ISSN 1648-4142 / eISSN 1648-3480

TRANSPORT

2023
Volume 38
Issue 4
Pages 231–242
https://doi.org/10.3846/transport.2023.20932

Original Article

FREIGHT RATE AND DEMAND FORECASTING IN ROAD FREIGHT TRANSPORTATION USING ECONOMETRIC AND ARTIFICIAL INTELLIGENCE METHODS

Edvardas LIACHOVIČIUS[1,2], Eldar ŠABANOVIČ[2], Viktor SKRICKIJ[2]✉

[1] *JSC "Girteka Logistics", Vilnius, Lithuania*
[2] *Transport and Logistics Competence Centre, Vilnius Gediminas Technical University, Vilnius, Lithuania*

Highlights:

- this article unveils cutting-edge forecasting techniques;
- a comparative analysis of econometric models is introduced for forecasting freight rates and demand;
- novel forecasting methodology for the freight transportation industry is showcased, leveraging ANN-based models;
- an innovative approach, based on the correlation between spot and contract rates, is introduced and thoroughly examined.

Article History:

- submitted 13 September 2023
- resubmitted 29 November 2023;
- accepted 8 December 2023.

Abstract. The digitisation of the transportation sector and data availability have opened up new opportunities to implement data-driven methods for improving company performance. This article analyses demand and freight rate forecasting techniques in the context of the road freight transportation company. The European market was analysed in this research, and direction from the Netherlands to Italy was selected for the case study. Performed investigation showed that econometric models such as Auto-Regressive Integrated Moving Average (ARIMA) used for demand prognosis provide good results. Freight rate forecasting is different; econometric models, including multivariate models ARIMA with exogenous variables (ARIMAX) and Seasonal ARIMAX (SARIMAX), do not perform satisfactorily under specified time intervals, therefore MultiLayer Perceptron (MLP) was used as a solution. It can be seen that Artificial Intelligence (AI) based methods provide better results. Despite its success, the AI-based approach alone is not recommended for practical implementation since forecasted input parameters are necessary. Lastly, the study uncovers a valuable insight. A strong correlation (0.86) between spot and contract rates was found, and the article shows how current spot rates can be used for contract rate forecasting.

Keywords: transportation, road freight transport, freight rate forecasting, demand forecasting, econometric models, artificial neural networks.

✉Corresponding author. E-mail: *viktor.skrickij@vilniustech.lt*

Notations

Abbreviations:

ADF – augmented Dickey–Fuller;
AI – artificial intelligence;
AIC – Akaike information criterion;
AICc – AIC corrected;
ANN – artificial neural network;
AR – auto-regressive;
ARIMA – auto-regressive integrated moving average;
ARIMAX – ARIMA with exogenous variables;
EU – European Union;
MA – moving average;
MAE – mean absolute error;
MAPE – mean absolute percentage error;
MFE – mean forecast error;
MLP – multilayer perceptron;
MSE – mean squared error;
NYMEX – New York mercantile exchange;
RMSE – root MSE;
SARIMA – seasonal ARIMA;
SARIMAX – SARIMA with exogenous variables;
US – United States;
VAR – vector equilibrium correction;
WTI – West Texas intermediate.

Variables and functions:

b – bias vector;
B – backshift (lag) operator;
$\hat{B}$ – the maximum value of the likelihood function of the model;
c – intercept;

d – order of the differences of the data-integrating part $I(d)$;
D – the seasonal differencing term is equal to 1 or 2;
k – the number of estimated parameters;
n – the sample size;
p – order of the $AR(p)$ model;
q – order of $MA(q)$ mode;
t – time;
$\mathbf{W}$ – matrix of the learnable input connection weight;
x – input vector;
$X_{i,t}$ – explanatory variable i at time moment t;
y – model innovations (random error);
y_t – current time-series value;
$y_{t-1}, y_{t-2}, ..., y_{t-p}$ – past time-series values;
β_i – coefficient;
ε_T – the output vector;
$\Phi_P(B)$ – non-seasonal AR polynomial;
$\Phi_P(B^S)$ – seasonal AR polynomial;
ϕ_i – an AR model coefficient;
Θ_q – the MA polynomial;
$\Theta_q(B)$ – non-seasonal MA polynomial;
$\Theta_Q(B^S)$ – seasonal MA polynomial;
σ – nonlinear activation function.

Introduction

Road freight transport remains the dominant mode in the field of freight transportation (Liachovičius *et al.* 2020). Despite its central role, the sector faces numerous challenges, including environmental concerns, political changes, and increasing competition, necessitating the development of innovative solutions. The European Road Transport Research Advisory Council (ERTRAC 2021) anticipates a shift in truck technology from traditional combustion engines to alternatives such as liquefied natural gas, hydrogen, and electric power. Concurrently, the industry is experiencing a transformation with a trend towards platform-based business models (Ruggieri *et al.* 2018; Zhao *et al.* 2020; Tauscher, Kietzmann 2017), alongside a digitalisation that transforms companies (Fernández-Portillo *et al.* 2022; Truant *et al.* 2021).

The field of road freight transportation is characterised by its diversity, with small carriers making up the largest market share (TI 2024). Large companies often outsource their transportation needs and do not own their fleet, leading to 2 primary order types in the market: less than truckload and full truckload. The former requires the consolidation of multiple shipments to fill a truck, while the latter involves transporting a full load directly from the pickup to the delivery point. In this study, only a full truckload case is taken into account.

The industry operates with 2 main types of contracts: spot and long-term contracts. Spot contracts are used for individual shipments, while long-term contracts, also known as forward freight agreements, typically span 12 months and are used for multiple shipments. The balance between spot and long-term contracts varies by company, with each carrier weighing the freight rate (price) and overall demand for shipments (quantity of shipments) in their operations.

To effectively grow a business, it is crucial to possess reliable forecasting models that enable the negotiation of mutually beneficial contracts between the customer and the transportation company. If the transportation costs are excessively high, customers may opt for alternative service providers. Conversely, setting prices too low may lead to financial losses for the transportation company. Therefore, finding the right balance is essential for sustainable business development.

The forecasting models proposed in this article centre around a comparative analysis of econometric and ANN-based models. These models are chosen for their potential to provide nuanced forecasting in the context of the European road freight market, which has been underexplored in existing literature.

This article presents a threefold contribution. Firstly, it proposes a comparative analysis of econometric models for the forecasting of freight rates and demand. Secondly, the introduction of forecasting methodologies utilising ANN-based models specifically tailored for the freight transportation industry is presented. Thirdly, the article develops an additional forecasting model by providing data-driven insights based on the correlation between spot and contract rates.

The rest of the article consists of 4 sections. In Section 1, a literature review of forecasting methods and available case studies are presented. In Section 2, a description of mathematical models used in this study is presented together with metrics used for evaluations. In Section 3, the results are presented and analysed, and in last section the principal findings and their implications for the industry are concluded.

1. Literature review

Forecasting in road freight transportation involves the integration of historical data, economic insights, and an understanding of the demand to make informed decisions regarding freight rate, demand, inventory, budget planning, and market expansion. Based on the literature review classification of qualitative, quantitative, and hybrid forecasting methods that can be used for the identified task have been developed (Figure 1).

