FORECASTING FUTURE TRENDS IN FREIGHT TRANSPORT IN SLOVENIA UNTIL THE YEAR 2030

Dejan Dragan a,*, Vlado Popović b, Abolfazl Keshavarz c, Borut Jereb a, Tomaž Kramberger a

a,* University of Maribor, Faculty of Logistics, Slovenia.
b University of Belgrade, Faculty of Transport and Traffic Engineering, Serbia.
c University of Kuala Lumpur, Faculty of Business and Administration, Malaysia.

Abstract: Slovenia regarding the freight traffic belongs at the very top of the EU, while every sixth truck is excessively overloaded. The main aim of the present study is to accurately forecast future freight traffic on the highways until the year 2030. For this purpose, the adequate Box-Jenkins time series model has been identified. In order to obtain the best possible structure and parameters of a final model, a unique heuristic modeling framework containing an entire composition of different rigorous statistical criteria has been applied. For making predictions of future freight transport, a Monte Carlo scenario-playing framework has been conducted. The results show that an interval forecast (16458, 20671) million Tkm with a 95% likelihood is expected at the beginning of 2030. Regarding the practical point of view, this study has contributed to a decision-making process while studying different planning approaches about adequate fees' systems for trucks on the highways.

Keywords: Transport Planning; Forecasting Models; Monte Carlo Procedure; Road Freight Transport in Slovenia.

1. INTRODUCTION

In terms of the share of international freight transport, Slovenia belongs to the very top of the members of the European Union; in addition, every sixth truck is overloaded. It is also estimated that the freight traffic will considerably increase over the next 15 years, with diverse projections that are quite different from one another. The main aim of this study is to try to predict, in the most professional and credible way, the increase in freight traffic by 2030.

According to the official facts, foreign trucks represent more than 90 percent of total freight traffic in Slovenia, while only Lithuania has a higher share. In many countries, for example in Italy and Germany, the ratio between foreign and domestic trucks is just the
opposite (Ocepek, 2013). In the case of Austria, a drop in transit traffic is evident, especially due to strict environmental policy. As a result of this fact, the bulk of transit freight transport has moved to the neighboring countries, particularly to Slovenia. In the countries of Central Europe, such as Slovakia, the Czech Republic, Hungary and Slovenia, there has been a significant increase in international freight transit traffic since 2004, when these countries joined the European Union (Ocepek, 2013).

The reasons for such an increase in transit freight transport are, for example, the enlargement of the EU and the subsequent Schengen area, as well as the significance of the strategic position of Slovenia being located at the crossroads of the 10th and 5th European transport corridor. In addition, some other reasons also contribute to an increase in foreign transit, such as, for example, price and low tolls, which is the reason why freight flows prefer to be redirected across Slovenia, rather than going on a shorter path (Ocepek, 2013). Latest state shows that highway driving is changing to endless overtaking of trucks on weekdays. The most heavily burdened sections of Slovenian highways transports are approaching to a load of seven thousand trucks weighing more than seven tons a day, so one every 12 seconds.

There are several ways to solve problems related to overloaded transit freight transport. One way is to gradually move freight from the roads to a railway network. Unfortunately, this is not possible at the current moment for Slovenia. The reasons are enormous costs, associated with a reconstruction of the obsolete and slow rail network. Also, other related investments, in this case, would be too costly and thus are unacceptable. The rising of tolls is also not a solution, mostly due to bilateral and multilateral political reasons on the relation (Slovenia – the EU, Slovenia – other EU countries). Thus, the only acceptable way has been the construction of a new and efficient toll system that can usually significantly decrease the existing traffic flows.

As a response to problems mentioned above, on the April 1st this year, a new free-flow electronic tolling system (ETS) has been initiated on Slovenian highways for heavy goods vehicles. While a couple of studies related to choosing of most appropriate technology for ETS have been carried out, a DSRC (Dedicated Short Range Communication) technology was finally selected (Kroflič, 2017; Miller et al., 2001).

As it is usual in wider planning of important transport projects, the planning of possible strategies must also take into account the most possible accurate long-term forecasting of future trends in freight transport. This is particularly important in the case of planning any of strategic future transport policies and strategies on the state level including a design of an appropriate tolling system on highways, where most freight traffic takes place, especially transit.

