is depicted by *Example 4*.
Example 4: Consider the following *tour C*:

C=(5 3 10 9 8 1 2 7 4) with cost=19,

as depicted in Figure 2. To apply IRGIBNNM:

Select two random genes, e.g. the third and seventh genes.

A=inversion mutation C.

The resulting offspring:

A=(5 3 10 2 1 8 9 7 4) with cost=18.2 (see Figure 3).

Apply RGIBNNM *A* as follows:

- Select a random gene from *A*, e.g. the random gene is the eighth gene, i.e. the *random city* is *7*.
- Find the nearest city to the *random city 7*, which is *city 3* in this case.
- Get a random city around *city 3* in the range (±5) (e.g we determine the range of the number of cities close to the chosen city, so if we choose 3, we mean only the nearest 3 cities and then we choose one of them); e.g. *city 9*.

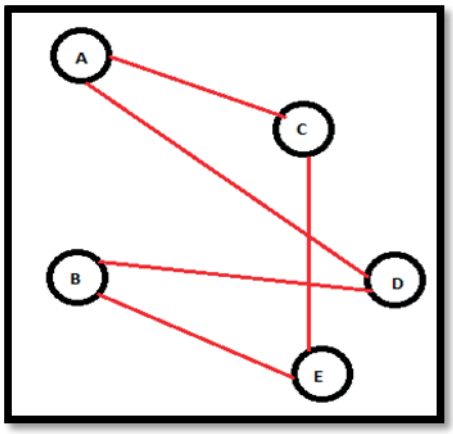


Figure 1 *Example of RGIBNNM*

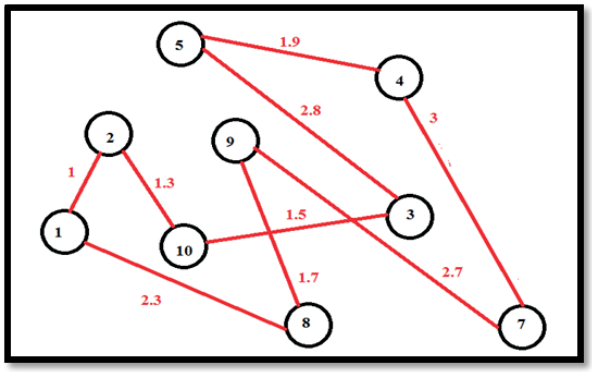


Figure 3 *Example of applying inversion mutation on C to get offspring A with cost=18.2*

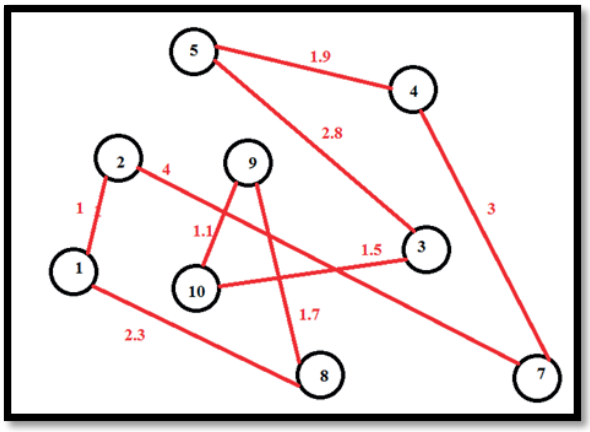


Figure 2 *Example of particular tour C with cost=19*

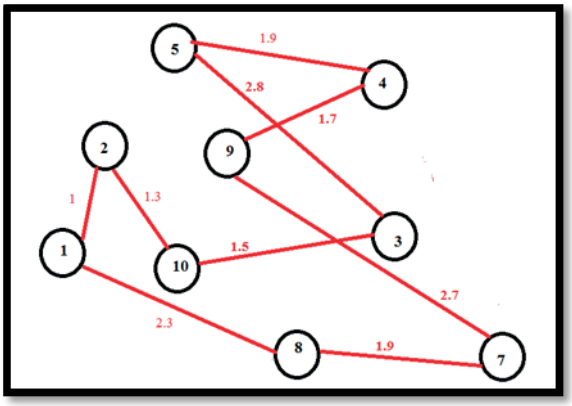


Figure 4 *Example of applying IRGIBNNM on A to get offspring with cost=17.1*

Apply the exchange mutation on chromosome *A* by swapping the *cities 7* and *9*, as shown in Figure 4. The final output offspring becomes:

Offspring=(5 3 10 2 1 8 7 9 4) with cost=(17.1).

3.5 Select the best mutation (SBM)

This strategy applies various mutation operators simultaneously to the same individual, and from each mutation produces one offspring; the "best" offspring that does not already exist in the population is added to the population [38]. For TSP the "best" solution, is the one with the minimum distance.

In this paper, we used three mutations only (slide mutation, inversion mutation, and the proposed IRGIBNNM), instead using several other mutations as proposed by [38].

A larger example is shown in Figures 5 and 6, which depicts the implementation of four mutations, in addition to the SBM strategy for 80 random cities.

A real data example is shown in Figure 7, which shows the implementation of the four mutations and SBM on a particular route of the TSP (eil51) taken from TSBLIB [45]. A closer look at Figures 5-7 shows significant improvements on the initial tour, particularly when using IRGIBNNM or SBM strategy.

4 Experimental setup and result

To evaluate the performance of the proposed mutation (IRGIBNNM) and the new SBM strategy, we conducted several experiments on twelve TSP instances, each having the known solution (optimal). Those instances were taken from the TSPLIB [45], and they include: eil51, a280, bier127, berlin52, KroA100, KroA200, ch150, rat195, st70, pr125, pr226 and lin318. Same experiments, under the same circumstances were conducted to examine the convergence to a minimum value of each operator separately, including the other mutations (slide, inversion)

We have implemented the new SBM strategy the same as proposed by [38], but using three mutations only (slide, inversion, and the proposed IRGIBNNM), instead of several other mutations, considering the best offspring to be added to the population. To prevent duplication of chromosomes, if the new offspring is found in the population, we consider the lower quality offspring, and if all of the three offspring found in the population the operation (on that chromosome) is canceled.

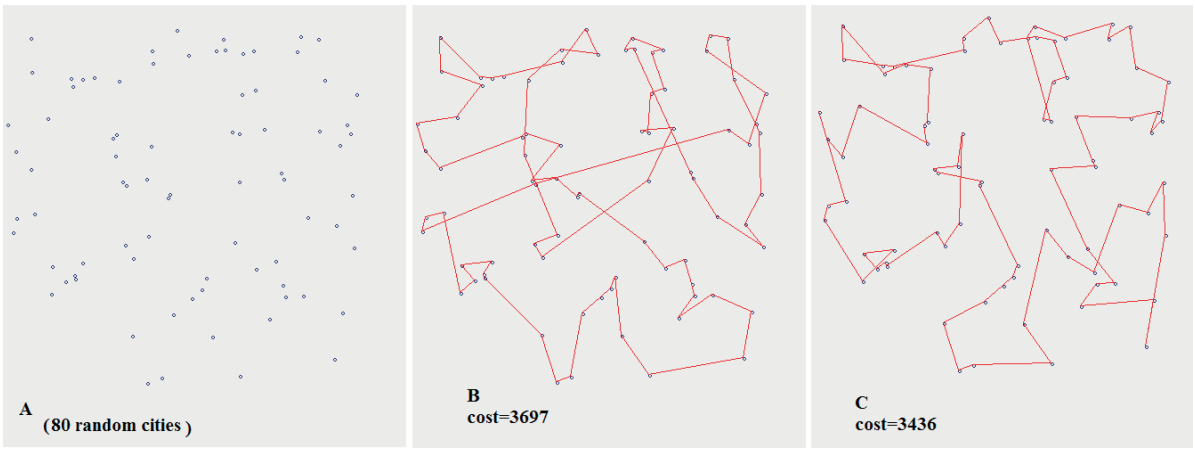


Figure 5 *Implementation of the three mutations for 80 random cities: A) 80 random cities, B) slide mutation, C) inversion mutation*

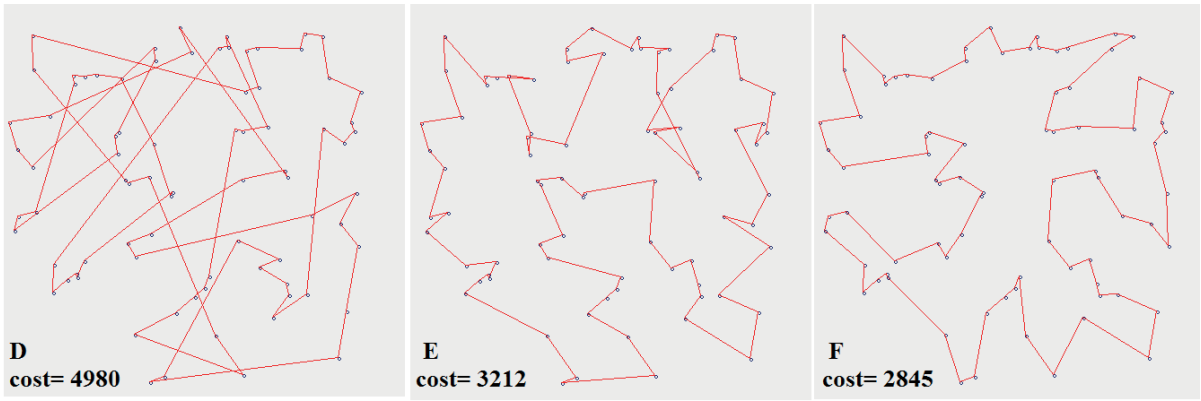


Figure 6 *Implementation of the three mutations on the same random cities of Figure 5: D) RGIBNNM, E) IRGIBNNM, F) SBM strategy*

