

Mathematical Evaluation of Passenger and Freight Rail Transport as Viewed Through the COVID-19 Pandemic and the War in Ukraine Situation

Anna Borucka¹, Edward Kozłowski^{2*}

¹ Faculty of Security, Logistics and Management, Military University of Technology, ul. Kaliskiego 2, 00-908 Warsaw, Poland

² Faculty of Management, Lublin University of Technology, ul. Nadbystrzycka 38, 20-618 Lublin, Poland

* Corresponding author's e-mail: e.kozlovski@pollub.pl

ABSTRACT

The main objective of the study was to develop a method for identifying the components of a time series disrupted by crisis events, allowing for evaluation, comparison and short-term forecasting of the impact of such situations. The article, using mathematical modelling, analyses and evaluates the impact of the pandemic and war on rail transport using Poland as an example. Three methods of local trend matching were used: locally estimated trend, locally estimated trend with seasonal components, locally estimated trend with locally estimated seasonal components. The study showed that the effects of the crisis are still felt today, but their impact on individual types of transport varied, and what's more, differences were also visible depending on the item being transported. Passengers, despite the introduction of high sanitary standards, severely limited their mobility. The freight transport market turned out to be more resistant to the impact of the pandemic, but both the pandemic and the war in Ukraine affected the volume of goods transported. The article presents a method that allows to identify time series disturbed by crisis events and therefore difficult to describe and make reliable forecasts. The method proposed by the authors is an innovative answer to the problems of identifying very variable series and can also be used in the analysis of other issues in which time series have been disturbed in a sudden and unexpected way.

Keywords: rail transport, time series disruptions, crisis situations, locally estimated scatterplot smoothing, harmonic analysis, Fourier series, weighted least squares.

INTRODUCTION

Crisis events disrupt global supply chains and transport flows. A perfect example is the COVID-19 pandemic [1]. War threats have similar effects. It is not without reason that Poland was chosen as the subject of the study, as it directly borders Ukraine, where there is currently an armed conflict. The war in Ukraine has a serious impact on the transport sector. Disruptions in the transport of goods cause huge delays in deliveries [2]. Transport instability and rapid changes in business force us to make quick decisions in a complex and demanding environment and should be supported by reliable mathematical analyses

[3, 4], not only identifying factors influencing changes in value streams, but also enabling short-term forecasting [5, 6]. The examples of forecasting of transport systems behaviour and managing crisis situations can provide excellent support for crisis teams, manufacturing and logistics enterprises [7]. They allow to determine the impact of selected factors on demand and supply, describe expected trends and provide information on possible disruptions. They also enable risk analysis, which becomes particularly important in a pandemic situation and is crucial for decision-making and strategy development [34, 39]. Its main task is to identify potential threats, assess their impact, and above all, the likelihood of occurrence. This

allows for reducing risk and minimizing negative consequences. In the event of armed conflicts, but also various types of disasters, they can provide support for companies, governmental and non-governmental agencies. Support in the form of mathematical modelling allows to make more informed decisions in crisis situations. In this article, the authors' research objective was to develop a method enabling the identification of time series disrupted by crisis events, and therefore difficult to describe and make reliable forecasts. The method proposed by the authors is an innovative answer to the problems of identifying time series in which the dynamics of the seasonality component is not constant (the dynamics of both trend and seasonality are influenced by external factors). The method presented can also be used in the analysis of other issues in which the time series were disrupted in a sudden and unexpected way.

BACKGROUND

The COVID-19 coronavirus pandemic has had a profound impact on all industries, including the rail transport sector. The literature in this area covers several main themes. First of all, it focuses on mitigating the consequences of the pandemic. In this context, authors strongly focus on limiting the transmission of the virus by proposing various models that describe this phenomenon [35, 37]. The main goal of such research is to limit the risk of infection and ensure the health of citizens through effective preventive strategies and intervention measures. These studies include the analysis of the effectiveness of protective measures and the assessment of the impact of health and social measures on controlling the pandemic, as well as elements of social functioning [36]. Moreover, researchers aim to identify risk factors associated with infection and evaluate the effectiveness of different diagnostic methods, treatment, and vaccines [41, 42]. The authors also identify strategies related to transportation systems to limit the consequences of this threat [8] or propose scenarios regarding the functioning of the transport system as a result of the COVID-19 pandemic [9, 10]. The impact of the pandemic on decisions about using transport has also been widely studied [11]. Crisis situations directly affect business strategy change, indirectly affecting transport and logistics flows [44]. The diversification strategy is one effective way to reduce the disruption resulting from the COVID-19

pandemic and the war in Ukraine. Changes in travel patterns and activity during the pandemic were evaluated. For this purpose, the methods of comparison and statistical analysis were most often used [12], or calculating appropriate coefficients [13]. The use of more advanced methods was also proposed. For example, Lucchesi et al. [14] used hybrid choice models, emphasizing the difficulty in studying this issue due to the presence of hidden variables [15]. Transport was also analysed in terms of the spread of the epidemic. Such research was conducted, for example, in Portugal [16], China [17] or Poland [18]. Research in this area was most popular due to the fact that the spread of the disease was one of the key problems in combating the threat [19]. The resistance of rail transport to threats was also assessed. For example, Wang et al. [20] studied the impact of COVID-19 on the Chinese economy using a forecasting model based on rail transport statistics. H. Nguyen [21] analysed indicator methods in relation to Vietnam. Selected publications concern the use of classic time series models, such as the stationary AR model [22], nonstationary ARIMA [23] and seasonal SARIMA [24], however, as the obtained results show, these models do not fully cope with identifying series containing changes caused by crisis situations.

Machine learning methods give better results using, for example, neural networks [25], support vectors (SVM) and random forests (RF) [26], however, require large data resources from a distant time horizon. Moreover, artificial intelligence (AI) may have difficulty adapting quickly in the event of rapid changes and algorithm complexity, which may lead to inaccurate predictions [27]. In [43], the authors present a special solution for determining alarm conditions based on AI techniques, including an expert system and developed algorithms for analysing certain parameters in biogas plant. This shows the need to create new solutions dedicated to the analysis of complex time series that can cope with sets with unexpected changes. That issue is addressed in this article. However, the authors keep in mind that identifying time series in which disturbances are caused by external factors is extremely difficult and requires reliable tools to ensure reliable results. As the literature shows, they are a huge hindrance to classical methods of identification and forecasting, and direct polynomial matching of the appropriate degree of the deterministic part for the entire time series or identification of the series using

ARIMA models is not always a rational action. Therefore, it is necessary to modify them and adapt known methods to take into account the complexity of the collected observations.