Qualitative methods encompass 2 groups of models. The 1st group comprises individual methods that can be performed by a single person. Scenario-building methods involve constructing narratives with plausible cause-and-effect links, connecting future conditions to the present while illustrating key decisions, events, and consequences (Glenn, Gordon 2009; Retek 2021). A simple review of available data sources, analysing the economic situation in the region, can be used for the forecast. The 2nd subgroup encompasses collective methods that require a team effort. The interview research methods involving interviewing experts and posing specific questions are widely used. Sur-

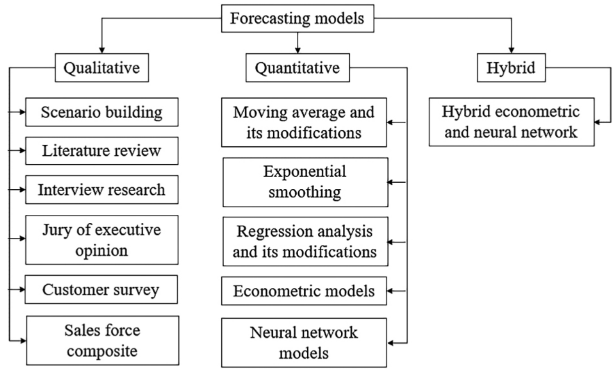

Figure 1. Forecasting models and methods that can be used for freight rate and demand forecasting

veys of customers are often used to qualitatively forecast as well (Claveria *et al.* 2017). One such survey method is the Delphi method, which involves multiple rounds of surveys where respondents can comment on questionnaires, and their answers can be adjusted after the presentation of the analysis result. Some convergence is observed after a few rounds (Kauko, Palmroos 2014). The sales force composite method involves sales managers forecasting sales in their respective territories, which are consolidated at the branch/region/area level to develop an overall company sales forecast (Mahlamäki *et al.* 2020). This method can be adapted for different areas by initially performing a low-level forecast and then aggregating the results at a higher level. It can be done using the jury of executive opinion method, which is another method (Wang, Chaovalitwongse 2011). Company executives performing in different areas come together and perform forecasting based on their experience, knowledge, and current situation in the market.

Quantitative methods are widely employed in forecasting. MA and its modifications, such as daily MA, normalised MA, and other MA-type indicators based on past rates for forecasting across various domains are widely used in forecasting (Wen *et al.* 2022). Simple exponential smoothing is suitable for situations without seasonal patterns or trend forecasting, where only the demand level needs to be determined (Dekker *et al.* 2004). Regression analysis-based methods can also be used for forecasting (Wang, Chaovalitwongse 2011), but econometric models are the most prevalent. The ARIMA model, incorporating AR, I, and MA components, is a classical and widespread algorithm suitable for stationary data. For seasonal data, SARIMA models are employed, and if forecasting data depends on several input parameters, multivariate models like ARIMAX or SARIMAX can be used.

Another group of quantitative methods is ANNs-based models. ANN is a data-driven method and learns process model from the data itself, therefore does not need a mathematical model (Goodfellow *et al.* 2016). ANNs as part of AI-based methods have gained prominence (Falatouri *et al.* 2022), often outperforming conventional approaches, as they can learn data-driven models connecting input and expected forecast data (LeCun *et al.* 2015). ANN-based models can also be updated with additional training on newly connected data without repeating the entire learning process (He *et al.* 2020). ANN-based models require a substantial amount of historical data, and not all the process data is available to researchers (Chen, Lin 2014).

In recent years the implementation of hybrid methods that combine econometric models, Kalman filtering, and AI methods has increased (Shukur, Lee 2015). Kalman filtering, being model-based, requires a mathematical model. A hybrid method based on combining the conventional econometric and ANN-based models can perform better than either of the models used separately (Ruiz-Aguilar *et al.* 2014). Simultaneously, the complexity of such solutions increases, and their practical application becomes more complex.

Vilutienė *et al.* (2014) defined requirements that forecasting models must correspond to, including: (1) models must be sufficiently simple to practically apply; (2) the observed phenomenon's tendencies in the future must remain the same as in the observation period; (3) the possibility of upgrading the models with new data; (4) the model must be multivariate; individual models can only reflect the possible process development accidentally; (5) the models must evaluate the stochastic nature of initial data; (6) the experience and intuition of the professionals are significant in evaluating the estimations. The last point could be adopted only for some of the qualitative methods.

There are several research articles dedicated to forecasting problems in freight transportation. Solutions appeared due to the digitisation of the sector and the appearance of databases. Batchelor *et al.* (2007) analysed the VAR, econometric and vector AR forecasting models. It was found that for spot rates and contracts, ARIMA performed better. Schramm & Munim (2021) used econometrics and VAR models for freight rate forecasting. They found out that the Logistics Confidence Index provided by Transport Intelligence in the ARIMAX model improves forecast performance greatly. However, this index is available for air and ocean transport modes, not road transport. Al Hajj Hassan *et al.* (2020) proposed using a reinforcement learning framework for freight demand forecasting. The margin of error in the long-term weekly forecasts is around 14% for long-term demand forecasting. The advantage of the proposed approach lies in its quick adaptability to recent events in the freight market and reasonable data requirements. Coyote Logistics (CL 2024) developed a freight rates forecasting model based on annual procurement, seasonal demand, and market capacity. Their multivariate model uses 5 indicators: truckload volume, industrial protection, imports, consumer spending and sales-to-inventory ratio. It satisfies the multivariate condition mentioned by Vilutienė *et al.* (2014). However, there are no details regarding the actual model, which is used for forecasting, analysis of forecasting accuracy is not performed either.

Some researchers analysed the dependence of spot rates on contracts and vice versa. In the early works dedicated to the shipping industry, Kavussanos *et al.* (2004b) show that contract prices 1 and 2 months prior to maturity are generally unbiased predictors of the realised spot prices. This exposes the forecasting problem of long-term contracts. Kavussanos *et al.* (2004a) argue that the freight rate defined in the contract has a general stabilising influence on the spot price volatility. Miller *et al.* (2021), found the dependence between spot price and contract. When aggregate spot rates increase by 20%, aggregate contract rates should be expected to increase by approximately 4% over the next few months, other things equal. Similar results are achieved by Coyote Logistics (CL 2020, 2024; Klujsza 2024) solution, which continuously provides truckload market forecasts with a 3 month forecast horizon. It should be mentioned that the majority of road freight transportation forecasting solutions (Miller *et al.* 2021; CL 2020, 2024; Klujsza 2024; Al Hajj Hassan *et al.* 2020), are performed for the US market. While in Europe, there are still data-related issues. First-of-all, there are no data aggregators that collect the data. Statistical data repositories provide data with great delay. For example, some input data used in this research appear with a delay of up to 9 months, while in the US it appears with a delay of less than 1 month.

Through the examination of existing literature, it becomes evident that econometric and ANN techniques are extensively employed for predictive purposes in transportation. However, a key limitation of the forecasting methods found in scientific literature is that they primarily focus on analysing past time intervals and known input indicators, disregarding the need to forecast these parameters, which is crucial in practical applications. Addressing data-related challenges is crucial for the European market and must be resolved to effectively implement the developed forecasting models in real-world scenarios.

This research article delves into the analysis of forecasting long-term freight rates and demand, taking into account spot prices and leveraging the advantages provided by digitalisation. A comprehensive forecasting methodology is meticulously developed and critically examined through the application of both econometric and ANN-based models.

2. Developing a mathematical methodology for freight rate and demand forecasting

In this section, the authors focus on mathematical models used for transportation demand and freight rates forecasting using internal company data as well as publicly available economic data. Internal data is actual freight demand and freight rates. External data used for freight rate forecasting are tonne-km, industrial production, imports, consumer spending, and fuel prices. It should be mentioned that there is no data aggregator in the EU for contract rates; some data can be found using Eurostat, Trading Economics, and Statista repositories.

During the models' development, a destination from the Netherlands to Italy was selected for the case study. For this destination shipment demand and contract freight rates since January 2015 were available and were used as input parameters for model training, with 78 months of data. It covers 6.5 years period prior to the middle of 2021. Data from months 79 to 90th (12 months) was used to analyse the accuracy of developed forecasting models. Month 90th is June 2022, as data for this period is already available, the forecasted values were compared with actual data directly.