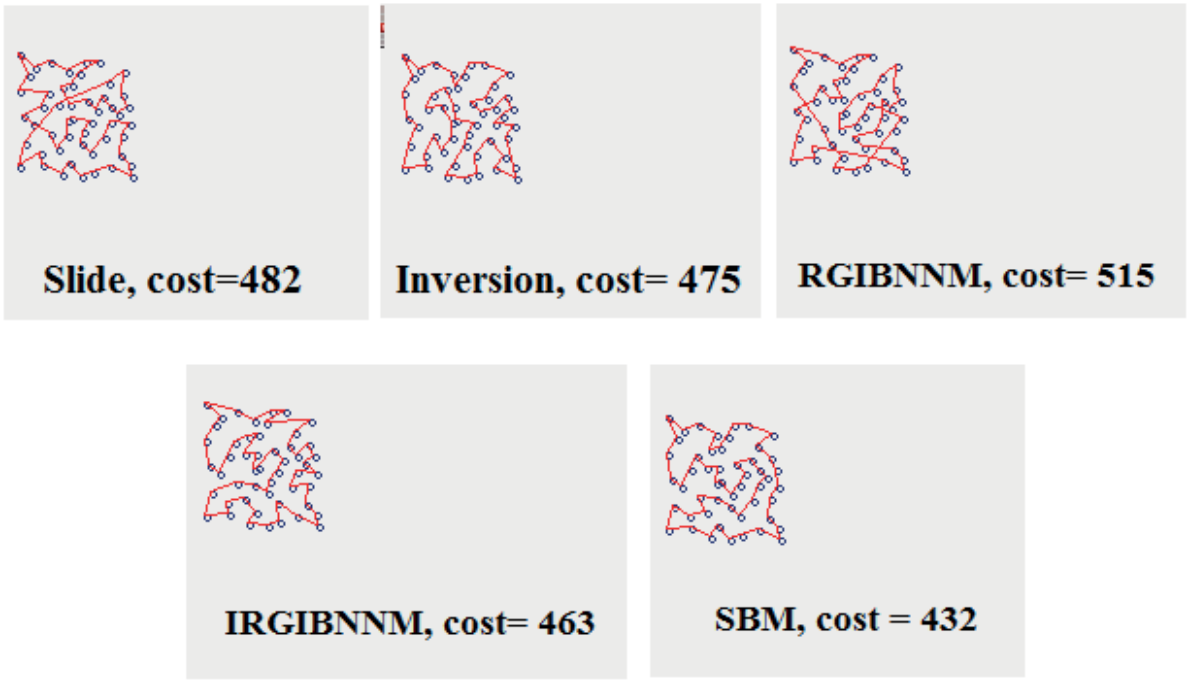


Figure 7 *The effect of applying the mutations on Eil51*

Table 1 *Results of TSP instances obtained by inversion and slide mutations after 2000 generations*

mutation type		inversion mutation			slide mutation		
instances	optimal	best fitness	worst fitness	average fitness	best fitness	worst fitness	average fitness
eil51	426	440	453	446.1	469	583	503.9
a280	2579	9811	10119	9974.2	9532	10522	9917.4
bier127	118282	167565	183857	172867.4	177720	193326	185276.6
kroA100	21282	30310	33413	31925.3	31800	36279	34120.6
berlin52	7542	7769	8515	8038.1	8498	10154	9334.6
kroA200	29368	80906	84555	81958	74586	90348	83529.8
pr125	73682	151643	168468	161445.4	170304	218119	192498
lin318	42029	185852	192611	188931.6	176935	185899	181978.5
pr226	80369	331572	353613	342094.3	345027	377088	360239.9
ch150	6528	13006	13670	13425.1	13129	15221	13778.1
st70	675	758	815	783	787	1004	882.4
rat195	2323	5548	5955	5836.5	5420	6169	5774.8

Table 2 *Results of TSP instances obtained by IRGIBNNM and RGIBNNM mutations after 2000 generations*

mutation type		IRGIBNNM			RGIBNNM		
Instances	optimal	best fitness	worst fitness	average fitness	best fitness	worst fitness	average fitness
eil51	426	448	463	455.3	518	603	575.5
a280	2579	7313	7846	7507.9	6543	8307	7526.5
bier127	118282	156903	169657	164072.9	205820	254541	234760.2
kroA100	21282	25941	29218	27418.7	43474	53903	48077.1
berlin52	7542	8098	8705	8354.2	9639	11105	10296.1
kroA200	29368	59802	63911	62136.9	88409	109892	97125.7
pr125	73682	111055	127783	121013.5	213526	270814	235064.1
lin318	42029	132899	145109	136569.5	159856	178241	173127.6
pr226	80369	191049	234720	216699	288421	380900	322855.1
ch150	6528	10517	11396	11111.9	15071	18435	16774.2
st70	675	733	772	753.4	1058	1296	1222.1
rat195	2323	4321	4758	4554.2	6203	7492	7081.5

In all experiments, our GA used the reinsertion method, which is an expansion sampling [46], where this method means, only the excellent half (from the new individuals and old generation) is selected as a population for the next generation. in other words, when creating a new generation, the old generation competes with the new individuals.

We repeated each experiment 10 times, the GA parameters used are as follows: the Population size=100, the probability of crossover=0% and all previous mutations occur 100%. The initial population is random based population seeding and selection strategy in all algorithms is random. The termination criterion is based on a fixed number of generations reached. For all of our experiments the maximum number of generations=2000.

The operators are coded in VC++, and the computer specifications: 1.66 GHz processor PC with 2 GB of RAM.

The results of the mutations evaluated on 12 instances from the TSP are summarized in Tables 1 and 2.

As can be seen from Tables 1 and 2, the best performance was recorded by the IRGIBNNM for 10 instances, followed by the inversion mutation, which also shows a better performance than both of the slide mutation and the RGIBNNM. The significant performance of the IRGIBNNM is justified by the exploiting of two mutations applied after each other on the same individual. The first provides random solutions and the second provides solutions based on some knowledge of the nearest neighbor. Randomness provided by the inversion mutation, and knowledge provided by

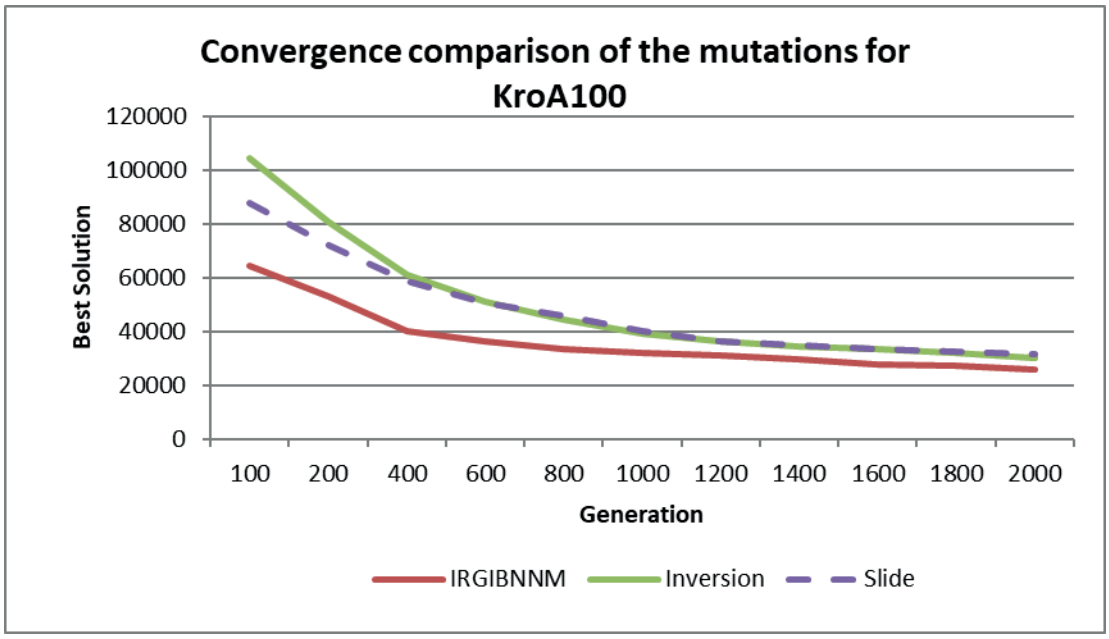


Figure 8 *Mutation's convergence to the minimum solution, kroA100*

Table 3 *Results of TSP instances obtained by SBM after 2000 generations*

Instances	Optimal	Best Fitness	Worst fitness	Average fitness
eil51	426	428	439	432.7
a280	2579	2898	3089	2974.9
bier127	118282	121644	128562	124492.5
kroA100	21282	21344	22788	21957.1
berlin52	7542	7544	8423	7890.7
kroA200	29368	30344	32103	31369
pr152	73682	74777	86240	77022.9
lin318	42029	47006	50033	48234.6
pr226	80369	82579	87006	84409.1
ch150	6528	6737	7044	6876
st70	675	677	723	694.8
rat195	2323	2404	2561	2481.9

the RGIBNNM allow for more diversity of good solutions, which leads to better results.

On result in Table 1 and 2 indicates that the SBM showed faster convergence to the minimum value followed by IRGIBNNM (at the level of mutation alone).

There are several performance factors used to investigate the significance of the importance of the different technique used to improve any GA, such as: computation time, error rate and average convergence [26] .

1- error rate (%): it could be defined as the percentage of difference in the fitness value of the solution with the known optimal solution for the problem. It can be given as:

*error rate (%)=((fitness-optimal fitness)/optimal fitness)*100%.* (1)

2- average convergence (%): it is defined as the average of the convergence rate of solutions.

*average convergence (%)=1-((average fitness - optimal fitness)/optimal fitness)*100.* (2)

Figure 8 shows the convergence to the minimum value recorded by each mutation. Again IRGIBNNM shows faster convergence to the minimum value than the other two mutations on KroA100. This faster convergence is due to the same randomness and knowledge afforded by the IRGIBNNM.