Our forecasting study was one of the key studies carried out during the process of choosing of most appropriate technology for ETS. Namely, the information about future trends of freight road transport on Slovenian highways was one of the key parameters during the decision-making process (DMP) regarding choosing the most suitable type of ETS. For the purpose of the study, the ARIMA model from the family of the Box-Jenkins time series models (Box et al., 2015) was used and integrated into the Monte Carlo scenario playing (MCSP) procedure for testing possible future scenarios. The entire ARIMA-MCSP framework represents a unique forecasting mechanism, with a composition of different rigorous criteria for selecting the most appropriate forecasting model. We believe that the paper might have included several contributions from the methodological and practical point of view, such as: a) the study presented here was one of the key studies...
during the DMP process of choosing the ETS system; b) to the best of our knowledge, practically none of the similar studies can be found in the existing literature; c) The unique ARIMA-MCSP framework was conducted for the case of traffic flow forecasting.

2. THE LITERATURE REVIEW

In general, we can classify traffic flow forecasting into two categories: a) short-term forecasting (STF); and b) long-term forecasting (LTF). The STF forecasting is particularly important for advanced traffic management and for information systems that belong to the intelligent transportation systems (ITS), mostly in the urban metropolitan areas (Bing et al., 2018). Here, the observation period and the forecasting horizon are typically considered in minutes’ or hours’ intervals. The forecasting methods in this respect can be roughly divided into parametric methods (PM) and nonparametric methods (NPM). The PM methods mostly include ARIMA and other Box-Jenkins time series models (Box et al., 2015), Kalman filtering-based models (Guo et al., 2014), and other parametric regression models. To overcome some limitations of PM models (e.g., normality of model’s residual), many studies have also conducted NPM methods, such as: nonparametric regression models, neural networks’ models, spectral analysis’s models, support vector machines’ models, and many more (Bing et al., 2018; Zhang, 2014).

Conversely to the STF models, there could not be detected many studies that would deal with a long-term traffic flow forecasting, particularly the ones that are also addressing the goods transport on the state level. In a study (Ishida and Okamoto, 2011), billions vehicle-km/year has been predicted regarding freight transport in Japan within a time interval [1980, 2030] by using log-linear regression models. The work of authors (Jha et al., 2016) has conducted the prediction of the total vehicular population in India by using different time intervals (up to the year 2021) and various ARIMA models. There have been also several other similar studies detected on the country (state, big city) level, such as for: 1) Australia; 2) India; 3) Illinois US; 4) Turkey; 5) Poland; 6) New Zealand; 7) North Carolina US; 8) Beijing, China; 9) Paris, France; 10) Spain; and 11) Pakistan.

Only some of these studies are dedicated exclusively to freight transport on the country (state) level. Their common denominator is that they possess a wider prediction time interval, which is monitored in terms of days, weeks, months or years ahead. The models for LTF forecasting vary from simple models (e.g., trend line models, linear-growth or log-linear models, exponential smoothing models), more complex classical models (e.g., Box Jenkins models, econometric models, ARMA-GARCH models, Dynamic Factor models, etc.), all over to advanced modern models (e.g., neural and fuzzy-neural models, agent-based models, models based on regression with support vectors, etc.). There are also a number of alternative models specially built for specific cases of traffic flow forecasting.

Details about these studies and applied models can be found in the literature (Chase et al., 2013; Choudhary et al., 2014; Chow et al., 2010; Chrobok, 2005; Dhingra et al., 1993; Dogan et al., 2018; Dorosiewicz, 2015; Gomez and Vassallo, 2015; IDFC, 2016; Jha et al., 2016; Jha et al., 2013; Kennedy and Wallis, 2007; Kulelpak and Sennaroglu, 2009; Li and Hensher, 2009; Matas and González-Savignat, 2009; Raikwar et al., 2017; Sabry et al., 2007; Slattery et al., 2004; Stone et al., 2006; Su et al., 2016; Toque et al., 2018; Tsekeris, 2011; Yu et al., 2017).
3. THE CONCEPTUAL FRAMEWORK OF A MODELING DESIGN

3.1 The characteristics of the data

Regarding historical data for the transport volume, it is worth to mention that traffic flows (in million-ton kilometers – mio Tkm) have been monitored in Slovenia already since the year 1954, with a higher level of precision since the year 1992 (Zanne, 2013). Concerning the systematic (automated) measurement of freight traffic, the latter has been monitored by the Directorate for Roads of the Republic of Slovenia since 2002 with automatic counters (Zanne, 2013). Trucks’ loads were measured in 2003 in 30 locations and in 2011 at 39 locations.