Due to the above, this article proposes to describe the dynamics occurring in the time series using the LOESS (locally estimated scatterplot smoothing) and LOWESS (locally weighted scatterplot smoothing) methods, which rely on local matching the trend occurring around the moment t based on realisations of the nearest neighbors, i.e. determining the trend based on the realization of sequence $\{x_s\}_{\max(1,t-k) \leq s \leq \min(t+k,n)}$. Trend identification involves the use of Weighted Least Squares (WLS). In a time series, in addition to the trend, there may also be seasonal trends. Local polynomial fitting allows only to estimate the trend but does not take into account seasonal fluctuations. Therefore, it was proposed to use LOESS both to determine the local trend and to estimate seasonal fluctuations, which is an innovative approach. Additionally, the matching of historical data was compared for the trend determined using LOESS, the trend and seasonal components using harmonic analysis, and the trend and seasonal components using the LOESS technique. For each of the above-mentioned methods, the model fit was evaluated for the training data set. Effective adaptation of the model to changing conditions was ensured by taking into account the effect of weights for empirical data, i.e. assigning greater weight to observations closer to the analysed moment, and less weight to observations further away from that moment. Then the forecast was determined and compared with the actual data. The following indicators were used to assess the matching and effectiveness of forecasts: root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), [28, 29]. Lower values of these indicators mean a better fit and a more accurate forecast. Accordingly, the contribution of this article to science is as follows:

- an innovative method was proposed to identify the dynamics of a time series in which seasonality disturbances are caused by external factors;
- an analysis of freight and passenger rail transport was made, presenting the impact of crisis situations on these flows;
- a method of local trend and seasonality matching was presented;
- the method of determining the local trend and adapting seasonal fluctuations using harmonic analysis and LOESS was presented;
- time series predictions and accuracy evaluation were made using Locally Estimated Trend, Locally Estimated Trend with Seasonal Components, Locally Estimated Trend with Locally Estimated Seasonal Components.

MATERIALS AND METHODS

LOESS and LOWESS

We consider the time series $\{x_s\}_{1 \leq s \leq n}$ around the moment t , $1 \leq t \leq n$, so the elements of the subsequence $\{x_s\}_{\max(1,t-k) \leq s \leq \min(t+k,n)}$ are presented in the form [30]:

$$x_s = \theta_0(t) + \theta_1(t)h_1(t, s) + \dots + \theta_m(t)h_m(t, s) + \varepsilon_s \tag{1}$$

for $\max(1, t - k) \leq s \leq \min(t + k, n)$, where $\{\varepsilon_i\}_{\max(1,t-k) \leq i \leq \min(t+k,n)}$ is a sequence of independent random variables with normal distribution $N(0, \sigma^2)$. The quantities $h_j(t, s)$, $1 \leq j \leq m$ represent transformations of the time variable around the moment t . If the deterministic part in (1) is approximated using a second degree polynomial ($m = 2$), we assume $h_j(t, s) = (t - s)^j$ for $1 \leq j \leq 2$. The relationship (1) between the endogenous variable and the predictors (transformations h_i , $1 \leq i \leq m$) is linear. For each moment t z the relationship (1) can be presented in the form:

$$X = Z_t \theta(t) + \varepsilon \tag{2}$$

where:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad Z_t = \begin{bmatrix} 1 & h_1(t, 1) & h_2(t, 1) & \dots & h_m(t, 1) \\ 1 & h_1(t, 2) & h_2(t, 2) & \dots & h_m(t, 2) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & h_1(t, n) & h_2(t, n) & \dots & h_m(t, n) \end{bmatrix}, \quad \theta(t) = \begin{bmatrix} \theta_0(t) \\ \theta_1(t) \\ \vdots \\ \theta_m(t) \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

The vector of structural parameters $\theta(t)$ is determined using the least squares method. To this purpose, for each moment t , $1 \leq t \leq n$ we solve the problem:

$$\min_{\theta(t)} (X - Z_t \theta(t))^T W_t (X - Z_t \theta(t)), \tag{3}$$

while the weight matrix has the form:

$$W_t = \begin{bmatrix} w(t, 1) & 0 & 0 & \dots & 0 \\ 0 & w(t, 2) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & w(t, n) \end{bmatrix} \tag{4}$$

In the analysed case, Gaussian weight was used:

$$w(t, s) = \exp\left(-\frac{\|t-s\|^2}{2\alpha^2}\right) \tag{5}$$

Solving problem (3) for each t , $1 \leq t \leq n$ by applying Gauss's theorem [28], we obtain the values of unknown parameters in Equation 1 as:

$$\hat{\theta}(t) = (Z_t^T W_t Z_t)^{-1} Z_t^T W_t X \tag{6}$$

Remark 1 For each moment t , $1 \leq t \leq n$ the sequence $\{\hat{\theta}_0(t)\}_{1 \leq t \leq n}$ defines the expected values of the series $\{x_t\}_{1 \leq t \leq n}$ of the form (2). For moments $\tau > n$ we estimate the forecast using the Equation:

$$\hat{x}_\tau = \hat{\theta}_0(n) + \hat{\theta}_1(n)h_1(n, \tau) + \dots + \hat{\theta}_m(n)h_m(n, \tau) \tag{7}$$

Remark 2 The sequence of differences $\{\eta_t\}_{1 \leq t \leq n}$, where: $\eta_t = x_t - \hat{\theta}_0(t)$ for $1 \leq t \leq n$ may contain non-random components (seasonality, integration, ARIMA series). When using the LOESS technique for a sequence of differences, we do not obtain positive effects, but the use of other transformations for model (2) allows us to identify, for example, seasonality.