Using the same GA parameters, the second set of experiments is conducted to measure the performance of the new SBM, and to show the effective use of more than one mutation at the same time by the GAs. The results are shown in Table 3 and Figures 9-11.

As can be seen from the results in Table 3, it is important to select the appropriate mutation, in particular for the SBM strategy, and in general for the GA, because the choice of those methods affects the results of the GA

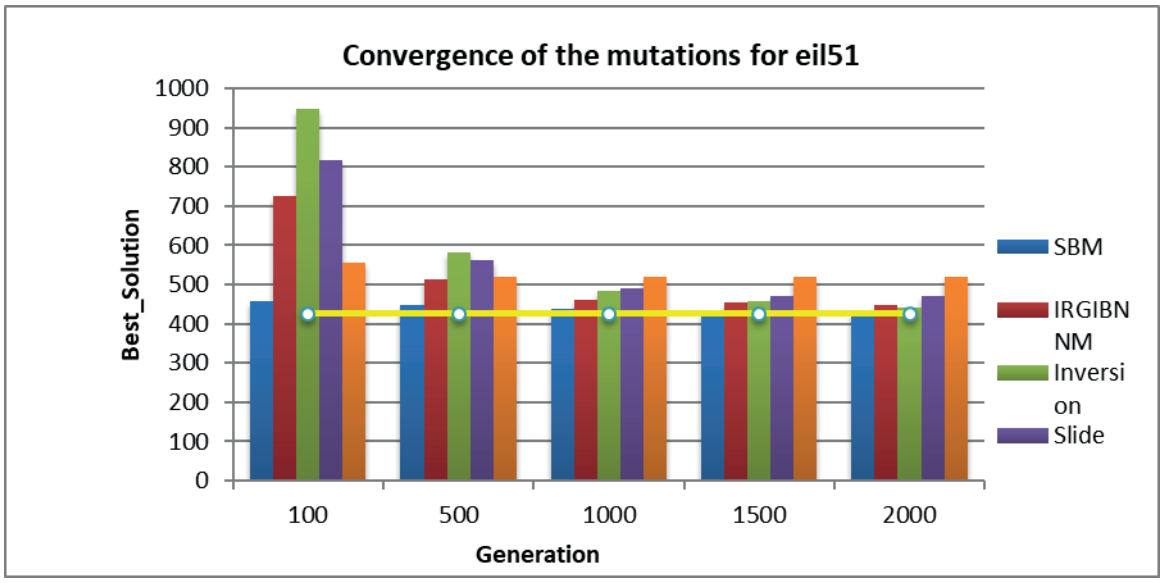


Figure 9 *Convergence Comparison for eil51*

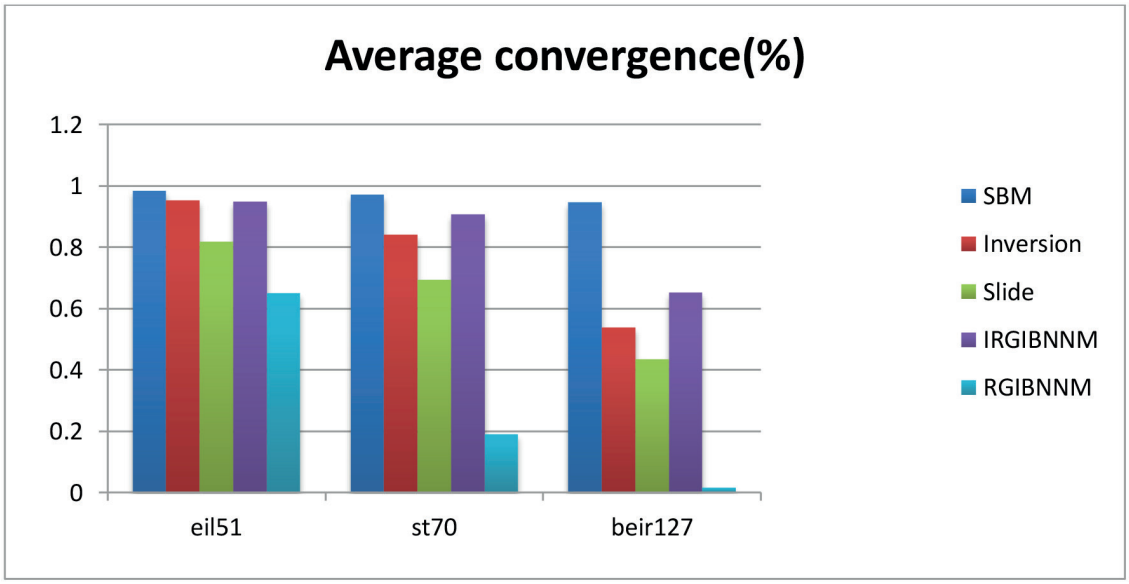


Figure 10 *Average Convergence of 4 mutations and SBM strategy for three instances from TSPLIB (eil51, st70, beir127)*

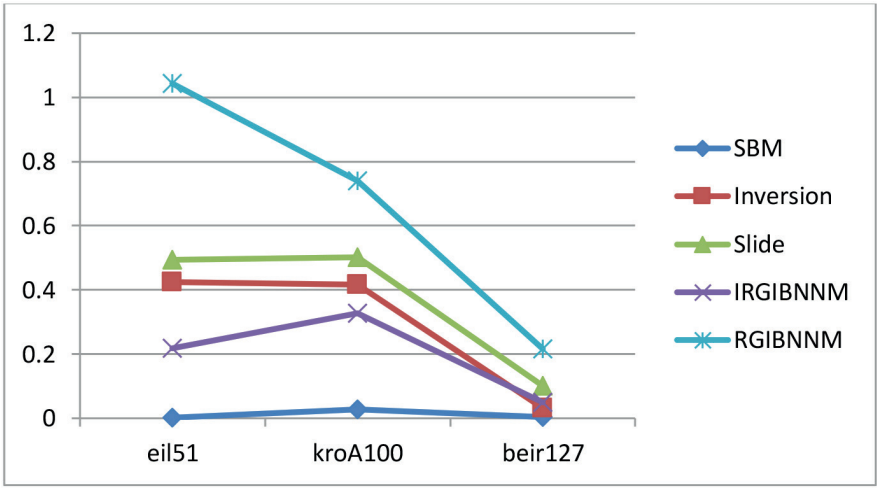


Figure 11 *Error rate of 4 mutations and SBM strategy for three instances from TSPLIB (eil51, st70, beir127)*

Table 4 *Results of new SBM compared to those of old SBM*

Instances	Optimal	New SBM 2000 Generations	Old SBM [38] 2000 Generations
eil51	426	428	443
a280	2579	2898	4824
bier127	118282	121644	175935
kroA100	21282	21344	33739
berlin52	7542	7544	8326
kroA200	29368	30344	51865
pr152	73682	74777	141114
lin318	42029	47006	94865
pr226	80369	82579	207167
ch150	6528	6737	8130
st70	675	677	809
rat195	2323	2404	3790

significantly. The best performance was recorded by the SBM algorithm, followed by the proposed IRGIBNNM, this is seen from Tables 2 and 3.

As seen in Figure 9, the SBM performs better than the other mutations, it is interesting to note that the solutions provided by the SBM are close to the optimal solutions for most of the TSP instances examined. A traditional genetic algorithm commonly uses one mutation operator. We propose using more than one mutation operation, anticipating that different operators will produce different patterns in the offspring and provide some sort of diversity in the population, so as to improve the overall performance of the genetic algorithm [47-48].

Results from Figures 10 and 11 shows the efficient use the three mutations together by the SBM, where the SBM achieves the highest convergence and less error with significant difference.

We justify the significant performance of the SBM as follows, intuitively, we have 2 options for the quality of a solution provided by any mutation, comparing to the average quality in the current population, a) a lower quality solution, and b) a higher quality solution; assuming that a solution with the same quality is considered as a higher quality solution. The new SBM uses 3 mutations, which are applied on the same chromosome, the probability to have them all fail, (i.e. to get lower quality outcomes (offspring) from all mutations used) is 1 out of 8 (low, low, low), while the probability to get a higher quality by any of them is 7 out of 8 possibilities, this high success rate justifies the significant performance of the SBM. Same justification applies to the good performance of the proposed IRGIBNNM, but with a lower success rate of 3 out of 4, since the IRGIBNNM uses only 2 mutations.

Despite the aim of this paper is not to find the optimal solution for TSP, the solutions of the hybrid mutation was close to optimal solutions in some cases, and none could achieve an optimal solution, this indicates the importance of the crossover operator and number of generation and operators ratios and other parameters along with mutation, due to the effective impact of their convergence to an optimal or near optimal solution.

The success of the new SBM is not attributed only to the use of multi mutations as described above, but also to the quality of the solutions provided by the mutations used by the SBM, and this pays attention to the proposed IRGIBNNM, which used by the SBM among the other two mutations. This conclusion is supported by comparing the results of the new SBM with the old SBM proposed by [38], see Table 4 and whith same genetic algorithm parameter, e.g (population size=100, the probability of crossover=0% and all previous mutations occur 100%. The initial population is random based population seeding and selection strategy in all algorithms is random. The termination criterion is based on a fixed number of generations reached. the maximum number of generations=2000.)

Comparing the proposed methods with the plethora of mutations found in the literature is not appreciated, because of the different parameters used by different GAs, such as the number of generations, the mutation rate, crossover rate, population size, selection method, initial population seeding, etc., since each of these parameters affects the results of the GA significantly.

Time complexity for most of mutations found in the literature designed for the TSP ranges from O(1) (such as the simple-random-swapping algorithms) to O(N) (for more complex mutations such as the slide, inversion and RGIBNNM mutations, where N is the number of cities in a TSP instance.