To find a better solution all the possible combination of econometric models has been tested, including AR, MA, ARMA, ARIMA, SARIMA, ARIMAX and SARIMAX variations. Below the equations of econometric models used in this research are presented. All the calculations can be repeated using the provided equations, or using specialised software with econometric toolboxes.

The generalised equation of p order ***AR model*** (ArunKumar *et al.* 2021) is expresses as follows:

$$y_t = c + \phi_1 \cdot y_{t-1} + \phi_2 \cdot y_{t-2} + \ldots + \phi_p \cdot y_{t-p} + \varepsilon_t. \tag{1}$$

The generalised equation of q order ***MA model*** is expressed as follows:

$$y_t = c + \varepsilon_t + \Theta_1 \cdot \varepsilon_{t-1} + \Theta_2 \cdot \varepsilon_{t-2} + \ldots + \Theta_q \cdot \varepsilon_{t-q}. \tag{2}$$

For non-seasonal time series data ARIMA (p, d, q) model can be formulated using the formula from (Markevičiūtė *et al.* 2022):

$$y_t = c + \phi_1 \cdot y_{t-1} + \ldots + \phi_p \cdot y_{t-p} + \\ \varepsilon_t + \Theta_1 \cdot \varepsilon_{t-1} + \ldots + \Theta_q \cdot \varepsilon_{t-q}. \tag{3}$$

Also, the ***ARIMA model*** can be presented in a different form:

$$\Phi_p(B) \cdot (1-B)^d \cdot y_t = c + \Theta_q(B) \cdot \varepsilon_t. \tag{4}$$

In Equation (4) polynomial B is defined as the backshift operator, sometimes in literature it is called lag operator:

$$B^k \cdot y_t = y_{t-k}. \tag{5}$$

Non-seasonal AR polynomial in Equation (4) can be described as follows:

$$\Phi_p(B) = 1 - \phi_1(B) - \phi_2(B^2) - \ldots - \phi_p(B^p). \tag{6}$$

Non-seasonal MA polynomial in Equation (4) can be described as follows:

$$\Theta_q(B) = 1 + \Theta_1(B) + \Theta_2(B^2) + \ldots + \Theta_q(B^q). \tag{7}$$

In freight transportation, seasonality is often observed; such time series can be analysed by incorporating seasonal fluctuations in classical ARIMA models as parameters (p, d, q)^S describing seasonal lags. The ***SARIMA model*** can be expressed in a form (ArunKumar *et al.* 2021):

$$\Phi_P(B^S) \cdot \Phi_p(B) \cdot (1-B^S)^D \cdot (1-B)^d \cdot y_t = \\ \Theta_Q(B^S) \cdot \Theta_q(B) \cdot \varepsilon_t. \tag{8}$$

Seasonal terms for seasonal AR in Equation (8):

$$\Phi_P\left(B^S\right)=1-\Phi_1\left(B^S\right)-\Phi_2\left(B^{2\cdot S}\right)-\ldots-\Phi_P\left(B^{P\cdot S}\right). \quad (9)$$

Seasonal terms for seasonal MA in Equation (8):

$$\Theta_Q\left(B^s\right)=1+\Theta_1\left(B^S\right)+\Theta_2\left(B^{2S}\right)+\ldots+\Theta_Q\left(B^{Q\cdot S}\right). \quad (10)$$

In real-world cases one system parameter usually depends on others, using multivariate models the forecasting accuracy may be increased. ***ARIMAX model*** is described as follows:

$$\Phi_P\left(B\right)\cdot\left(1-B\right)^d\cdot y_t=c+\Theta_q\left(B\right)\cdot\varepsilon_t+\sum_{i=1}^{n}X_{i,t}\cdot\beta_i. \quad (11)$$

The ***SARIMAX model*** can be presented as follows:

$$\Phi_P\left(B^S\right)\cdot\Phi_P\left(B\right)\cdot\left(1-B^S\right)^D\cdot\left(1-B\right)^d\cdot y_t=$$
$$\Theta_Q\left(B^s\right)\cdot\Theta_q\left(B\right)\cdot\varepsilon_t+\sum_{i=1}^{n}X_{i,t}\cdot\beta_i. \quad (1\cdot2)$$

The econometric models presented above need to be tested for the model performance's goodness in explaining the relationships between the variables (ArunKumar *et al.* 2021). It can be done using AIC:

$$AIC=2\cdot k-2\cdot\ln\left(\hat{B}\right). \quad (13)$$

Due to the specific task under investigation, small datasets are available. To estimate the econometric model for such datasets, the AICc for the small samples can be used (Markevičiūtė *et al.* 2022):

$$AICc=AIC+\frac{2\cdot k^2+2\cdot k}{n-k-1}=$$
$$2\cdot k-2\cdot\ln\left(\hat{B}\right)+\frac{2\cdot k^2+2\cdot k}{n-k-1}. \quad (14)$$

Before using the econometric methods described above, time-series data must be tested for stationarity. Data differentiation (single or double) solves an issue in most practical cases if the data is not stationary. ADF test is used in applied statistics and econometrics to analyse time series for stationarity. ADF tests the null hypothesis (H0) that a unit root is present in a time series sample. If $p \leq 0.05$, the null hypothesis (H0) is rejected, the data does not have a unit root and is stationary. If $p > 0.05$, the null hypothesis (H0) fails to reject, the data has a unit root and is non-stationary (Brownlee 2020).

During the next stage, MLP ANN is chosen to solve this problem because its structure is the easiest and most flexible. The MLP used in this investigation has one hidden layer of neurons. Each neuron has a weighted connection to all inputs and each neuron's output is connected to the output neuron layer. The connection weights are tuned during the learning process. The disadvantage of this flexibility is the high computational complexity since each connection requires an additional multiplication. For this case study, the data amount is rather small, therefore computational power and time are low. An MLP layer with any number of neurons can be expressed by the following equation:

$$y=\sigma\cdot\left(\mathbf{W}\cdot x+b\right). \quad (15)$$

The hyperbolic tangent (tanh) function was used to activate all MLP layers in this article.

The data, which included: tonne-km, industrial production, imports, consumer spending, and fuel prices, was supplied as a time series flattened to a one-dimensional array, this model outputs freight rate. Input indicators were defined based on the literature review and practical experience of investigators.

Different metrics can be used to evaluate the forecasting accuracy of the proposed models. The commonly used are (Nwokike *et al.* 2020; Liu *et al.* 2020, Jierula *et al.* 2021): (1) correlation *R* between the actual value and forecasted value; (2) forecast error; (3) MFE; (4) MSE; (5) RMSE; (6) MAE; (7) MAPE. Regarding the investigation provided by (Jierula *et al.* 2021), the sensitivity of some accuracy metrics can be ranked as follows: MSE > MAPE > MAE > RMSE > *R*. The more sensitive the metric is, the more suitable it is for comparing the accuracy of different predictions. The *R* and RMSE were used in this article (Liu *et al.* 2020):

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}\left(y_t-\hat{y}_t\right)^2}{n}}. \quad (16)$$

RMSE measures the average error magnitude between the predicted and actual values (Jierula *et al.* 2021). During ANN-based model development process, the MSE served as the metric for testing. In the context of forecasting contract freight rates through spot rates, the primary metric employed was the correlation coefficient *R*. The article does not provide explicit formulas for these metrics, assuming that existing software can automatically compute them.

This section has comprehensively introduced a spectrum of models suitable for forecasting both freight rates and demand. It has delineated and described the prerequisites for data pre-processing, offering clarity on the essential steps during data preparation. Furthermore, the section has presented the commonly employed metrics for results evaluation. All models from this section were deployed in the investigation presented in this article and achieved results are presented in the following section.

3. Results

This section delves into the analysis of the European road freight market, with a case study focused on the destination from the Netherlands to Italy. It evaluates the forecasting capabilities of econometric and ANN-based models, presenting the procedures for data stationarity, model selection based on AICc, and the precision of the resulting forecasts. This section aims to distil the comparative effectiveness of the employed models in capturing the dynamics of freight rates and demand within this specific destination.