Harmonic analysis

The next component is the seasonal component [31], which can be identified using phase trends, the Holt-Winters method [32, 33], or harmonic analysis [31]. However, these techniques assume that the seasonality pattern is constant. In fact, in many cases, elements of the seasonal component evolve over time. The phase trends method allows for the identification of the dynamics of phases that are part of seasonality, while the Holt-Winters method allows for the identification of the seasonality component using exponential smoothing. In this way, we obtain a seasonality component that mimics the fluctuations of recent periods. Harmonic analysis assumes that the time series is a sum of trigonometric functions with different frequencies [31]. Using Fourier expansion, we identify harmonics, where it is assumed that the amplitudes of the harmonics do not change in time. Below, the method of identifying seasonality proposed in the article will be presented, based on harmonic analysis, taking into account the dynamics of changes in harmonics occurring in the series. We are considering a series from which various types of trend functions, polynomial functions, etc. have been eliminated, while periodic fluctuations still occur in it. In order to identify the time series $\{\eta_t\}_{t \in \mathbb{N}}$ we first create a Fourier series. The analysed time series is presented in the form:

$$\eta_t = \alpha_0 + \sum_{k=1}^T \left(\alpha_k \cos\left(\frac{2\pi}{n} kt\right) + \beta_k \sin\left(\frac{2\pi}{n} kt\right) \right) + \epsilon_t \tag{8}$$

where: $\{\epsilon_t\}_{1 \leq t \leq n}$ is a sequence of independent random variables with a normal distribution $N(0, \sigma^2)$. Based on the realisation of sequence $\{\eta_t\}_{1 \leq t \leq n}$ we identify the series using the expansion of (8) and assume that the number of possible harmonics occurring in this series $T \ll \left\lfloor \frac{n}{2} \right\rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part. The expansion (8) shows that seasonality can be represented as the sum of harmonic components. The estimators of the unknown parameters $\alpha_0, \alpha_1, \beta_1, \dots, \alpha_T, \beta_T$ in the Equation 8 are determined using the least squares method and have the form:

$$\hat{\alpha}_0 = \frac{1}{n} \sum_{t=1}^n \eta_t \tag{9}$$

$$\hat{\alpha}_k = \frac{2}{n} \sum_{t=1}^n \eta_t \cos\left(\frac{2\pi}{n} kt\right) \tag{10}$$

$$\hat{\beta}_k = \frac{2}{n} \sum_{t=1}^n \eta_t \sin\left(\frac{2\pi}{n} kt\right) \tag{11}$$

for $k = 1, 2, \dots, T$. If a trend is separated from series (1), the sequence $\{\eta_t\}_{1 \leq t \leq n}$ oscillates around the zero level, therefore $\hat{\alpha}_0 \approx 0$. We estimate the sequence of residuals $\{\epsilon_t\}_{1 \leq t \leq n}$ as:

$$\epsilon_t = \eta_t - \hat{\alpha}_0 - \sum_{k=1}^T \left(\hat{\alpha}_k \cos\left(\frac{2\pi}{n} kt\right) + \hat{\beta}_k \sin\left(\frac{2\pi}{n} kt\right) \right) \tag{12}$$

The standard deviation of the sequence of residuals $\{\epsilon_t\}_{1 \leq t \leq n}$ is equal to

$$\sigma_\epsilon = \sqrt{\frac{1}{n} \sum_{t=1}^n \eta_t^2 - \hat{\alpha}_0^2 - \frac{1}{2} \sum_{k=1}^T (\hat{\alpha}_k^2 + \hat{\beta}_k^2)} \tag{13}$$

For the k -th harmonic, the quantity: $\omega_k = \frac{2\pi}{n} k$ means the angular velocity, while $A_k = \sqrt{\alpha_k^2 + \beta_k^2}$ an amplitude. By analysing the sequence of amplitudes $\{A_k\}_{1 \leq k \leq T}$ it is possible to determine harmonics that have significantly impact on seasonality. As the set of indices of the harmonics occurring in the time series we select:

$$H = \{i: A_i \geq q_\alpha, 1 \leq i \leq T\} \tag{14}$$

while q_α means quantile of the order α , $0 < \alpha < 1$ for the sequence of amplitudes $\{A_k\}_{1 \leq k \leq T}$. We determine the forecast values related to seasonality as:

$$\hat{\eta}_t = \hat{\alpha}_0 + \sum_{k \in H} \left(\hat{\alpha}_k \cos(\omega_k t) + \hat{\beta}_k \sin(\omega_k t) \right) \tag{15}$$

Application of LOESS to identify seasonality

The method of identifying the seasonal component using LOESS is called LESC (locally estimated seasonal components). In the case under consideration, we first identify the harmonic components occurring in the time series (i.e. we determine the set H given by Equation 14. Then, using the WLS method, we determine the coefficients of the harmonic components for each moment. We analyse the time series $\{\eta_s\}_{1 \leq s \leq n}$ containing a seasonal component around the moment t , $1 \leq t \leq n$. The structural parameters in model (8) depend on the moment t . Let us assume $m \in \mathbb{N}$ the number of harmonics determining the seasonality, i.e. $H = \{k_1, k_2, \dots, k_m\}$, therefore the elements of this series are presented in the form:

$$\eta_s = \alpha_0(t) + \alpha_1(t)h_1^c(s) + \dots + \alpha_m(t)h_m^c(s) + \beta_1(t)h_1^s(s) + \dots + \beta_m(t)h_m^s(s) + \epsilon_s \tag{16}$$

where:

$h_i^c(s) = \cos\left(\frac{2\pi}{n} k_i s\right)$ and $h_i^s(s) = \sin\left(\frac{2\pi}{n} k_i s\right)$ for $\max(1, t - k) \leq s \leq \min(t + k, n)$ and $1 \leq i \leq m$. We can represent relationship (16) using Equation 2, where:

$$X = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}, \quad Z_t = \begin{bmatrix} 1 & h_1^c(1) & \dots & h_m^c(1) & h_1^s(1) & \dots & h_m^s(1) \\ 1 & h_1^c(2) & \dots & h_m^c(2) & h_1^s(2) & \dots & h_m^s(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & h_1^c(n) & \dots & h_m^c(n) & h_1^s(n) & \dots & h_m^s(n) \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}.$$

while the vector of structural parameters has the form:

$$\theta(t) = [\alpha_0(t), \alpha_1(t), \dots, \alpha_m(t), \beta_1(t), \dots, \beta_m(t)]^T$$

Then, solving problem (3) taking into account the weight matrix (4), we obtain estimates of the structural parameters $\theta(t)$ for $1 \leq t \leq n$ using Equation 6.