The time complexity of the proposed IRGIBNNM mutation is O(2N), since it uses two mutations of order N, namely, the inversion mutation and the RGIBNNM where each consumes O(N) time.

Accordingly, the Time complexity of the new SBM is O(4N), because it uses 3 mutations, namely slide mutation with O(N), inversion mutation with O(N), and the proposed IRGIBNNM with O(2N), comparing to the old SBM, which has O(10N) as it uses ten O(N) mutations. Asymptotically speaking, both of the proposed methods are of O(N), but

in practice they definitely consume more time than most of the mutations found in the literature. Surprisingly, both algorithms might be used to speed up the GA; this is due to their fast convergence to a minimum solution. See Figures 9 and 10, using just the first 100 iterations the GA converged to high quality solutions.

5 Conclusion

In this paper, we propose a hybrid mutation based on knowledge of the TSP and random swapping) called "IRGIBNNM" to enhance the performance of the GA for solving the TSP. We have compared the performance of the IRGIBNNM with three existing mutations, in addition to the SBM, which in this work used three mutations including the proposed one.

The experimental results of 12 TSP instances show the efficiency of the proposed mutation, and the strength of the new SBM, both of the proposed methods benefit from randomness and knowledge provided by the nearest neighbor approach. Also, both methods benefit from the increased probability of getting new high quality solutions due to the use of more than one mutation.

The high quality solutions for the TSP obtained by a GA, which used only the mutation operator, without using other advanced options that used GA by state-of-the-art such as advanced crossovers, initial seeding, advanced selection methods, adaptive change of population size and mutation/crossover rates, etc. The future work will focus on employing the proposed method with other advanced operators to further enhance the performance of the GA when applied for solving the TSP.

References

[1] SONI, N., KUMAR, T. Study of various mutation operators in genetic algorithms. *International Journal of Computer Science and Information Technologies* [online]. 2014, **5**(3), p. 4519-4521. ISSN 0975-9646. Available from: http://ijcsit.com/docs/Volume%205/vol5issue03/ijcsit20140503404.pdf

[2] POTVIN, J.-Y. Genetic algorithms for the traveling salesman problem. *Annals of Operations Research* [online]. 1996, **63**(3), p. 337-370. ISSN 0254-5330, eISSN 1572-9338. Available from: https://doi.org/10.1007/BF02125403

[3] HASSANAT, A. B. A., ALKAFAWEEN, E. On enhancing genetic algorithms using new crossovers. *International Journal of Computer Applications in Technology* [online]. 2017, **55**(3), p. 202-212. ISSN 0952-8091, eISSN 1741-5047. Available from: https://doi.org/0.1504/IJCAT.2017.084774

[4] SAMANTA, S., DE, A., SINGHA, S. Solution of traveling salesman problem on scx based selection with performance analysis using genetic algorithm. *International Journal of Engineering Science and Technology* [online]. 2011, **3**(8), p. 6622-6629. eISSN 0975-5462. Available from: http://www.ijest.info/docs/IJEST11-03-08-258.pdf

[5] RAO, A., HEGDE, S. K. Literature survey on travelling salesman problem using genetic algorithms. *International Journal of Advanced Research in Eduation Technology* [online]. 2015, **2**(1), p. 42-45. ISSN 2394-6814, eISSN 2394-2975. Available from: http://ijaret.com/wp-content/themes/felicity/issues/vol2issue1/anitha.pdf

[6] RAI, K., MADAN, L., ANAND, K. Research paper on travelling salesman problem and it's solution using genetic algorithm. *International Journal of Innovative Research in Technology* [online]. 2014, **1**(11), p. 103-114. ISSN 2349-6002. Available from: http://ijirt.org/master/publishedpaper/IJIRT101672_PAPER.pdf

[7] FREISLEBEN, B., MERZ, P. A genetic local search algorithm for solving symmetric and asymmetric traveling salesman problems. In: IEEE International Conference on Evolutionary Computation : proceedings [online]. 1996. ISBN 0-7803-2902-3. Available from: https://doi.org/10.1109/ICEC.1996.542671

[8] LARRANAGA, P., KUIJPERS, C. M. H., MURGA, R. H., INZA, I., DIZDAREVIC, S. Genetic algorithms for the travelling salesman problem: a review of representations and operators. *Artificial Intelligence Review* [online]. 1999, **13**(2), p. 129-170. ISSN 0269-2821, eISSN 1573-7462. Available from: https://doi.org/10.1023/A:1006529012972

[9] DORIGO, M., GAMBARDELLA, L. M. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation* [online]. 1997, **1**(1), p. 53-66. ISSN 1089-778X, eISSN 1941-0026. Available from: https://doi.org/10.1109/4235.585892

[10] KARKORY, F. A., ABUDALMOLA, A. A. Implementation of heuristics for solving travelling salesman problem using nearest neighbour and minimum spanning tree algorithms. *International Journal of Computer and Information Engineering* [online]. 2013, **7**(10), p. 1524-1534. ISNI 0000000091950263. Available from: https://publications.waset.org/17101/pdf

[11] MALEK, M., GURUSWAMY, M., PANDYA, M., OWENS, H. Serial and parallel simulated annealing and tabu search algorithms for the traveling salesman problem. *Annals of Operations Research* [online]. 1989, **21**(1), p. 59-84. ISSN 0254-5330, eISSN 1572-9338. Available from: https://doi.org/10.1007/BF02022093

[12] GENDREAU, M., HERTZ, A., LAPORTE, G. A tabu search heuristic for the vehicle routing problem. *Management Science* [online]. 1994, **40**(10), p. 1276-1290. ISSN 0025-1909, eISSN 1526-5501. Available from: https://doi.org/10.1287/mnsc.40.10.1276

[13] SHI, X. H., LIANG, Y. C., LEE, H. P., LU, C., WANG, Q. X. Particle swarm optimization-based algorithms for TSP and generalized TSP. *Information Processing Letters* [online]. 2007, **103**(5), p. 169-176. ISSN 0020-0190, eISSN 1872-6119. Available from: https://doi.org/10.1016/j.ipl.2007.03.010

[14] DURBIN, R., SZELISKI, R., YUILLE, A. An analysis of the elastic net approach to the traveling salesman problem. *Neural Computation* [online]. 1989, **1**(3), p. 348-358. ISSN 0899-7667, eISSN 1530-888X. Available from: https://doi.org/10.1162/neco.1989.1.3.348

[15] AARTS, E. H., STEHOUWER, H. P. Neural networks and the travelling salesman problem. In: ICANN'93 : proceedings. 1993.

[16] ALKAFAWEEN, E. *Novel methods for enhancing the performance of genetic algorithms*. Master thesis. Jordan: Mutah University, 2015.

[17] GOLDBERG, D. E. *Genetic algorithms in search, optimization, and machine learning*. 1.ed. Boston: Addison-Wesley Professional, 1989. ISBN 978-0201157673.

[18] HOLLAND, J. H. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, MA: MIT Press, 1975. ISBN 9780262082136.

[19] KOREJO, I., YANG, S., LI, CH. A directed mutation operator for real coded genetic algorithms. In: European Conference on the Applications of Evolutionary Computation 2010 : proceedings [online]. Berlin, Heidelberg: Springer, 2010. ISBN 978-3-642-12238-5, eISBN 978-3-642-12239-2. p. 491-500. Available from: https://doi.org/10.1007/978-3-642-12239-2_51

[20] MOHAMMED, A. A., NAGIB, G. Optimal routing in ad-hoc network using genetic algorithm. *International Journal of Advanced Networking and Applications* [online]. 2012, **3**(05), p. 1323-1328. eISSN 0975-0282. Available from: https://www.ijana.in/papers/V3I5-4.pdf

[21] BENKHELLAT, Z., BELMEHDI, A. Genetic algorithms in speech recognition systems. In: 2012 International Conference on Industrial Engineering and Operations Management : proceedings. 2012. ISBN 9781629939117, p. 853-858.

[22] PAULINAS, M., USINSKAS, A. A survey of genetic algorithms applications for image enhancement and segmentation. *Information Technology and Control*. 2015, **36**(3), p. 278-284. ISSN 1392-124X, eISSN 2335-884X.

[23] SRIVASTAVA, P. R., KIM, T.-H. Application of genetic algorithm in software testing. *International Journal of Software Engineering and its Applications*. 2009, **3**(4), p. 87-96. ISSN 1738-9984.

[24] MICHALEWICZ, Z. *Genetic algorithms+data structures=evolution programs* [online]. 3. ed. Berlin Heidelberg: Springer Science & Business Media, 2013. ISBN 978-3-540-60676-5, eISBN 978-3-662-03315-9. Available from: https://doi.org/10.1007/978-3-662-03315-9

[25] YUGAY, O., KIM, I., KIM, B., KO, F. I. Hybrid genetic algorithm for solving traveling salesman problem with sorted population. In: 2008 Third International Conference on Convergence and Hybrid Information Technology : proceedings [online]. 2008. ISBN 978-0-7695-3407-7. Available from: https://doi.org/10.1109/ICCIT.2008.373

[26] HASSANAT, A., PRASATH, V., ABBADI, M., ABU-QDARI, S., FARIS, H. An improved genetic algorithm with a new initialization mechanism based on regression techniques. *Information* [online]. 2018, **9**(7), p. 167-196. eISSN 2078-2489. Available from: https://doi.org/10.3390/info9070167

[27] KOREJO, I. A., YANG, S., BROHI, K., KHUHRO, Z. U. A. Multi-population methods with adaptive mutation for multi-modal optimization problems. *International Journal on Soft Computing, Artificial Intelligence and Applications* [online]. 2013, **2**(2), p. 1-11. ISSN 2319-4081, eISSN 2319-1015. Available from: https://doi.org/10.5121/ijscai.2013.2201

[28] FRIEDRICHS, T., OLIVETO, P. S., SUDHOL, D., WITT, C. Analysis of diversity-preserving mechanisms for global exploration. *Evolutionary Computation* [online]. 2009, **17**(4), p. 455-476. ISSN 1063-6560, eISSN 1530-9304. Available from: https://doi.org/10.1162/evco.2009.17.4.17401

[29] YANG, S. Adaptive non-uniform mutation based on statistics for genetic algorithms. In: 4th Annual Conference on Genetic and Evolutionary Computation GECCO'02 : proceedings. Part II. 2002. ISBN 978-1-55860-878-8, p. 650-657.