3.1. Results achieved using quantitative individual econometric models

This subsection is dedicated to the comprehensive evaluation of econometric models, with a particular focus on normalised freight rates and demand data. It is internal company data, which has been collected during the last years as was described in Section 2. The dataset contains 90 months, data from months 1 to 78 is used for model development, and data from 79 to 90 months is used for testing. Figure 2 provides a visual representation of the average monthly values used in the investigation. To ensure the possibility of using econometric models ADF tests have been performed, the results of which are detailed in Table 1. Notably, the data was initially found to be non-stationary. The procedure of data differentiation was applied and the ADF test was repeated. It was found that after the differentiation procedure, the data became stationary and econometric models can be used.

During the 1st steps, econometrics models presented in Section 2 were used to forecast the demand. The investigation progresses by delving into various configurations of the ARIMA model, with a keen emphasis on the selection process guided by the AICc for optimal forecasting performance. Multiple econometric models underwent testing, ranging from simple 1st-level AR and MA models to the culmination with the ARIMA (4, 2, 4) model (comprising 74 combinations for each dataset). Similarly, the procedure was repeated for SARIMA models. The best results were achieved using a seasonality of 12 months.

The top 3 ARIMA models were chosen based on the minimal AICc value, the same procedure was repeated using the SARIMA model, and the results are shown in Table 2. Figures 3 and 4 offer a visual depiction of the demand prognosis for the subsequent 12 months (from 79 to 90 months) using models from Table 2, which are compared with actual data. In Figures 3 and 4 prediction intervals have also been shown. Notably, the ARIMA models exhibit superior performance, with the majority of estimated values falling within prediction intervals. The minimal RMSE value was obtained using the ARIMA (3, 0, 3) model, recording a numerical value of 0.0118 (as illustrated in Table 2). It is essential to highlight the overall commendable performance of all ARIMA models, as evidenced by the maximum RMSE of 0.0155. In contrast, SARIMA, while achieving favourable results, exhibited a maximal RMSE of 0.08888, with the most optimal performance obtained using the SARIMA (4, 0, 3)(4, 0, 0)^{12} model – RMSE = 0.01671.

After selecting the best econometric models, data from Figures 3 and 4 was integrated and compared with actual data (Figure 5). Graphical representations of the modelling results, presented in Figure 5, underscore the superior performance of the ARIMA model compared to SARIMA.

From the achieved results, we conclude that ARIMA models are suitable for demand forecasting in road freight transportation.

During the next step, similar procedure was repeated to find out the best models for freight rate forecasting. The freight data was used from Figure 2. Similarly, data from 1 to 78 months was used for model tuning, and the last 12 months (from 79 to 90 months) were used for testing.

Freight rate prognoses are presented in Figures 6 (ARIMA models) and Figure 7 (SARIMA). For the selected period, both ARIMA and SARIMA models yielded suboptimal results. The best models are presented in Table 3. However, prognosis data consistently fell outside prediction intervals. Received data was integrated, and the results are presented in Figure 8. Interestingly, when the actual data (not differentiated) was considered (Figure 8), the SARIMA model outperformed ARIMA in freight rate prognosis. However, the obtained results underscore the

Table 1. ADF test result

Data	Null rejected	*P*-value	Test statistics	Critical value	Significance level
Demand	false	0.610	–0.106	–1.945	0.05
Demand after differentiation	true	0.010	–10.249		
Freight rate	false	0.898	0.887		
Freight rate after differentiation	true	0.010	–10.707		

Table 2. Best econometric models for demand modelling

Model	ARIMA (2, 0, 4)	ARIMA (3, 0, 3)	ARIMA (4, 0, 4)	SARIMA (2, 0, 2)(4, 0, 0)^{12}	SARIMA (4, 0, 1)(4, 0, 0)^{12}	SARIMA (4, 0, 3)(4, 0, 0)^{12}
AICc	–2.11	–381.5	–381.7	–411.5	–408.6	–414.2
RMSE for prognosis values	0.0127	0.0118	0.0155	0.03288	0.08888	0.01671

Table 3. Best econometric models for freight rate modelling

Model	ARIMA (1, 0, 1)	ARIMA (3, 0, 4)	ARIMA (4, 0, 4)	SARIMA (3, 0, 3)(4, 0, 0)^{12}	SARIMA (3, 0, 4)(4, 0, 0)^{12}	SARIMA (4, 0, 3)(4, 0, 0)^{12}
AICc	–652.95	–659.76	–659.68	–641.06	–631.06	–631.37
RMSE for prognosis values	0.0346	0.0312	0.0308	0.0222	0.0222	0.0225

limitations of econometric models in forecasting freight rates during the specified 12 month period (July 2021 – June 2022). It is crucial to note that this period was marked by heightened market volatility, primarily influenced by the COVID-19 pandemic and geopolitical uncertainties. As a consequence, alternative methodologies may be warranted for more accurate freight rate forecasting during such turbulent market conditions.

Performed investigation showed that freight rate forecasting is challenging, and econometric models that are suitable for demand forecasting are not applicable for freight rate (price) forecasting.

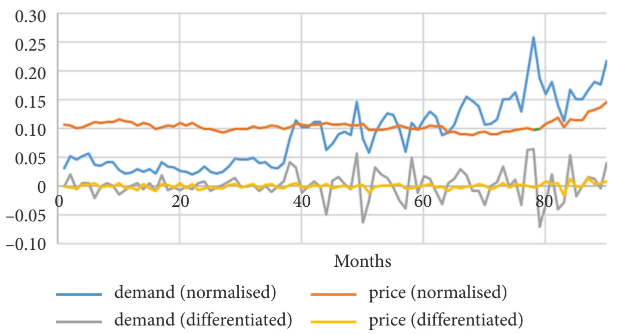

Figure 2. Demand and price (freight rate) for shipments

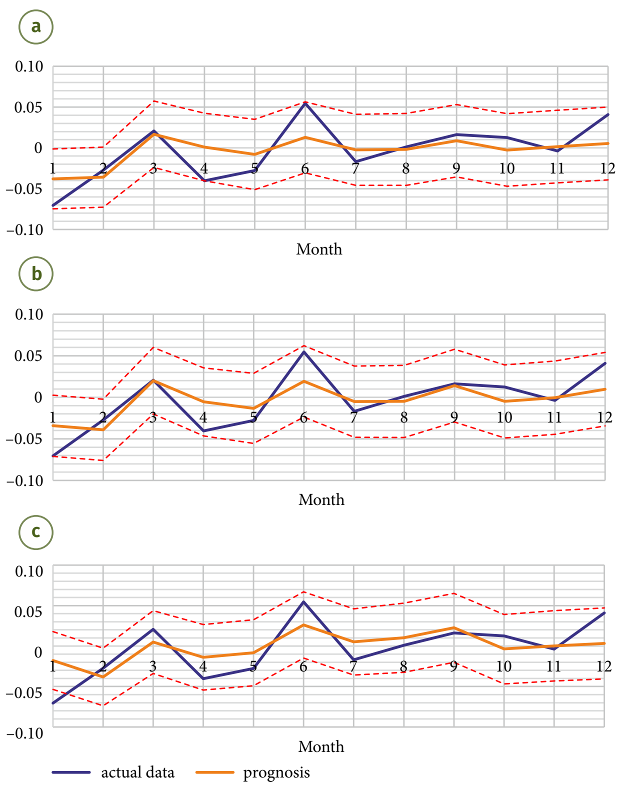

Figure 3. Prognosis with prediction intervals using ARIMA models for differentiated demand data:

a – ARIMA (2, 0, 4);

b – ARIMA (3, 0, 3);

c – ARIMA (4, 0, 4)