Remark 3 For each moment t , $1 \leq t \leq n$ the sequence $\{\hat{\eta}_t^{LOESS}\}_{1 \leq t \leq n}$:

$$\hat{\eta}_t^{LOESS} = \hat{\alpha}_0(t) + \hat{\alpha}_1(t)h_1^c(t) + \dots + \hat{\alpha}_m(t)h_m^c(t) + \hat{\beta}_1(t)h_1^s(t) + \dots + \hat{\beta}_m(t)h_m^s(t) \tag{17}$$

determines the expected values of the series (8). The sequence $\{\epsilon_t\}_{1 \leq t \leq n}$, where $\epsilon_t = \eta_t - \hat{\eta}_t^{LOESS}$ for $1 \leq t \leq n$, represents the differences between the empirical values of the series (8) and the expected values obtained using LOESS. For the moments $\tau > n$ we estimate the forecast using the Equation:

$$\hat{\eta}_\tau^{LOESS} = \hat{\alpha}_0(n) + \hat{\alpha}_1(n)h_1^c(\tau) + \dots + \hat{\alpha}_m(n)h_m^c(\tau) + \hat{\beta}_1(n)h_1^s(\tau) + \dots + \hat{\beta}_m(n)h_m^s(\tau) \quad (18)$$

Accuracy analysis metrics

The models presented above have been fitted to goods (freight mass) and passengers data and the forecasts have been estimated. The following metrics were used to compare the quality of model fit and the forecasting accuracy [28, 31]. Root Mean Square Error (*RMSE*) means the root of mean of squared differences between the actual and model predicted values and is given by the Equation:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (19)$$

where: x_i is the actual value of i -th observation and \hat{x}_i is the predicted value, $1 \leq i \leq n$.

The *MAE* (Mean Absolute Error) denotes the average absolute errors between actual values and predicted ones and is estimated as follows

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (20)$$

The *MAPE* (Mean Absolute Percentage Error) is a measure that calculates the average absolute percentage difference between actual and expected values. It is often used to quantify the prediction error as a percentage and is determined as follows

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left\| \frac{x_i - \hat{x}_i}{x_i} \right\| \quad (21)$$

Smaller values of presented above metrics mean superior model performance.

RAIL TRANSPORT ANALYSIS

The article analyses rail transport regarding both goods (freight mass) and passengers (number of travellers). The data comes from <https://utk.gov.pl/> and concerns the years 2012–2023 [38]. Additionally, a short-term prediction was made for the period February–April 2023 and the obtained results were compared with the actual values. Identification and prediction were performed for the LET, LETSC and LETLESC models.

Transport of goods by rail

Freight transport by rail mainly concerns coal, crude oil and its products, natural gas, steel and metal products, cement, grain, etc. These products are necessary for the proper functioning of the economy, so countries place particular emphasis on maintaining supplies with the least disruption possible. Such efforts are also made during crises, but it is not possible to completely eliminate disturbing factors, as shown in Fig. 1 – black line – showing the transported mass of goods in the period under study. The use of LOESS allowed us to estimate the trend occurring in the time series. On this basis, predictions were made for the period February–April 2023 (training set).

In Figure 1, the matching of the locally estimated trend $\{\hat{\theta}_0(t)\}_{1 \leq t \leq n}$ from the Equation 6 for the period from January 2012 to January 2023 is marked with a solid dark blue line, while the prediction $\{\hat{\theta}_0(\tau)\}_{\tau > n}$ estimated from Equation 7 is marked with a dashed line for the period February–April 2023.

The next figure (Fig. 2) shows the analysis of the seasonality of the residuals. Analysing the polar chart (left graph), a clear increase in transport is visible in March and October, while a decrease in January, February and December, which confirms the existence of seasonality in the time series. Additionally, the right graph shows that for the 4 harmonic the amplitude values are above the 0.95 quantile (see Equation 14). Therefore, 4 harmonic components were selected for further analysis, which to the greatest extent explains the behaviour of seasonality in a series of residuals. Corresponding to greatest amplitude values the angular velocities for these harmonics (red points) are marked in Figure 2.