[30] HASSANAT, A., ALMOHAMMADI, K., ALKAFAWEEN, E. ABUNAWAS, E., HAMMOURI, A., PRASATH, V. B. S. Choosing mutation and crossover ratios for genetic algorithms - a review with a new dynamic approach. *Information* [online]. 2019, **10**(12), p. 390-426. eISSN 2078-2489. Available from: https://doi.org/10.3390/info10120390

[31] NICOARA, E. S. Mechanisms to avoid the premature convergence of genetic algorithms. *Petroleum - Gas University of Ploiesti Bulletin - Mathematics, Informatics, Physics Series* [online]. 2009, **61**(1), p. 87-96. ISSN 1224-4899, eISSN 2067-242X. Available from: http://bulletin-mif.unde.ro/docs/20091/12NICOARA_SIMONA.pdf

[32] ABDOUN, O., ABOUCHABAKA, J. A comparative study of adaptive crossover operators for genetic algorithms to resolve the traveling salesman problem. *International Journal of Computer Applications* [online]. 2011, **31**(11), p. 49-57. ISSN 0975-8887. Available from: https://arxiv.org/ftp/arxiv/papers/1203/1203.3097.pdf

[33] SIVANANDAM, S. N., DEEPA, S. N. *Introduction to genetic algorithms* [online]. 1. ed. Berlin Heidelberg: Springer -Verlag, 2008. ISBN 978-3-540-73189-4, eISBN 978-3-540-73190-0. Available from: https://doi.org/10.1007/978-3-540-73190-0

[34] BANZHAF, W. The "molecular" traveling salesman. *Biological Cybernetics* [online]. 1990, **64**(1), p. 7-14. ISSN 0340-1200, eISSN 1432-0770. Available from: https://doi.org/10.1007/BF00203625

[35] MICHALEWICZ, Z. *Genetic algorithms+data structures=evolutionary programs* [online]. 1. ed. Berlin Heidelberg: Springer-Verlag, 1992. ISSN 1431-0066, eISBN 978-3-662-02830-8. Available from: https://doi.org/10.1007/978-3-662-02830-8

[36] FOGEL, D. A. A parallel processing approach to a multiple travelling salesman problem using evolutionary programming. In: Fourth annual Symposium on Parallel Processing : proceedings. IEEE Orange County Computer Society, 1990.

[37] LOUIS, S. J., TANG, R. Interactive genetic algorithms for the traveling salesman problem. In: GECCO-99: Genetic and Evolutionary Computation Conference : a Joint Meeting of the Eighth International Conference on Genetic Algorithms (ICGA-99) and the Fourth Annual Genetic Programming Conference (GP-99) : proceedings. Vol. 1. Morgan Kaufmann Publishers, 1999. ISBN 1558606114, 9781558606111.

[38] HASSANAT , A. A. B., ALKAFAWEEN, E., AL-NAWAISEH, N. A., ABBADI, M. A., ALKASASSBEH, M., ALHASANAT, M. B. Enhancing genetic algorithms using multi mutations: experimental results on the travelling salesman problem. *International Journal of Computer Science and Information Security* [online]. 2016, **14**(7), p. 785-801. ISSN 1947-5500. Available from: https://arxiv.org/ftp/arxiv/papers/1602/1602.08313.pdf

[39] HONG, T. P., WANG, H. S., LIN, W. Y., LEE, W. Y. Evolution of appropriate crossover and mutation operators in a genetic process. *Applied Intelligence* [online]. 2002, **16**(1), p. 7-17. ISSN 0924-669X, eISSN 1573-7497. Available from: https://doi.org/10.1023/A:1012815625611

[40] DENG , Y. LIU, Y., ZHOU, D. An improved genetic algorithm with initial population strategy for symmetric TSP. *Mathematical Problems in Engineering* [online]. 2015, Article ID 212794. ISSN 1024-123X, eISSN 1563-5147. Available from: https://doi.org/10.1155/2015/212794

[41] HILDING, F., WARD, K. Automated crossover and mutation operator selection on genetic algorithms. In: 9th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems : proceedings [online]. Berlin Heidelberg: Springer-Verlag, 2005. ISBN 978-3-540-28897-8, eISBN 978-3-540-31997-9. Available from: https://doi.org/10.1007/11554028

[42] KATAYAMA, K., SAKAMOTO, H., NARIHISA, H. The efficiency of hybrid mutation genetic algorithm for the travelling salesman problem. *Mathematical and Computer Modelling*. 2000, **31**(10-12), p. 197-203. ISSN 0895-7177.

[43] HONG, T. P., WANG, H. S., CHEN, W. CH. Simultaneously applying multiple mutation operators in genetic algorithms. *Journal of Heuristics* [online]. 2000, **6**(4), p. 439-455. ISSN 1381-1231, eISSN 1572-9397. Available from: https://doi.org/10.1023/A:1009642825198

[44] SUN, W. A novel genetic admission control for real-time multiprocessor systems. In: 2009 International Conference on Parallel and Distributed Computing, Applications and Technologies : proceedings [online]. 2009. ISSN 2379-5352, eISBN 978-0-7695-3914-0. Available from: https://doi.org/10.1109/PDCAT.2009.10

[45] REINELT, G. TSPLIB [online]. University of Heidelberg, 1996. Available from: http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95

[46] DONG, M., WU, Y. Dynamic crossover and mutation genetic algorithm based on expansion sampling. In: International Conference on Artificial Intelligence and Computational Intelligence AICI 2009 : proceedings [online]. Lecture Notes in Computer Science, vol. 5855. Berlin, Heidelberg: Springer, 2009. ISBN 978-3-642-05252-1, eISBN 978-3-642-05253-8. Available from: https://doi.org/10.1007/978-3-642-05253-8_16

[47] CONTRERAS-BOLTON, C., PARADA, V. Automatic combination of operators in a genetic algorithm to solve the traveling salesman problem. *PlOS one* [online]. 2015, **10**(9). eISSN 1932-6203. Available from: https://doi.org/10.1371/journal.pone.0137724

[48] SPEARS, W. M. Adapting crossover in evolutionary algorithms. In: *Evolutionary Programming IV: Proceedings of the Fourth Annual Conference on Evolutionary Programming*. MCDONNELL, J. R., REYNOLDS, R. G., FOGEL, D. B. (eds.). 1995. eISBN 9780262290920. Available from: https://doi.org/10.7551/mitpress/2887.001.0001

Gero Gerber

ALARM SYSTEMS IN BUILDINGS - INVESTIGATION OF EFFECTIVENESS DEPENDING ON BUILDING TYPE AND USER PROFILE

Alarm systems are used to warn passengers and employees in railway stations, airports, logistics facilities and administrative buildings of hazards; in this way, they are instructed to rescue themselves. A large online survey conducted in Germany in 2018 has shown that often sounding the alarm does not have the desired effect. The research presented in this article investigates dependence of an alarm effectiveness on type of building, user profile and type of the alarm system.

In this paper, general and building-specific measures, using which the effective sounding of the alarm and optimal behaviour of the building users in the case of danger can be achieved, are presented.

Keywords: *hazard warning, alarm systems, effectiveness, building type, user profile*

1 Introduction

Goods and passengers are mainly transported by road, rail, waterway or air. Nevertheless, the majority of transport workers do not work in vehicles, but in buildings. These include:

- railway stations,
- airports,
- logistics halls,
- administration buildings.

All the aforementioned types of buildings are special buildings. Compared to normal residential buildings, various factors can lead to an increased risk of danger for these types.

In the case of danger, all the affected persons must quickly and effectively be informed about the dangerous situation to give them the opportunity to bring themselves to safety. In large and complex buildings, rapid information transmission only works with the help of alarm systems. The experience reports of numerous experts have shown that in many cases, even well-functioning systems do not lead to the intended reaction.

Example: In 2015, a medium-sized regional railway station in southern Germany was renovated. Acoustic signalling devices were installed throughout the entire building to provide an alarm in the event of a fire. The functional test of the alarm was performed without prior notice during a normal station operation. The warning signals sounded throughout the entire building. But how did people react? Neither the passengers nor the employees left the station building. Nobody took the alarm signal seriously, and nobody reacted as intended.

Such and similar experiences led us to the idea of scientifically investigating the effectiveness of alarm systems depending on the building type, user profile and type of the alarm system. The technical literature presents studies on the effect and design of danger signals. The books titled 'Human factors in alarm design' [1] and 'Design of danger signals' [2] contain a well-structured summary of the research results. Influence of the acoustic parameters of signals on the reaction time was investigated in a British Standard study [3].

The investigation of the effectiveness of entire alarm systems broke new ground.

2 Tasks of alarm systems

The actions required in a hazardous situation depend on the nature of the hazard (Figure 1). In the event of a fire or bomb alarm, all people should leave the building as quickly as possible. If the danger is outside (e.g. chemical alarm or storm), people must remain inside the building. In an amok situation, all endangered people must seek shelter in the building and barricade themselves in.

The aim of a technical alarm is always to save time and enable self-rescue [4, p. 245-247].