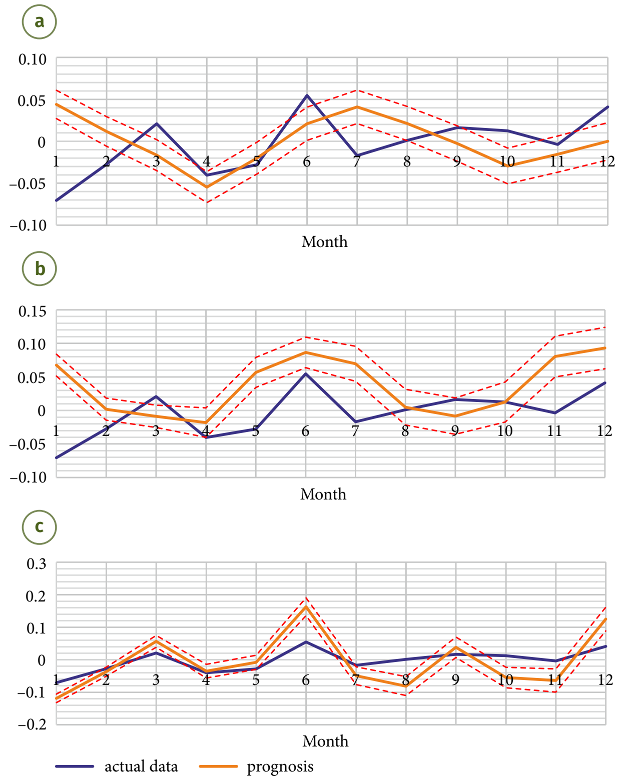

Figure 4. Prognosis with prediction intervals using SARIMA models for differentiated demand data:

a – SARIMA $(4, 0, 3)(4, 0, 0)^{12}$;

b – SARIMA $(4, 0, 1)(4, 0, 0)^{12}$;

c – SARIMA $(2, 0, 2)(4, 0, 0)^{12}$

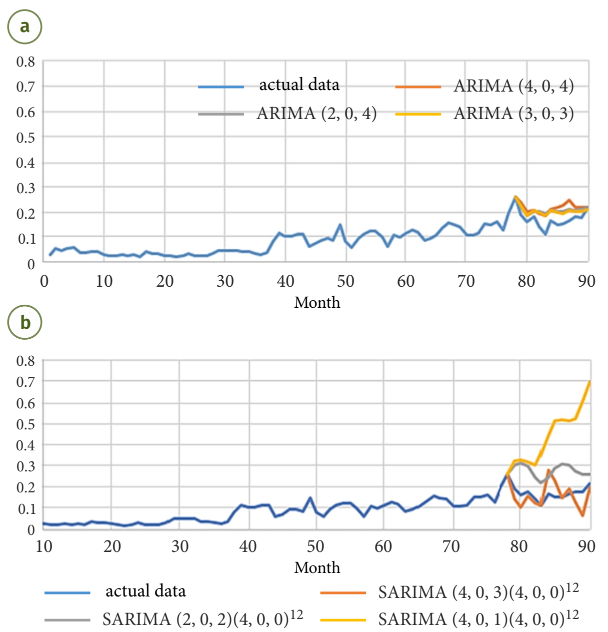

Figure 5. Prognosis and actual demand data using models:

a – ARIMA;

b – SARIMA

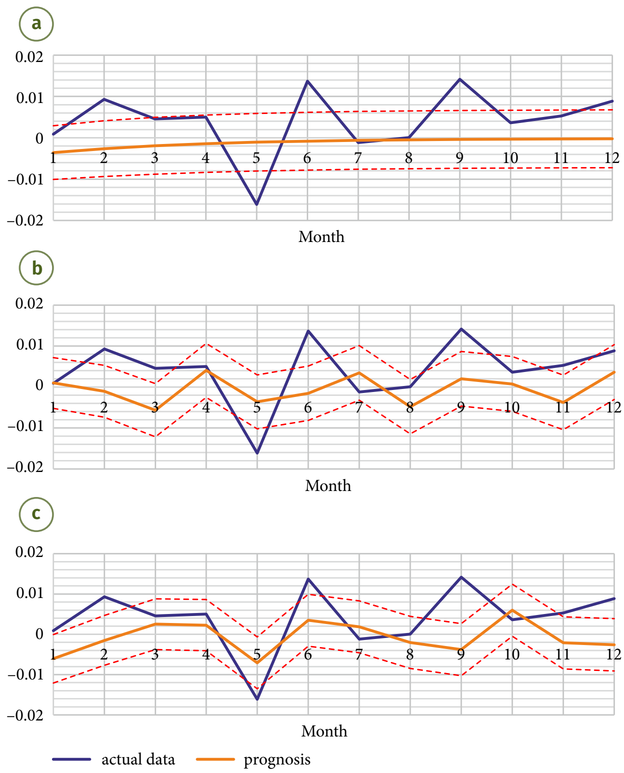

Figure 6. Prognosis with prediction intervals using ARIMA models for differentiated freight rate data:
a – ARIMA (1, 0, 1);
b – ARIMA (3, 0, 4);
c – ARIMA (4, 0, 4)

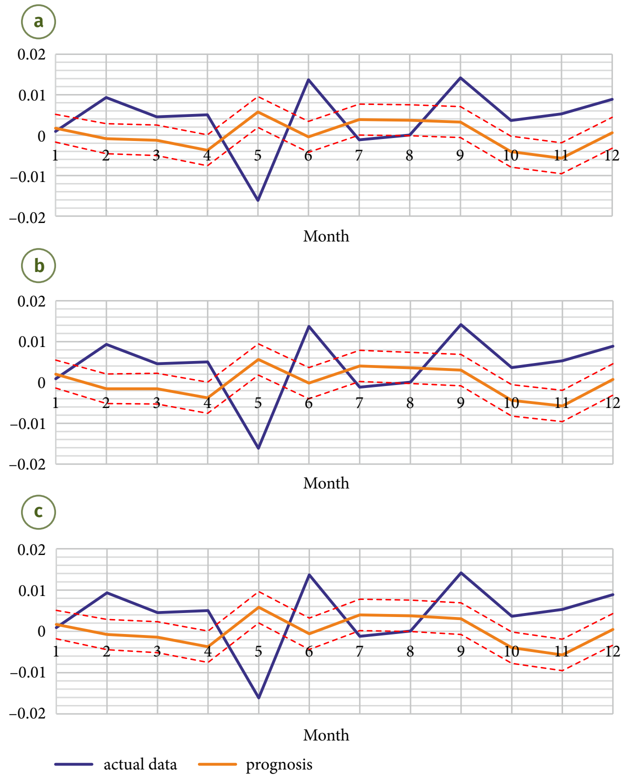

Figure 7. Prognosis with prediction intervals using SARIMA models for differentiated freight rate data:
a – SARIMA (3, 0, 4)(4, 0, 0)^{12};
b – SARIMA (4, 0, 3)(4, 0, 0)^{12};
c – SARIMA (3, 0, 3)(4, 0, 0)^{12}