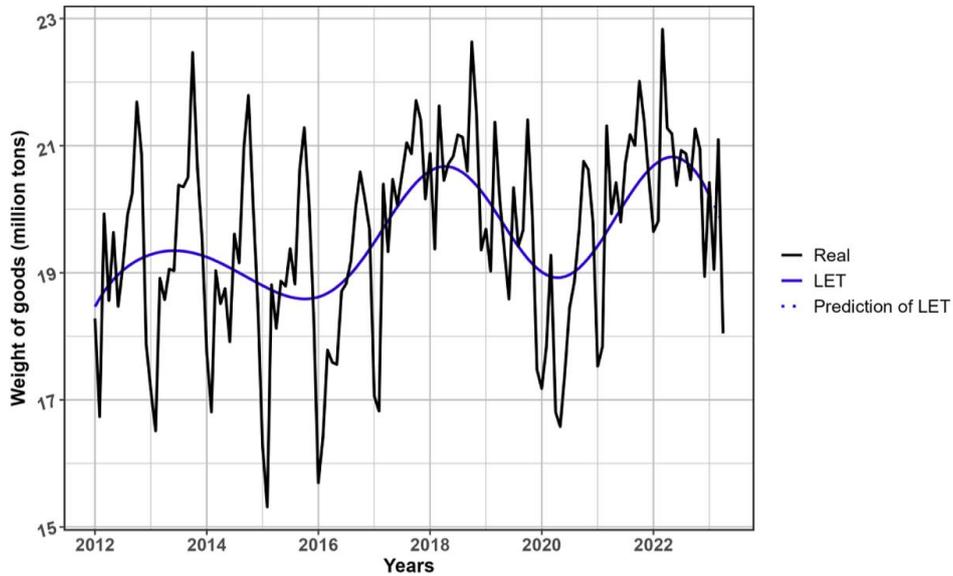


Figure 1. Determining the trend using LOESS

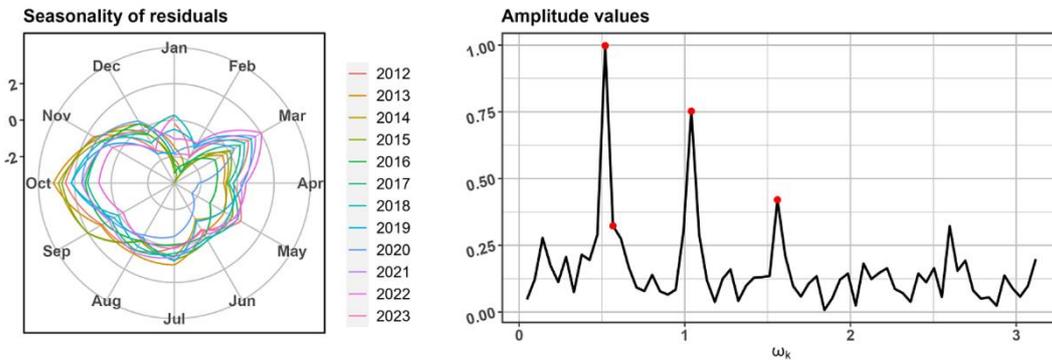


Figure 2. Seasonality analysis of residuals

In Fig. 3, in addition to the results from Fig. 1, the composition of the locally estimated trend and the seasonal component (LETSC) $\{\hat{\theta}_0(t) + \hat{\eta}_t\}_{1 \leq t \leq n}$ was plotted (see Equations 6 and 15), and the red dashed line shows the prediction $\{\hat{\theta}_0(\tau) + \hat{\eta}_\tau\}_{\tau > n}$ in the period February–April 2023. Additionally, the solid green line shows the composition of the locally estimated trend and the locally estimated seasonal component (LETLESC) $\{\hat{\theta}_0(t) + \hat{\eta}_t^{LOESS}\}_{1 \leq t \leq n}$ determined using Equations 7 and 17, while the green dashed line marks the prediction $\{\hat{\theta}_0(\tau) + \hat{\eta}_\tau^{LOESS}\}_{\tau > n}$ in the period February–April 2023 estimated using (7) and (18).

Next, the proposed models were assessed. Table 1 presents the values RMSE, MAE, MAPE for all three models. The LET method takes into account only the local trend in the series, while the LETSC and LETLESC methods additionally take into account the seasonality factor, therefore the error values for LETSC and LETLESC are much smaller. The method that additionally takes into account the seasonality locally estimated seasonal components (LETLESC) most accurately reflects the behaviour of the mass of transported goods. For LETLESC, the root mean square error is 0.5027 (million tons), the mean absolute error is 0.5421 (million tons), and the mean percentage error is 2.83%. Therefore, increasing the complexity of the model and taking into account additional factors, the quality of the matching improves. This is especially important in series that experience dynamic and rapid changes.

Additionally, error values were calculated for the forecasts allow us to evaluate predictive abilities of presented methods. Comparing the RMSE, MAE, MAPE values presented in Table 2 for LET, LETSC and LETLESC methods we can conclude that a more accurate prediction for the period February–April 2023 was also obtained using LETLESC, the mean square error is 2.0633 (million tons), the mean absolute error is 1.0918 (million tons), while the mean percentage error 5.89%.

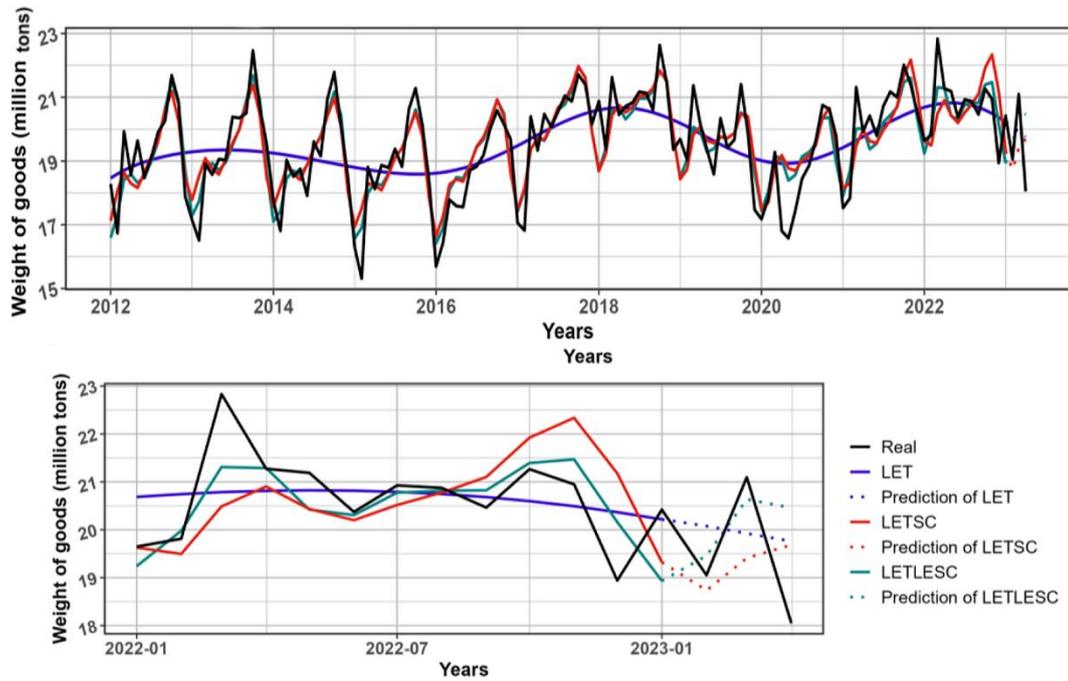


Figure 3. Matching and prediction obtained using LET, LETSC and LETLESC

Table 1. Trend matching and seasonality analysis using LET, LETSC and LETLESC

Method	RMSE	MAE	MAPE
LET	1.6031	0.9812	0.0515
LETSC	0.6809	0.6231	0.0326
LETLESC	0.5027	0.5421	0.0283

Table 2. Forecast analysis using LET, LETSC and LETLESC

Method	RMSE	MAE	MAPE
LET	1.7837	1.3025	0.0680
LETSC	1.8744	1.2139	0.0625
LETLESC	2.0633	1.0918	0.0589

Therefore, methods that took into account the seasonality component gave better forecasts than the LET method. The values of the RMSE, MAE and MAPE indices in the case of forecasts are higher than in the case of matching, but it should be emphasized that the only factor analysed in this case is time. The proposed models do not take into account additional factors (e.g. geopolitical) that may affect the weight of the transported goods.

Carriage of passengers by rail

Rail passenger transport has been more severely disrupted due to the Covid-19 pandemic. Restrictions on movement, remote work, restrictions on the tourism industry and, above all, fear of possible infection have significantly influenced the decline in the number of people using this form of travel. This is clearly visible in Fig. 4, where the black line shows the number of passengers using rail transport. A characteristic feature is a decrease in the number of passengers during the lockdown period and an increase at the beginning of the war with Ukraine. The use of LOESS allowed to estimate the local trend occurring in the time series. In Figure 4, similarly to the case of freight transport, the locally estimated trend (solid line) and prediction (dashed line) are shown in dark blue.