3 Possibilities of sounding the alarm

In small- and medium-sized buildings, it is usually sufficient for people to alert each other of a danger. This can be done at random or according to organisational specifications (e.g. with alarm chains or evacuation

Gero Gerber
Hohenfelden, Germany
E-mail of corresponding author: Gero.Gerber@factum-gmbh.com

Table 1 *Distribution of hazard potentials among building types*

Risk caused by / Type of building	High number of persons	Lack of local knowledge	Long or unclear escape routes	Increased risk of fire formationor fire propagation	Increased risk of accidents	Difficult access for rescue forces	Limited mobility orresponsiveness	Many people in need of supervision and guidance
Railway stations Airports	*	* 1)	•	•	•	•	•	•
Logistics halls								
High-rise and admin-buildings	•	• 2)	•	•	•	•		

Legend:

*	typical	1)	for passengers and external personnel
•	possible	2)	for temporary employees
	unlikely	3)	for customer and guests

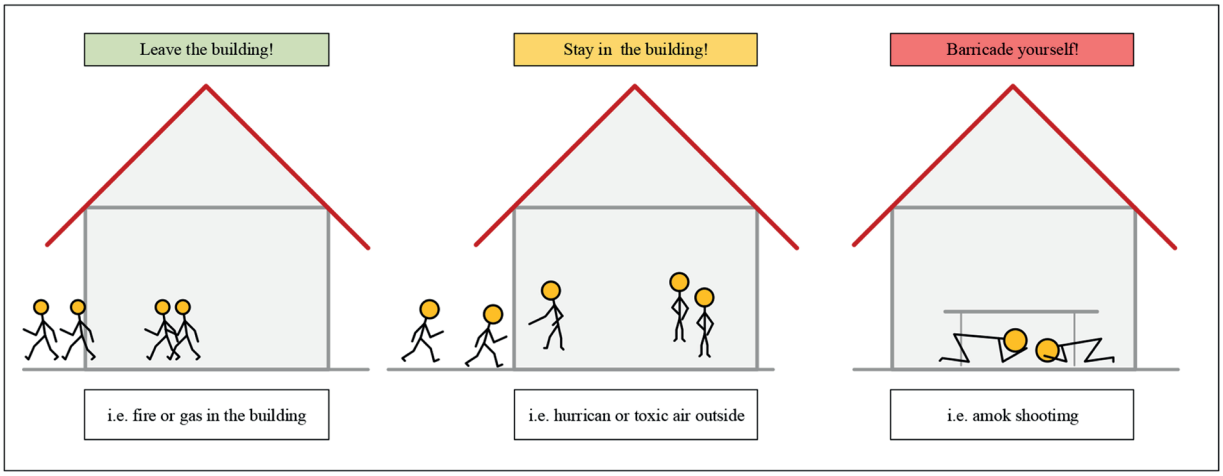


Figure 1 *Intended reactions in the case of an alarm*

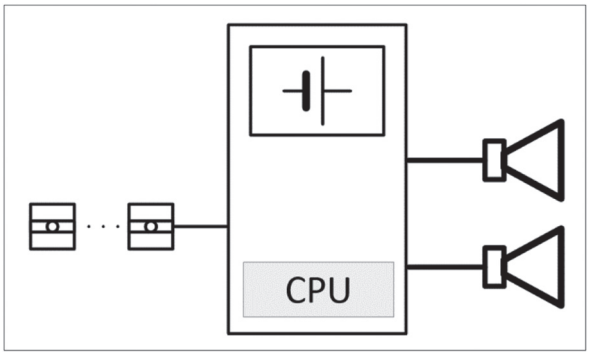


Figure 2 *Components of an alarm system*

Table 2 *Alarm types in various special buildings*

Type of building	Voice alarmsystem	Acoustic andvisual signallers	Silent alerting
Railway stations	X		(X)
Airports	X		(X)
Logistic halls		X	
Administration buildings		X	
Hospitals and nurseries			X

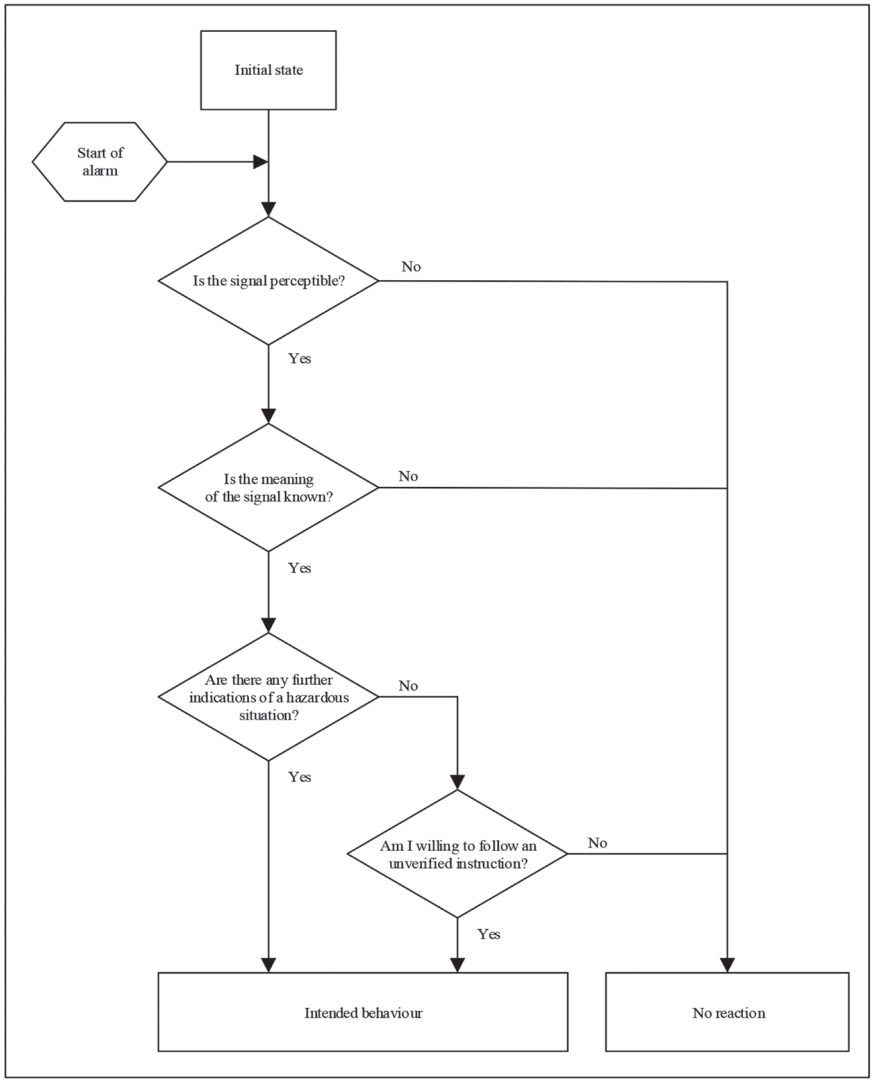


Figure 3 *Decision making*

assistants). A personal alarm will take too long if the buildings become larger. Technical aids, such as fire bells, sirens or megaphones, must be used. The range of simple alarm devices for extensive and complex buildings is too short. Alarm systems must be installed to quickly reach all the endangered persons. Those systems must consist of one or more triggering devices, a control centre with a power supply and the alarm devices (Figure 2). Often, the voice alarm systems are preferred in buildings without a fixed user group. The book titled 'Speech Alarm Systems and Electroacoustic Emergency Warning Systems' contains detailed basic information and notes on planning [5].

In practice, the following plant types are particularly used:

- systems with acoustic and/or optical signalling devices,
- voice alarm systems,
- systems for the silent alarming the personnel.

Table 2 shows the typical areas of application. Fields with "X" show typical applications. An "(X)" means that the solution is rarely used.

4 Normative requirements

Standards of the EN54 series "Fire alarm systems" contain the technical requirements for the components of the fire alarm and voice alarm systems. EN 54 Part 32 contains specifications for the internal alarm of fire alarm systems. EN54 Part 16 specifies the requirements for voice alarm control panels [6]. Requirements for planning, installation and operation of the systems are defined in national standards. In Germany, DIN 14675 "Fire alarm systems - Construction and operation" [7] and DIN VDE 0833-4 "Hazard alarm systems - Specifications for voice alarm systems" are the standards [8]. In the US, alarm systems are built according to the guideline, NFPA72. Chapter 18 of the NFPA is about notification appliances [9, p. 109-115].

The standards described give recommendations for alarm sound levels and speech intelligibility. None of the above standards contains criteria for assessing the effectiveness of the alarm systems.

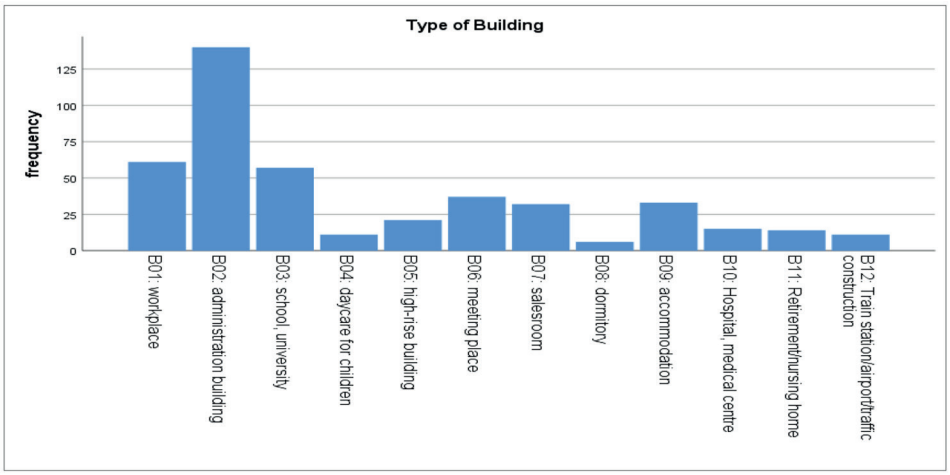


Figure 4 *Frequencies of the building types in the case reports*

5 Human behaviour in hazardous situations

Human behaviour patterns in a danger situation are ancient and essential for survival. An individual recognises a threat and has to decide what to do in a very short time. The two main options are **flight** or **fight**. In social communities of humans, but also of animals, it has proven successful to warn the other individuals of the group, as well. The warning can be an invitation to flight or a call for support in the fight.