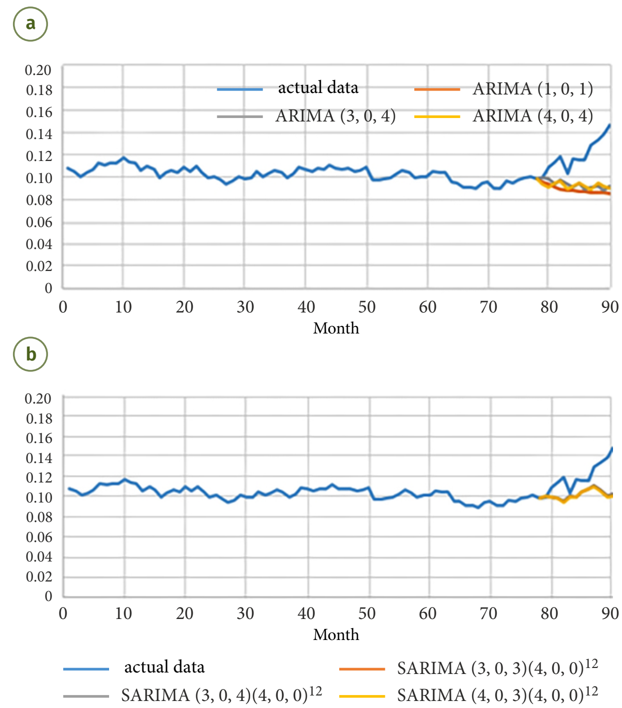

Figure 8. Prognosis and actual freight rate data using models:
a – ARIMA;
b – SARIMA

3.2. Results achieved using quantitative multivariate econometric and AI-based models

After finding out that both ARIMA and SARIMA models do not provide reliable results during freight rate forecasting. Multivariate models have been developed. Consistent with the findings in the literature, multivariate models offer a robust framework for forecasting tasks, addressing the complexities of multiple interrelated variables. This section extends the analysis begun in Sub-Section 3.1, utilising the same data from Netherlands to Italy direction (Figure 2). The approach here is to harness the predictive power of multivariate econometric models, exploring how these models can encapsulate the intricacies of forecasting when multiple factors are at play. In this research, key indicators used for freight rate forecasting are tonne-km, industrial production, imports, consumer spending, and fuel prices.

The main disadvantage of multivariate models is that key indicators used for prognosis also should be forecasted. For the case study, the data was available, as the results achieved using the developed model were compared with already available data. The forecast was performed for 12 months (from months 79 to 90 in Figure 2), from July 2021 to June 2022. Only for input parameter tonne-km, the data was unavailable and needed to be forecasted. This parameter is collected quarterly and presented by Eurostat with a delay of 6...9 months. The freight intensity index was used to forecast tonne-km. This indicator is defined as the ratio between tonne-km and gross domestic product. It was indexed in 1995; for each EU country and was equal to 1.

In Figure 9, freight intensity from 1 quarter of 2015 is presented. It can be seen that it has a linear trendline. For the Netherlands, the correlation coefficient is 0.87 (R^2 = 0.76); for Italy, the correlation coefficient is 0.82 (R^2 = 0.68). The prognosis for the next quarters can be built using trendlines presented in Figure 9.

After estimating freight intensity for the forecasting period, tonne-km can be evaluated. The rest input parameters (industrial production, imports, consumer spending, and fuel prices) were available, and an econometric model was developed using historical data.

Different combinations of ARIMAX and SARIMAX models have been tested in a similar manner as described in Sub-Section 3.1. Achieved results showed that the ARIMAX model performed better than SARIMAX for a selected case study. The best results were achieved using the ARIMAX (1, 0, 2) model.

At the same time, nonlinear AR MLP models have been developed. The main difference in using ANN is that they do not require stationarity of initial data. So raw data can be used without any procedures. Levenberg–Marquardt algorithm was used for training. The number of neurons in the hidden layer was changed and results were compared using MSE as the main metric, which is calculated automatically using the MATLAB ANN toolbox.

The best results were achieved for a model with 10 neurons in one hidden layer. Using more neurons, the tendency of the ANN towards overlearning was noticeable. MSE values decreased for the training dataset, but results were unsatisfactory while using the testing dataset.

Freight rate forecasting using 2 best models MLP and ARIMAX are presented in Figure 10.

The smaller achieved RMSE value for prediction is 0.01934 using the ANN model. It can be seen that ANN provide more accurate results by the end of the prognosis period; it is a minimal value achieved using different models for price prognosis.

The primary drawback of implementing the proposed solution lies in the requirement to forecast 5 parameters, which increases the potential for errors. To exemplify this concern, let's consider the case of the input parameter – fuel prices. In the US, transportation companies commonly rely on WTI crude oil price forecasts and NYMEX prognoses during the forecasting of fuel prices. The same forecasting model was employed for this investigation, as global oil trends tend to exhibit similarities.

However, it is worth noting that the prognosis for 2022 was entirely incorrect. In January 2022, the forecast predicted a decrease in oil prices from nearly $80 per barrel observed in the previous year to approximately $60 per barrel by the end of 2022. The situation unfolded differently, with the actual price soaring to nearly $120 per barrel by mid-year and then receding to $80 per barrel by year-end.

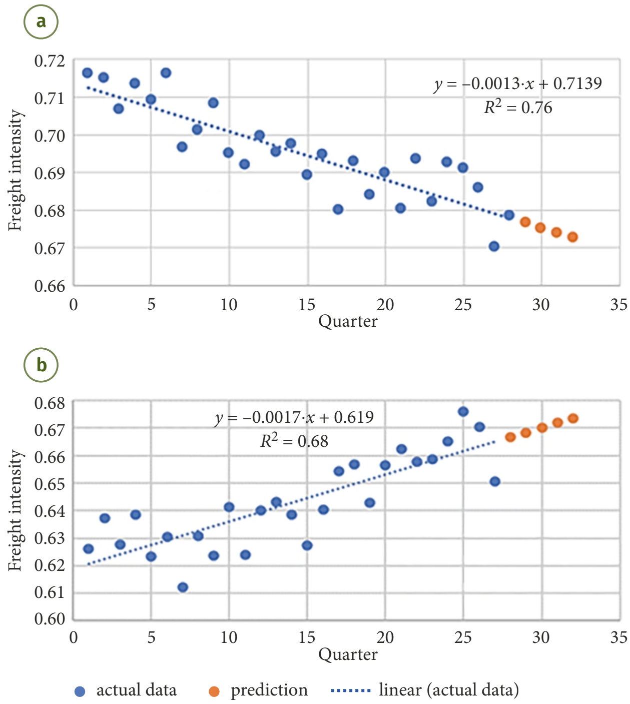

Figure 9. Freight intensity (2015–2021) in:
(a) – Netherlands;
(b) – Italy

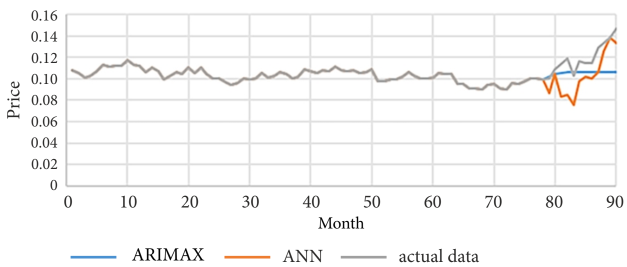

Figure 10. Prognosis and actual freight rate data using ARIMAX and ANN model

Consequently, the idea suggesting that multivariate models would outperform conventional individual models must be approached with the utmost caution, particularly considering the need to forecast other parameters. In times of market volatility, such an approach introduces significant uncertainty. Therefore, a robust solution is still needed to tackle these challenges effectively.

3.3. Forecasting contract freight rates using spot rates

This subsection tests the idea that contract freight rates may correlate with spot rates with some delay. In some industries, this correlation was found previously. If spot rates change, contract rates should be expected to change over the next several months.

The next attempt to perform a freight rate prognosis was to separate spot and contract rates. Unfortunately, such data separation was not previously used in a company under investigation and was launched only in 2020.

Data was collected separating contract and spot rates including 24 month period. The direction is the same (from the Netherlands to Italy). After that, rates were compared using correlation *R* as a primary metric. In Table 4, the results are presented. Month lag 0 means that the contract rate of month one is compared with the spot rate of month one, month lag 1 means that month 2's contract rate is compared with month one's spot rate, etc. It can be seen that the highest correlation is achieved with a lag equal to 6 months. The rate curve of the contract repeats the tendencies of the spot.