The next figure shows the analysis of the seasonality of the residuals. There is a slight decrease in transports in February and December and an increase in October, but no clear differences are visible on the polar chart.

In accordance with the adopted methodology, four harmonic components were selected again, the amplitudes of which are above the 0.95 quantile, assuming that these harmonic components explain the seasonality behaviour to the greatest extent. The angular velocities for these harmonics corresponding to greatest amplitude values are shown in Figure 5.

In Figure 6, similarly to the freight transport results, the empirical data is marked in black and the matches are plotted with solid lines in individual colours: LET $\{\hat{\theta}_0(t)\}_{1 \leq t \leq n}$ (dark blue), LETSC

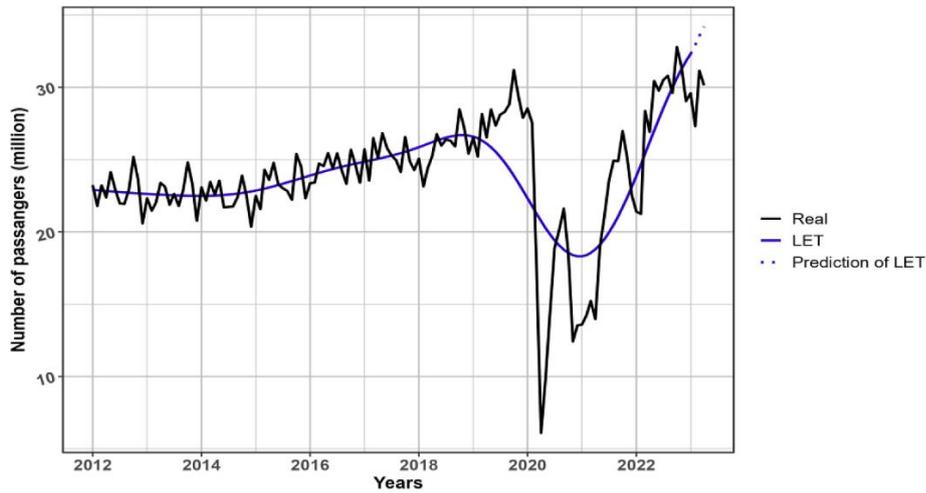


Figure 4. Determining the trend using LOESS

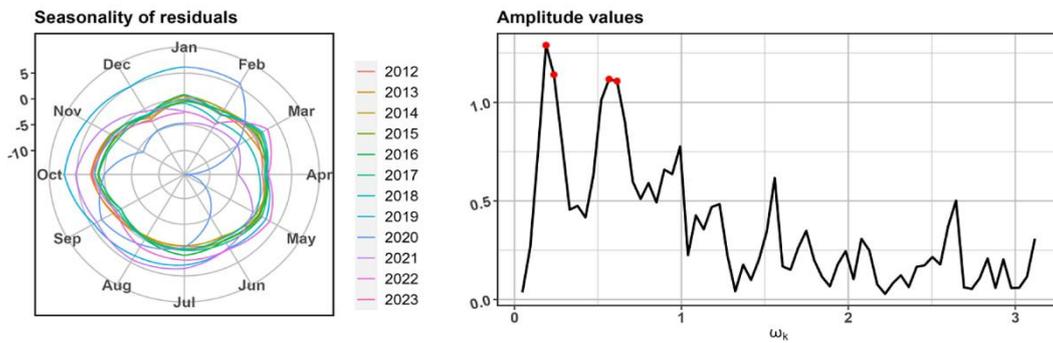


Figure 5. Seasonality analysis of residuals

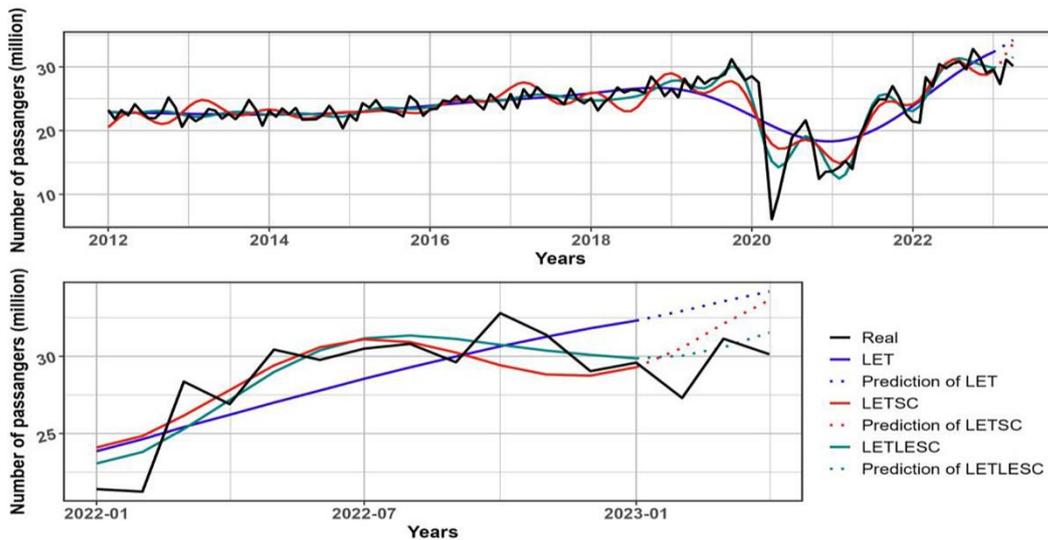


Figure 6. Matching and prediction obtained using LET, LETSC and LETLESC

$\{\hat{\theta}_0(t) + \hat{\eta}_t\}_{1 \leq t \leq n}$ (red) and LETLESC $\{\hat{\theta}_0(t) + \hat{\eta}_t^{LOESS}\}_{1 \leq t \leq n}$ (green). Similarly, predictions are marked with dashed lines in appropriate colours (for the period from February to April 2023): LET $\{\hat{\theta}_0(\tau)\}_{\tau > n}$ (dark blue), LETSC $\{\hat{\theta}_0(\tau) + \hat{\eta}_\tau\}_{\tau > n}$ (red) and LETLESC $\{\hat{\theta}_0(\tau) + \hat{\eta}_\tau^{LOESS}\}_{\tau > n}$ (green).