The term "alarm" has many meanings in psychology, technology and colloquial language. Psychologist Dr Neville Stanton writes about this in the introductory article of his anthology "Human Factors in ALARM DESIGN": A frequently given definition of an alarm is "a significant attractor of attention"! [1, p. 2].

The first prerequisites for a correct reaction is the perception of the alarm signals. Further preconditions are:

- knowledge of the meaning of the danger signal,
- willingness to act.

Figure 3 shows possible ways of decision-making.

6 Expectations regarding the effectiveness of alarm systems

Effectiveness of a measure describes the ratio of the real result to the desired goal and can take values between 0 and 1 or 0% and 100%.

$$\text{Effectiveness} = \frac{\text{result}}{\text{goal}}. \tag{1}$$

The present paper focuses on effectiveness of the alarm systems. The highest effectiveness is achieved when all the alarmed persons behave as intended. If no one reacts, the effectiveness is zero. Effectiveness of an alarm system can be described with the following formula:

$$\text{Effectiveness} = \frac{\text{Number of persons behaving as intended}}{\text{Number of persons to be alarmed}}. \tag{2}$$

However, two questions should be answered: What are expectations regarding the effectiveness of an alarm system? How can this alarm system be evaluated? These questions are answered by conducting interviews with 12 experts in this field, including professors at universities, fire protection planners, senior fire brigade employees and test experts for safety systems.

The experts agreed that alarm systems must be planned and built in such a way that, normally, every person present can receive the alarm signal. According to most experts, exceptions are permitted if people with headphones and loud music or by taking alcohol, drugs or sleeping pills are able to disconnect themselves from their environment. For people with hearing impairments and those who wear hearing protection at work, the alarm must be ensured by means of organisation or other aids.

The experts also agreed that the objective is for all the alarmed persons to react as intended, even if this is not always possible in practice. Certain groups of people may not be able to react as intended. These include

- small children,
- mentally ill people,
- persons with physical disabilities.

The required effectiveness of an alarm in a building depends on how many people who do not respond as intended can be rescued in an acceptable time by the emergency services.

Example: If up to 50 people can be rescued by the fire brigade in an acceptable time in the event of a fire in a cinema with 1000 people, the effectiveness of the alarm must be 95%.

It is not advisable to specify a fixed set point of effectiveness for all the building types. Even for a concrete building it is difficult to fix a certain number in the planning phase. Nevertheless, a great agreement among the respondents was that in order to achieve the protection goal, the effectiveness achieved must be 90% or higher, in most cases.

When evaluating the case reports from the survey, all examples with an effectiveness of <90% were rated as insufficient.

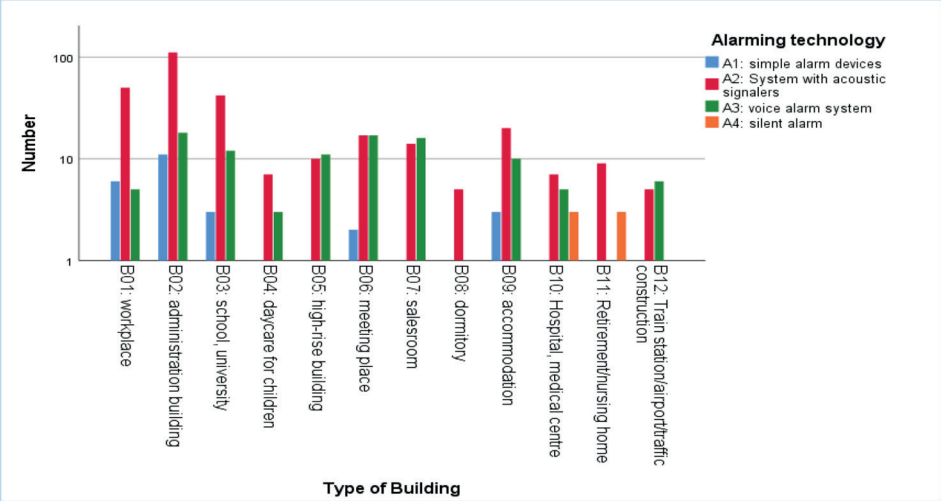


Figure 5 *Distribution of alarm systems in buildings*

7 Investigation of effectiveness

The central part of this research work was the collection and evaluation of the case reports of real alarms in buildings. The data were collected in a Germany-wide online survey between 30.11.2017 and 25.01.2019. The participants were all the people with professional experience in fire protection or security technology. Of the 609 data records, 438 fulfilled all the formal requirements to be included in the survey. Case reports from the second hand or with incomplete information were not used. Only 16.5% of the case reports were related to an announced practice alarm, indicating that in 83.5% of the cases, the alarm came as a surprise to those affected. The persons behaved as in a real danger case.

Figure 4 shows distribution of the building types within the case reports. A total of 11 case reports represent railway stations, airports and traffic structures. The category 'workplaces' contains at least 11 case reports from freight forwarding and logistics halls. The proportion of transport companies cannot be exactly determined in the case of administrative buildings. However, the specific user sector is likely to have only a minor influence on the behaviour of users in the event of an alarm; hence, the results of the study can also be applied to administrative buildings in the transport sector.

The survey results showed in what frequency the types of alarm systems are used in different building types (Figure 5).

The second step was to analyse how the effectiveness of an alarm system depends on the building type, user profile and type of alarm system.

Examination of the three factors with the Pearson Chi-square test showed no significant correlation within the three factors. All the three factors had to be individually examined.

The user profile and the building type have a significant influence on effectiveness of the alarm. The type of alarm system has only a minor influence.

Overall, the average effectiveness of the alarm systems is below 90% for all the user profiles and in all the building types, with the exception of day care centres for children and is therefore inadequate. Buildings with a fixed circle of users (e.g. logistics halls and administration buildings) performed significantly better at around 80% than buildings with many non-local users (e.g. airports and railway stations), which reached around 40%. (Figures 6 and 7).

The user profile and building type are closely related. Nevertheless, the statistical investigation has shown that they are independent variables. The error bar displayed in the columns shows that 95% of all the results are within this range.

Interestingly, there were case reports for each type of building and user profile in which the alarm worked very well. This indicates that there are other important influencing factors in addition to the factors investigated. In order to narrow these down, the expert interviews and the comments of the survey were examined qualitatively.

Both the experts interviewed and the survey participants cited organisational deficits as the main reason for the insufficient effectiveness.

In cooperation with the Chair of Psychology at the University of Erfurt, the psychological aspects of alerting were investigated and the following findings were obtained:

- Previous experiences with alarm situations showed no significant influence on the behaviour in the case of a fire alarm.
- Occurrence of frequent alarms without any real danger results in unserious responses to danger signals [10, p. 33-35].
- By changing the acoustic properties of the alarm signal, only slight improvements in reaction time are possible [11, p. 20-23].

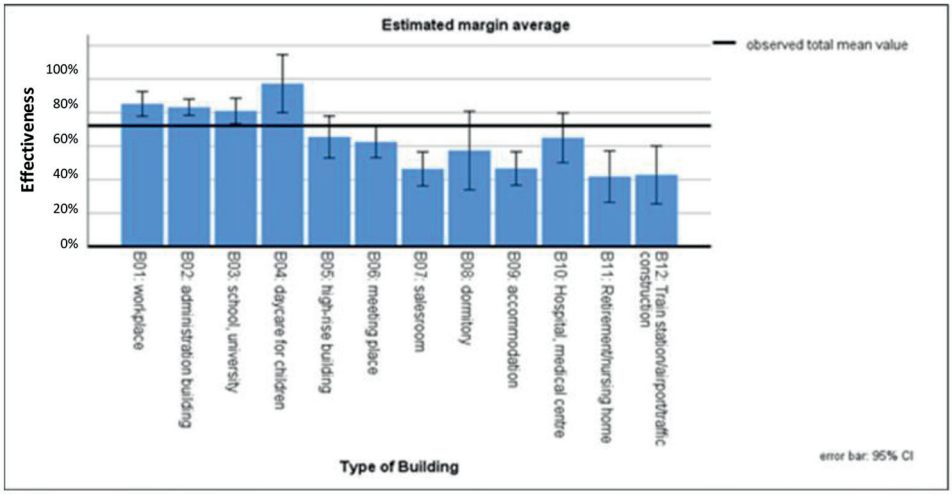


Figure 6 *Effectiveness of alarm systems for various building types*

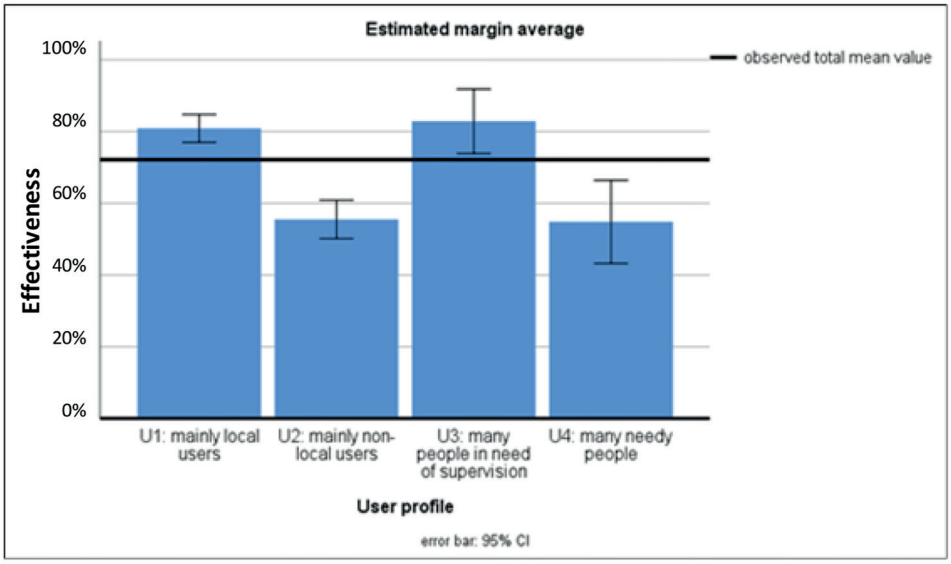


Figure 7 *Effectiveness of alarm systems for various user profiles*

8 Ways to improve effectiveness of the alarm systems

The following recommendations for planning and operation of the alarm systems are given based on the expert interviews, hints of the survey participants and own experiences.