For the 6 month lag, the correlation coefficient is equal to 0.86; graphically, this data is presented in Figure 11.

For the selected case study it was found that spot rates could be used as a primary metric for evaluating contract rates. First-of-all, such an approach may be used during market volatility, while the accuracy of more advanced multivariate solutions decreases. The main advantage of this technique is simplicity and robustness. However, this is not a universal solution for all the cases. For example in some directions 12 month seasonability is possible. An example can be the direction where fruit or vegetables are transported. As these products are seasonal, and the season repeats every 12 months, a defined 6 month lag will not work. Before forecasting contract freight rates using spot rates initial investigation need to be performed using historical data.

Conducting this investigation revealed the formidable challenge of forecasting freight rates and demand in road transportation. There is no one-size-fits-all solution applicable to all market scenarios. Econometric models prove valuable in demand forecasting, as demand exhibits fewer fluctuations compared to price during market volatility.

Table 4. Correlation of spot rates with contract rates

Month lag	Correlation
0	0.71
1	0.64
2	0.54
3	0.61
4	0.76
5	0.72
6	0.86
7	0.78
8	0.72
9	0.50

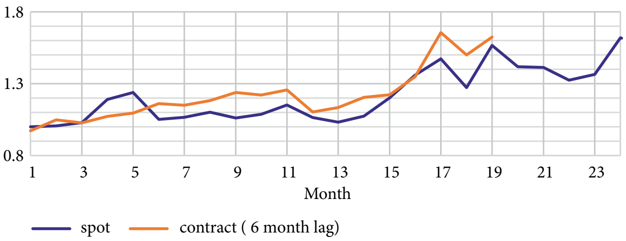

Figure 11. Spot vs contract rate with a 6 month lag

In this article, ANN demonstrates greater accuracy in forecasting freight rates compared with results achieved using econometric models. Another critical consideration, particularly in the European market, is the delayed availability of statistical data compared to the US market. High market volatility poses a significant drawback during the forecasting process, potentially leading to inaccuracies in input parameters, which are themselves subject to forecasting. Spot price analysis is a potent tool for forecasting contract prices. However, it is essential to account for the nature of the transported goods. Seasonal products, for instance, may render this method less effective. Summarising this section, it can be stated that the mathematical models developed can serve as supplementary tools in the decision-making process, complementing each other but not eliminating the necessity for human judgment.

Conclusions

An in-depth analysis of forecasting techniques has been conducted, focusing on demand and rate forecasting in road freight transportation. The demand forecasting results revealed that the ARIMA (3, 0, 3) model achieved the best outcomes. It should be noted that this model is not universally applicable, and the other ARIMA models must be tested for each particular direction. In any way, other ARIMA models also provide good results for demand forecasting. Conversely, using SARIMA models resulted in inferior forecasting results for all the cases examined.

Freight rate forecasting proved challenging due to the significant impact of COVID-19 and geopolitical factors that caused substantial rate fluctuations during the investigation period (2021–2022). Standard ARIMA and SARIMA models performed unsatisfactorily for this task. Instead, multivariate models, such as ARIMAX and SARIMAX and ANN-based solutions, were utilised to forecast rates by incorporating input data: tonne-km, industrial production, imports, consumer spending, and fuel prices. This approach yielded better results compared to the previously described methods. However, even such results were only achieved using actual values of additional input data. Using forecasted values of these parameters (additional input data) in practical applications may lead to improper performance due to their forecast errors.

A different approach was adopted to address the issue of freight rate forecasting for transportation companies. Freight rate data was collected, separating freight rate types into spot and contract over 24 months. Analysis revealed that contract rates mirrored the trends observed in spot rates but with a 6 month delay. The correlation coefficient between spot and contract rates for the case under investigation was found to be 0.86. It means that the company must collect the contract and spot data separately. Such a simple approach can aid decision-making by allowing transportation companies to forecast future contract rates based on the trends observed in the corresponding spot rates.

Despite market uncertainty, econometric models can still be employed to forecast demand, as the demand for goods tends to exhibit less drastic changes. However, freight rate forecasting poses unique challenges due to the substantial fluctuations in the parameters used as inputs for multivariate models. In light of this, we propose a 2 pronged approach for freight rate forecasting. Firstly, we suggest utilising ANN-based models in stable situations. Secondly, when confronted with volatile market conditions, it becomes crucial to analyse and consider spot prices in the forecasting process. Contract prices typically follow the trends observed in spot prices, albeit with a certain delay. By combining ANN-based models for stable situations and incorporating spot price analysis for volatile market conditions, a comprehensive and robust framework for freight rate forecasting can be established.

Summarising this investigation, it can be stated that developed mathematical models can serve as supplementary tools in the decision-making process during the freight rate and demand forecasting process, complementing each other but not eliminating the necessity for human judgment.

Author contributions

Conceptualisation – *Edvardas Liachovičius*, *Eldar Šabanovič*, *Viktor Skrickij*.

Mathematical modelling – *Edvardas Liachovičius*, *Eldar Šabanovič*, *Viktor Skrickij*.

Formal analysis – *Viktor Skrickij*, *Edvardas Liachovičius*.

Data curation – *Edvardas Liachovičius*, *Viktor Skrickij*.

Writing (original draft) – *Viktor Skrickij*.

All authors have read and agreed to the published version of the manuscript.

Disclosure statement

The authors declare no conflict of interest.

References

Al Hajj Hassan, L.; Mahmassani, H. S.; Chen, Y. 2020. Reinforcement learning framework for freight demand forecasting to support operational planning decisions, *Transportation Research Part E: Logistics and Transportation Review* 137: 101926. https://doi.org/10.1016/j.tre.2020.101926

ArunKumar, K. E.; Kalaga, D. V.; Kumar, C. M. S.; Chilkoor, G.; Kawaji, M.; Brenza, T. M. 2021. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: auto-regressive integrated moving average (ARIMA) and seasonal auto-regressive integrated moving average (SARIMA), *Applied Soft Computing* 103: 107161. https://doi.org/10.1016/j.asoc.2021.107161

Batchelor, R.; Alizadeh, A.; Visvikis, I. 2007. Forecasting spot and forward prices in the international freight market, *International Journal of Forecasting* 23(1): 101–114. https://doi.org/10.1016/j.ijforecast.2006.07.004

Brownlee, J. 2020. *Introduction to Time Series Forecasting with Python: How to Prepare Data and Develop Models to Predict Future*. Machine Learning Mastery. 367 p.

Chen, X.-W.; Lin, X. 2014. Big data deep learning: challenges and perspectives, *IEEE Access* 2: 514–525. https://doi.org/10.1109/ACCESS.2014.2325029

CL. 2020. *2020 Recession Outlook: 5 Key Indicators Driving Truckload Rates*. Coyote Logistics (CL). Available from Internet: https://coyotelogistics.medium.com/2020-recession-outlook-5-key-indicators-driving-truckload-rates-4a92aa6b0b91

CL. 2024. *Contract vs. Spot Rates: What's the Difference in Truckload Freight Shipping?* Coyote Logistics (CL). Available from Internet: https://resources.coyote.com/source/contract-vs-spot-rates

Claveria, O.; Monte, E.; Torra, S. 2017. Using survey data to forecast real activity with evolutionary algorithms. A cross-country analysis, *Journal of Applied Economics* 20(2): 329–349, https://doi.org/10.1016/S1514-0326(17)30015-6

Dekker, M.; Van Donselaar, K.; Ouwehand, P. 2004. How to use aggregation and combined forecasting to improve seasonal demand forecasts, *International Journal of Production Economics* 90(2): 151–167. https://doi.org/10.1016/j.ijpe.2004.02.004

ERTRAC. 2021. Carbon-Neutral Road Transport 2050: a Technical Study From a Well-to-Wheels Perspective. European Road Transport Research Advisory Council (ERTRAC). 37 p. Available from Internet: https://www.ertrac.org/wp-content/uploads/2022/12/ERTRAC-PPT-Carbon-Neutral-Road-Transport-2050_Workshop_April_29.pdf

Falatouri, T.; Darbanian, F.; Brandtner, P.; Udokwu, C. 2022. Predictive analytics for demand forecasting – a comparison of SARIMA and LSTM in retail SCM, *Procedia Computer Science* 200: 993–1003. https://doi.org/10.1016/j.procs.2022.01.298

Fernández-Portillo, A.; Almodóvar-González, M.; Sánchez-Escobedo, M. C.; Coca-Pérez, J. L. 2022. The role of innovation in the relationship between digitalisation and economic and financial performance. A company-level research, *European Research on Management and Business Economics* 28(3): 100190. https://doi.org/10.1016/j.iedeen.2021.100190

Glenn, J. C.; Gordon, T. J. 2009. *Futures Research Methodology – Version 3.0*. The Millennium Project. 1300 p.