Table 3. Trend matching and seasonality analysis using LET, LETSC and LETLESC

Method	RMSE	MAE	MAPE
LET	7.5016	1.7858	0.0975
LETSC	4.7799	1.5270	0.0802
LETLESC	2.5937	1.0995	0.0583

The models were again evaluated using RMSE, MAE, MAPE presented in the Table 3. Since the LET method only takes into account the local trend, the RMSE, MAE and MAPE values are much higher. The LETSC and LETLESC methods additionally take into account the seasonality factor, so the indices values are lower.

The above analysis clearly shows that the LETLESC method most accurately reflects the behaviour of rail passenger numbers. Even with such drastic disruptions (the introduction of a lockdown and the start of the war in Ukraine), the root mean square error is 2.5937 (million people), the mean absolute error is 1.0995 (million people), and the mean percentage error is 5.83%. Prediction indices for models concerning rail passengers are presented in Table 4. Consistently the best results are obtained with the LETLESC model. The mean square error is 3.2529 (million people), the average absolute error is 1.5662 (million people), and the average percentage error is 5.49%.

Taking into account the trend and seasonal changes allowed for more accurate passenger transport forecasts. Analysing Figure 8 and the RMSA, MAE, MAPE indices, it is visible that using the classic LET method, which does not take into account the seasonality component, the prediction of the number of passengers is much weaker than using the LETSC and LETLESC methods. Among the methods that take into account the seasonality component, better forecasts were obtained for LETLESC.

The authors' intention was to obtain a predictive model that would describe the changes occurring over time as accurately as possible, taking into account not only the trend, but also seasonality. What is particularly important is not only the consideration of the seasonality component, but also the possibility of adapting this component, i.e. its local adjustment. This allowed for accurate capture of the changes taking place and allowed to obtain better results in both, matching to empirical data and forecasts, as shown in Figure 4 and Figure 8 and the results in Tables 1–4.

Table 4. Forecast analysis using LET, LETSC and LETL

Method	RMSE	MAE	MAPE
LET	17.9998	4.0363	0.1395
LETSC	7.9181	2.5746	0.0888
LETLESC	3.2529	1.5662	0.0549

CONCLUSIONS

Time series characterized by significant dynamics and rapid changes occurring over time are difficult to identify and forecast. Therefore, the authors of this article proposed a method that is based on local fitting and adaptation of a given component of the series to the current observation. This approach was applied to both the trend and the seasonal component. To demonstrate the accuracy of the adopted assumptions, three modelling techniques were used and compared, i.e. locally estimated trend (LET), locally estimated trend with seasonal components (LETSC) and locally estimated trend with locally estimated seasonal components (LETLESC). For each method, identification was made to the training data and the matching was evaluated, and then a prediction was determined and its compliance with the test set was assessed. Such research was carried out in relation to freight and passenger transport by rail. In each case, it turned out that better results of matching to the test set and a more accurate short-term forecast were obtained for the model that took into account both the locally estimated trend and the LETLESC. Thus, the authors' objective, which was to develop a method for identifying the components of a time series disrupted by crisis events, has been achieved. Local fitting methods have been shown to perform well on such empirical data and provide accurate predictions. The predictive model obtained describes well the changes occurring over time, taking into account not only the trend, but also seasonality. Therefore, the proposed method of local trend and seasonality fitting gives better forecasting results. The main advantage of the method ensuring effective adaptation to changing conditions was taking into account the effect of weights for empirical data, which – as already mentioned – relied on assigning greater weight to observations closer to the analysed moment, and less weight to observations further away from that moment. This technique is the answer to the problem of

identifying time series in which external factors cause pattern changes.

The method was presented using the data regarding rail transport in Poland as an example, but it is universal and can be used for similar time series with high dynamics. It is also worth emphasizing that, in accordance with the authors' intention, the presented concept only concerns the influence of the time factor on changes in the time series. Although the model is adapted in parallel to the data, which allows for taking into account changes in model parameters over time in order to construct more accurate forecasts, other factors (e.g. geopolitical changes, economic situation, etc.) were not taken into account. Therefore, future research will develop models to also take into account additional variables.

REFERENCES

1. Sudan T., Taggar, R. Recovering supply chain disruptions in post-COVID-19 pandemic through transport intelligence and logistics systems: India's experiences and policy options. *Frontiers in Future Transportation* 2021; 2: 660116.
2. Andrych-Zalewska M., Chlopek Z., Pielecha J., Merkisz J. Investigation of exhaust emissions from the gasoline engine of a light duty vehicle in the Real Driving Emissions test. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2023; 25(2): 165880.
3. Murawski J., Szczepański E., Jacyna-Gołda I., Izdebski M., and Jankowska-Karpa D. Intelligent mobility: A model for assessing the safety of children traveling to school on a school bus with the use of intelligent bus stops. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2022; 24(4): 695–706.
4. Oszczypała M., Ziółkowski J., Małachowski J. Semi-Markov approach for reliability modelling of light utility vehicles. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2023; 25(2): 161859.
5. Kukulski J., Lewczuk K., Góra I., Wasiak M. Methodological aspects of risk mapping in multimode transport systems. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2023; 25(1): 19.
6. Szaciłło L., Jacyna M., Szczepański E., Izdebski M. Risk assessment for rail freight transport operations. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2021; 23(3): 476–488.
7. Zhang Y., Zhao M. An integrated approach to estimate storage reliability with masked data from series system. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2023; 25(4): 172922.
8. Zhu G., Chou M.C., Tsai C.W. Lessons learned from the COVID-19 pandemic exposing the shortcomings of current supply chain operations: A long-term prescriptive offering. *Sustainability* 2020; 12(14): 5858.
9. Ciuffini F., Tengattini S., Bigazzi A.Y. Mitigating increased driving after the COVID-19 pandemic: an analysis on mode share, travel demand, and public transport capacity. *Transportation Research Record: Journal of the Transportation Research Board* 2023, 2677(4): 154–167.
10. Merk O., Hoffmann J., Haralambides H. Post-COVID-19 scenarios for the governance of maritime transport and ports. *Maritime Economics & Logistic* 2022; 24(4): 673–685.
11. Eisenmann C., Nobis, C., Kolarova V., Lenz B., Winkler C. Transport mode use during the COVID-19 lockdown period in Germany: The car became more important, public transport lost ground. *Transport Policy* 2021; 103: 60–67.
12. Shamshiripour A., Rahimi E., Shabanpour R., Mohammadian A.K. How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago. *Transportation Research Interdisciplinary Perspective* 2020; 7: 100216.
13. Grechi D., Ceron M. COVID-19 lightening the load factor in railway transport: Performance analysis in the north-west area of Milan. *Research in Transportation Business & Management* 2022; 43: 100739.
14. Lucchesi S.T., Tavares V.B., Rocha M.K. Larranaga A.M., Public Transport COVID-19-Safe: New Barriers and Policies to Implement Effective Countermeasures under User's Safety Perspective. *Sustainability* 2022; 14(5): 2945.
15. Kozłowski E., Borucka A., Oleszczuk P., Jałowicz T. Evaluation of the maintenance system readiness using the semi-Markov model taking into account hidden factors. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2023; 25(4): 172857.
16. Severo M., Ribeiro A.I., Lucas R., Leão T., Barros H. Urban rail transportation and SARS-Cov-2 infections: an ecological study in the Lisbon metropolitan area. *Frontiers in Public Health* 2021; 9: 611565.
17. Lu J., Lin A., Jiang C., Zhang A., Yang Z. Influence of transportation network on transmission heterogeneity of COVID-19 in China. *Transportation Research Part C: Emerging Technologies* 2021; 129: 103231.
18. Burdzik R., Speybroeck N. Study on the estimation of Sars-Cov-2 virus pathogens' transmission probabilities for different public bus transport service scenarios. *Transport Problems* 2023; 18(3): 199–211.
19. Liu R., Li D., Kaewunruen, S. Role of Railway Transportation in the Spread of the Coronavirus: Evidence From Wuhan-Beijing Railway Corridor. *Frontiers in Built Environment* 2020; 6: 590146.
20. Wang Y., Fang Z., Gao W. Covid-19's impact on China's economy: a prediction model based on railway transportation statistics. *Disasters* 2021; 45: 76–96.