8.1 General measures for all building types

In the event of a danger, an alarm system can only be highly effective by a coordinated interaction of technical and organisational measures. The most important technical tasks include the following:

- ensuring signal perceptibility in the entire alarm area (volume, speech intelligibility and visibility),
- use of suitable signals (known tone sequences, multilingual text announcements and comprehensible visual displays),
- high plant availability and replacement measures in the event of malfunctions.

The most important organisational measures include the following:

- regular training and instruction of all the permanent users of the building,
- regular exercises under realistic conditions (unannounced test alarms),
- consistent evaluation of exercises (commendation, criticism and deficit identification and elimination).

8.2 Administrative buildings

Administrative buildings do not pose a major fire protection challenge in many cases. The probability of a hazardous situation is low. Everyone is awake, mobile and able to quickly react when needed; therefore, no fire alarm or alarm systems are installed in many small- and medium-sized administrative buildings. Accordingly, investigation of

the alarm systems' effectiveness is primarily related to large and complex administrative buildings.

Use of the alarm systems with simple acoustic and, if necessary, optical signalling devices is sufficient in most cases. Once the perceptibility of the signal has been ensured at all the workplaces, break rooms and sanitary rooms, only the organisational measures need to take effect. Unfortunately, the willingness to react to an alarm signal is not particularly pronounced among office employees. Hence, regular instruction and exercises, which are also evaluated, are particularly important.

8.3 Workplaces

Workplaces always have a fixed circle of users. The above-described measures are applicable. Cargo ports and logistic halls often have other acoustic signals in addition to the alarm signal. The following technical measures can be used to achieve an effective alarm system in the event of a danger:

- use of additional optical alarm devices (flashlights or large-surface illuminated displays),
- plant shutdown,
- use of displays on the control station of the plant technology,
- cyclic switching of room lighting,
- use of pager or portable phones with a vibration alarm,
- equipping of evacuation assistants with inscribed warning vests and whistles,
- additional use of powerful motor sirens.

8.4 Railway stations and airports

The problem in railway stations and airports is that a very small group of permanent employees is confronted with a large group of people unfamiliar with the place. The success of the alarm depends to a large extent on the correct behaviour of this small group. Regularly instruction and training this group are the basic prerequisites for achieving the protection goals. Recommendations for special building types are given below.

Large traffic structures, such as railway stations and airports, are designed for the simultaneous and short-term presence of many people who are unfamiliar with the place and are mostly in motion. Many foreign travellers can stay in large cities and transport hubs.

Alerting travellers is a special challenge. Passengers who hear a hazard warning, but do not recognise a hazard would much rather proceed in their aircraft or train than follow a presumably unfounded instruction.

Experience has shown that simple acoustic signalling devices are ineffective in such facilities. They garner attention briefly; then, they are ignored. After a few minutes, they are perceived as annoying. Even a high-quality voice alarm system does not guarantee effective alerting if not accompanied with organisational measures.

The following recommendations could be extracted from the experiences of the survey participants and the interviewed experts:

- If possible, the first step should only be to alarm the staff (e.g. via coded announcements).
- If a hazard condition is verified, a personal situation-related announcement should be made by the person responsible for safety.
- An automatic hazard announcement should be made if the pre-alarm has not been acknowledged within a predefined time.
- Evacuation helpers, who are recognisable as authority figures, should call on travellers, employees and shop assistants to go to safety. This task can be performed by uniformed railway officials, police officers, security personnel or employees with marked warning vests.

9 Summary

Alarm systems are designed to quickly inform people in large buildings about a hazardous situation and induce them to take a certain action. In practice, this goal is not achieved in many cases. The situation is particularly critical in buildings without the fixed user groups, such as railway stations and airports.

Building- and user-specific organisational measures are required in addition to suitable alarm systems to achieve the desired protection goal. The decisive factors herein are regular and practical exercises, their professional evaluation and elimination of identified deficits.

Acknowledgement

The research was supervised by Prof. Dr Ales Dudacek from the Technical University of Ostrava. Prof. Dr Jorg Reintsema from Cologne University of Applied Sciences was recruited as the second supervisor. Professor Dr Tilmann Betsch of the University of Erfurt took over the consultation to the psychological aspects. I would like to thank all professors, the interview partners and the participants of the survey for their support.

References

[1] STANTON, N. A human factors approach. In: *Human factors in alarm design*. STANTON, N. (ed.). London, Washington, D.C: Taylor & Francis, 1994, p. 1-14.

[2] MALTER, B., GUSKI, R. *Gestaltung von Gefahrensignalen / Design of hazard signals*(in German). Bremerhaven: Wirtschaftsverl. NW, 2001. Schriftenreihe der Bundesanstalt fur Arbeitsschutz und Arbeitsmedizin Fb/ Series of publications from the Federal Institute for Occupational Safety and Health. ISBN 935-3-89701-780-6.

[3] BSI. DD 240, 1997, Fire Safety Engineering in Buildings.

[4] GERBER, G. *Brandmeldeanlagen. Planen, Errichten, Betreiben / Fire alarm systems. Planning, building, operating* (in German). 5. ed. Heidelberg: Huthig, 2018. ISBN 978-3-8101-0464-9.

[5] SIMON, A. *Sprach alarm anlagen und elektroakustische Notfallwarnsysteme. Neuausgabe des beliebten Fachbuchs "Fachkraft fur Sprach alarm anlagennach DIN 14675" / Voice alarm systems and electro-acoustic emergency warning systems. New edition of the popular specialist book "Specialist for voice alarm systems according to DIN 14675"* (in German). Heidelberg: Huthig, 2018. ISBN 978-3-8101-0454-0.

[6] DIN Deutsches Institut fur Normung. EN54-16, DIN EN 54-16:2008-06, Brandmeldeanlagen_- Teil_16: Sprachalarmzentralen; Deutsche Fassung EN_54-16:2008 / German institute for standardization. EN54-16, DIN EN 54-16: 2008-06, fire alarm systems_- part_16: voice alarm centers; German version EN_54-16:2008 (in German). Berlin: Beuth Verlag GmbH.

[7] DIN DeutschesInstitut fur Normung. 14675-1, Brandmeldeanlagen / German institute for standardization. 14675-1, fire alarm systems (in German). Berlin: Beuth Verlag GmbH.

[8] DKE. 0833-4, Gefahren meldeanlagen fur Brand, Einbruch und Uberfall / Alarm systems for fire, burglary and robbery (in German). Berlin: VDE-Verlag.

[9] NFPA 72 National fire alarm and signaling code handbook. 8 ed. Quincy, MA: National Fire Protection Association, 2015. ISBN 978-145591201-8.

[10] HEHN, K. *Decision making during fire alarms. The role of experience and perceived risk*. Bachelor thesis. Erfurt, 2019.

[11] DIECKMANN, T. J. *Das Spiel mitdem Feuer. Wie Alarmsignale die menschliche Reaktionsbereitschaft und die wahrgenommene Dringlichkeit beeinflussen* / Playing with fire. How alarm signals affect human responsiveness and perceived urgency (in German). Bachelor thesis. Erfurt, 2019.

komunikácie
COMMUNICATIONS

VEDECKÉ LISTY ŽILINSKEJ UNIVERZITY
SCIENTIFIC LETTERS OF THE UNIVERSITY OF ZILINA

VOLUME 22 Number 3

https://doi.org/10.26552/com.J.2020.3

Author guidelines

- All papers have to deal with the topic of transport and be submitted strictly within one of the listed subtopics. Please, refer to list of topics and subtopics here and indicate it clearly when submitting your paper.
- Submitted papers must be unpublished and must not be currently under review for any other publication.
- Manuscripts written in good English must include abstract and keywords also written in English. The abstract should not exceed 10 lines.
- Submitted manuscripts should not exceed 20 pages including figures and graphs
- Submission should be sent by e-mail – as an attachment – to the following address: komunikacie@uniza.sk.
- The author‘s exact mailing address, full names, E-mail address, telephone or fax number, the name and address of the organization and workplace (also written in English) must be enclosed.
- For all manuscripts a double-blind peer review by at least two independent reviewers and language correction is mandatory.
- After reviewing and incorporating the editor‘s comments, the final draft (before printing) will be sent to authors for final review and minor adjustments.

The full author quidelines are available at:
http://komunikacie.uniza.sk/index.php/communications/guidelines

Address of the editorial office:
University of Zilina
EDIS – Publishing House
Univerzitna 8215/1
010 26 Zilina, Slovakia

E-mail: komunikacie@uniza.sk

Individual issues of the journal can be found on:
http://komunikacie.uniza.sk

Each paper was reviewed by two reviewers.

Journal is excerpted in **SCOPUS** and **EBSCO.**

Published quarterly by University of Zilina in
EDIS – Publishing House of University of Zilina

Registered No: EV 3672/09

ISSN (print version) 1335-4205
ISSN (online version) 2585-7878

ICO 00397 563

July 2020