Goodfellow, I.; Bengio, Y.; Courville, A. 2016. *Deep Learning*. MIT Press. 800 p.

He, Y.; Henze, J.; Sick, B. 2020. Continuous learning of deep neural networks to improve forecasts for regional energy markets, *IFAC-PapersOnLine* 53(2): 12175–12182. https://doi.org/10.1016/j.ifacol.2020.12.1017

Jierula, A.; Wang, S.; OH, T.-M.; Wang, P. 2021. Study on accuracy metrics for evaluating the predictions of damage locations in deep piles using artificial neural networks with acoustic emission data, *Applied Sciences* 11(5): 2314. https://doi.org/10.3390/app11052314

Kauko, K.; Palmroos, P. 2014. The Delphi method in forecasting financial markets – an experimental study, *International Journal of Forecasting* 30(2): 313–327. https://doi.org/10.1016/j.ijforecast.2013.09.007

Kavussanos, M. G.; Visvikis, I. D.; Batchelor, R. 2004a. Over-the-counter forward contracts and spot price volatility in shipping, *Transportation Research Part E: Logistics and Transportation Review* 40(4): 273–296. https://doi.org/10.1016/j.tre.2003.08.007

Kavussanos, M. G.; Visvikis, I. D.; Menachof, D. A. 2004b. The unbiasedness hypothesis in the freight forward market: evidence from cointegration tests, *Review of Derivatives Research* 7(3): 241–266. https://doi.org/10.1007/s11147-004-4811-7

Klujsza, K. 2024. *Q1 2024 Truckload Market Forecast: Spot & Contract Freight Rate Trends*. Coyote Logistics (CL). Available from Internet: https://resources.coyote.com/source/us-truckload-market-guide

LeCun, Y.; Bengio, Y.; Hinton, G. 2015. Deep learning, *Nature* 521(7553): 436–444. https://doi.org/10.1038/nature14539

Liachovičius, E.; Skrickij, V.; Podviezko, A. 2020. MCDM evaluation of asset-based road freight transport companies using key drivers that influence the enterprise value, *Sustainability* 12(18): 7259. https://doi.org/10.3390/su12187259

Liu, H.; Li, C.; Shao, Y.; Zhang, X.; Zhai, Z.; Wang, X.; Qi, X.; Wang, J.; Hao, Y.; Wu, Q.; Jiao, M. 2020. Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011–2019 using the seasonal autoregressive integrated moving average (SARIMA) and exponential smoothing (ETS) models, *Journal of Infection and Public Health* 13(2): 287–294. https://doi.org/10.1016/j.jiph.2019.12.008

Mahlamäki, T.; Storbacka, K.; Pylkkönen, S.; Ojala, M. 2020. Adoption of digital sales force automation tools in supply chain: customers' acceptance of sales configurators, *Industrial Marketing Management* 91: 162–173. https://doi.org/10.1016/j.indmarman.2020.08.024

Markevičiūtė, J.; Bernatavičienė, J.; Levulienė, R.; Medvedev, V.; Treigys, P.; Venskus, J. 2022. Attention-based and time series models for short-term forecasting of COVID-19 spread, *Computers, Materials & Continua* 70(1): 695–714. https://doi.org/10.32604/cmc.2022.018735

Miller, J. W.; Scott, A.; Williams, B. D. 2021. Pricing dynamics in the truckload sector: the moderating role of the electronic logging device mandate, *Journal of Business Logistics* 42(4): 388–405. https://doi.org/10.1111/jbl.12256

Nwokike, C. C.; Offorha, B. C.; Obubu, M.; Ugoala, C. B.; Ukomah, H. I. 2020. Comparing SANN and SARIMA for forecasting frequency of monthly rainfall in Umuahia, *Scientific African* 10: e00621. https://doi.org/10.1016/j.sciaf.2020.e00621

Retek, M. 2021. Scenario building in an interactive environment and online communication, *Technological Forecasting and Social Change* 162: 120395. https://doi.org/10.1016/j.techfore.2020.120395

Ruggieri, R.; Savastano, M.; Scalingi, A.; Bala, D.; D'Ascenzo, F. 2018. The impact of digital platforms on business models: an empirical investigation on innovative start-ups, *Management & Marketing* 13(4): 1210–1225. https://doi.org/10.2478/mmcks-2018-0032

Ruiz-Aguilar, J. J.; Turias, I. J.; Jiménez-Come, M. J. 2014. Hybrid approaches based on SARIMA and artificial neural networks for inspection time series forecasting, *Transportation Research Part E: Logistics and Transportation Review* 67: 1–13. https://doi.org/10.1016/j.tre.2014.03.009

Schramm, H.-J.; Munim, Z. H. 2021. Container freight rate forecasting with improved accuracy by integrating soft facts from practitioners, *Research in Transportation Business & Management* 41: 100662. https://doi.org/10.1016/j.rtbm.2021.100662

Shukur, O. B.; Lee, M. H. 2015. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA, *Renewable Energy* 76: 637–647. https://doi.org/10.1016/j.renene.2014.11.084

Tauscher, K.; Kietzmann, J. 2017. Learning from failures in the sharing economy, *MIS Quarterly Executive* 16(4): 2. Available from Internet: https://aisel.aisnet.org/misqe/vol16/iss4/2/

TI. 2024. *European Road Freight Transport 2023*. Report ID 1459895. Transport Intelligence (TI). 193 p.

Truant, E.; Broccardo, L.; Dana, L.-P. 2021. Digitalisation boosts company performance: an overview of Italian listed companies, *Technological Forecasting and Social Change* 173: 121173. https://doi.org/10.1016/j.techfore.2021.121173

Vilutienė, T.; Podvezko, V.; Ambrasas, G.; Šarka, V. 2014. Forecasting the demand for blue-collar workers in the construction sector in 2020: the case of Lithuania, *Economic Research – Ekonomska Istraživanja* 27(1): 442–462. https://doi.org/10.1080/1331677X.2014.966972

Wang, S.; Chaovalitwongse, W. A. 2011. Evaluating and comparing forecasting models, in J. J. Cochran, L. A. Cox, P. Keskinocak, J. P. Kharoufeh, J. C. Smith (Eds.). *Wiley Encyclopedia of Operations Research and Management Science*, eorms0307. https://doi.org/10.1002/9780470400531.eorms0307

Wen, D.; Liu, L.; Wang, Y.; Zhang, Y. 2022. Forecasting crude oil market returns: enhanced moving average technical indicators, *Resources Policy* 76: 102570. https://doi.org/10.1016/j.resourpol.2022.102570

Zhao, Y.; Van Delft, S.; Morgan-Thomas, A.; Buck, T. 2020. The evolution of platform business models: exploring competitive battles in the world of platforms, Long Range Planning 53(4): 101892. https://doi.org/10.1016/j.lrp.2019.101892