21. Nguyen H.K. Application of mathematical models to assess the impact of the COVID-19 pandemic on logistics businesses and recovery solutions for sustainable development. *Mathematics* 2021; 9(16): 1977.
22. Arai Y., Niwa Y., Kusakabe T., Honma K. How has the Covid-19 pandemic affected wheelchair users? Time-series analysis of the number of railway passengers in Tokyo. *Humanities and Social Sciences Communications* 2023; 490.
23. Saxena A., Yadav A.K. Examining the Effect of COVID-19 on rail freight volume and revenue using the ARIMA forecasting model and assessing the resilience of Indian railways during the pandemic. *Innovative Infrastructure Solutions* 2022; 7: 348.
24. Kallidoni M., Katrakazas C., Yannis G. Modelling the relationship between covid-19 restrictive measures and mobility patterns across Europe using time-series analysis *European Journal of Transport and Infrastructure Research* 2022; 22(2): 161–182.
25. Liu Q., Huang Z. Research on intelligent prevention and control of COVID-19 in China's urban rail transit based on artificial intelligence and big data. *Journal of Intelligent & Fuzzy Systems* 2020; 39(6): 9085–9090.
26. Asad S.M., Dashtipour K., Rais R.N.B., Hussain S., Abbasi Q.H., Imran M.A. Employing machine learning for predicting transportation modes under the COVID-19 pandemic: A mobility-trends analysis. In: *International Conference on UK-China Emerging Technologies (UCET)*, Chengdu, China 2021; 235–240.
27. Rodríguez-Pérez R., Bajorath J. Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. *Journal of Computer- Aided Molecular Design* 2020; 34: 1013–1026.
28. Wooldridge M. *Introductory Econometrics: A Modern Approach*. Cengage learning, 2015.
29. Li K., Ren L., Li X., Wang Z. Remaining useful life prediction of equipment considering dynamic thresholds under the influence of maintenance. *Eksplatacja i Niezawodność – Maintenance and Reliability* 2024; 26(1): 174903.
30. Cleveland W.S., Susan J.D. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of American Statistical Association* 1988; 83(403): 596–610.
31. Hamilton J.D. *Time Series Analysis*. Princeton University Press, 2020.
32. Borucka A, Kozłowski E, Oleszczuk P, Świdorski A. Predictive analysis of the impact of the time of day on road accidents in Poland, *Open Engineering* 2020; 11(1): 142–150.
33. Holt C. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting* 2004; 20(1): 5–10.
34. Burdzik R. *Epidemic Risk Analysis and Assessment in Transport Services: COVID-19 and Other Viruses*. CRC Press, 2021.
35. Burdzik R. Probability of transmission of SARS-CoV-2 virus pathogens in long-distance passenger transport. *Archives of Transport* 2023; 68(4): 21–39.
36. Jaroń A. Borucka, A., Parczewski R. Analysis of the impact of the COVID-19 pandemic on the value of CO₂ emissions from electricity generation. *Energies* 2022; 15: 4514.
37. Burdzik R. An Application of the DHI Methodology for a Comparison of SARS-CoV-2 Epidemic Hazards in Customer Delivery Services of Smart Cities. *Smart Cities* 2023; 6: 965–986.
38. Office of Rail Transport (UTK), railway data 2012–2023, UTK statistical portal, <https://utk.gov.pl/>
39. Zhang C., Zhang Y., Dui H., Wang S., Tomovic M. Component maintenance strategies and risk analysis for random shock effects considering maintenance costs. *Eksplatacja i Niezawodność – Maintenance and Reliability* 2023; 25(2): 162011.
40. Niewczas A., Móravski Ł., Rymarz J., Dębicka E., Hołyszko P. Operational risk assessment model for city buses. *Eksplatacja i Niezawodność – Maintenance and Reliability* 2023; 25(1): 14.
41. Benmalek E., Elmhamdi J., Jilbab A., Jbari A. A cough-based Covid-19 detection with gammatone and mel-frequency cepstral coefficients. *Diagnostyka* 2023; 24(2): 2023214.
42. Bensid K., Lati A., Benlamoudi A., Ghoul B.E., Senoussi M.L. Efficient Covid-19 disease diagnosis based on cough signal processing and supervised machine learning. *Diagnostyka* 2023; 24(1): 2023103.
43. Rybak G., Kozłowski E., Król K., Rymarczyk T., Sulimierska A., Dmowski A., Bednarczyk, P. Algorithms for optimizing energy consumption for fermentation processes in biogas Production. *Energies* 2023; 16: 7972.
44. Pluciński M., Kotowska I., Mańkowska M., Filina-Dawidowicz L. Research on diversification strategies of terminal operators – evidence from Polish Seaports. *Sustainability* 2023; 15: 5644.