3.2 The applied forecasting model and relations with our previous work

While a procedure for searching the best ARIMA model’s structure and parameters was carried out, a unique heuristic-based approach for model selection was designed. The latter has incorporated a whole spectrum of different rigorous criteria, from the statistical-based, residual-based, all over to information-based criteria. The main working mechanism of this approach was previously introduced and applied in several other studies from the field of Maritime logistics (Dragan et al., 2017a; Intihar et al., 2015, 2017). For making predictions of long-term future freight transport in Slovenia, an MCSP framework was applied combined with a previously identified ARIMA model (Dragan, 2016). The main concept of combining different models with the MCSP concept was in-depth explained in a couple of other similar research studies, e.g., works (Dragan et al., 2017b; Ivanuša et al., 2018).

3.3 The conceptual framework

Figure 1 shows the conceptual framework of an entire modeling design. Firstly, the 100 quarterly data from 1990 to 2014 was collected (Block A). These data were assigned to the observed time series variable denoted by $z(t), t = 1990Q, ..., 2014Q$. The identification of the best ARIMA model has followed in the next step (Block B). The best model was found among an entire family of the model candidates with different models’ parameters, structures and orders, for which an estimation interval (EI) $t \in \{1990Q, ..., T = 2007Q\}$ was taken into account. For this purpose, the unique heuristic algorithm for model selection (firstly introduced in (Intihar et al., 2015)) represented in Block C was applied in a slightly adjusted form for the case of traffic flow forecasting (c.f. figure 2 (Intihar et al., 2017)). This heuristic also provided an appropriate validation of the best model (Block D) on the test interval (TI) $t \in \{T + Q, ..., 2014Q\}$. The interested reader can find the details of the working mechanism of this algorithm in our previous work (e.g., (Intihar et al., 2017)). The model with its output denoted by $\hat{z}(t), t \in t_p$ was afterwards injected into the MC algorithm (Block E), which has produced several hundred ($i = 1, ..., M_s$) possible future scenarios $\hat{z}(t,i), i = 1, ..., M_s, t \in t_p$ on the prediction interval (PI) $t_p \in \{2015Q, ..., 2030Q\}$ (Block F). As a final step, an averaging of future scenarios (Block G) was carried out across entire prediction interval with a purpose of obtaining of average (mean) values of the interval forecasts $\bar{z}(t,i) \pm 2 \cdot STD(\hat{z}(t,i)), i = 1, ..., M_s, t \in t_p$,
where $\bar{z}(t_{i}, i)$ are the averaged future point forecasts. Moreover, the normality test was also conducted for future scenarios $i = 1, ..., M_s$ in the final time point $T_p = 2030Q_t$. As it turned out, the future scenarios have been approximately normally distributed here, which means:

$$
\bar{z}(2030, i) \approx N(\bar{z}(2030, i), \text{STD}[\bar{z}(2030, i)]), i = 1, ..., M_s, t = 2030
$$

$$
P\left[ \bar{z}(2030, i) - 2 \times \text{STD}[\bar{z}(2030, i)] \leq \tilde{z}(2030, i) \leq \bar{z}(2030, i) + 2 \times \text{STD}[\bar{z}(2030, i)] \right] \approx 95\%
$$

(1)

**Figure 1.** The conceptual framework of an entire modeling design.

**Figure 2.** The heuristic algorithm for model selection to find the best ARIMA model.
4. THE ARIMA MODELS

In general, the structure of the ARMA models family can be represented by the following expression (Box et al., 2015; Ljung, 1998, 2000):

\[ z(t) + \sum_{i=1}^{n_a} a_i \cdot z(t-i) = \varepsilon_c(t) + \sum_{k=1}^{n_c} c_k \cdot \varepsilon_c(t-k) \]  

(2)

where the AR part’s order \( n_a \) and MA part’s order \( n_c \) refer to the oldest delay of output \( z(t) \) and the oldest delay of random noise \( \varepsilon_c(t) \), respectively. The noise is supposed to have the properties of the white noise. The equation (2) can also be rewritten in the more compact form of a transfer function (TF) equation, if the backshift operator \( q^{-1} \) is conducted (Ljung, 1998, 2000):

\[ A(q) \cdot z(t) = C(q) \cdot \varepsilon_c(t), \text{ where:} \]

\[ A(q) = 1 + a_1 \cdot q^{-1} + \ldots + a_{n_a} \cdot q^{-n_a} \]

\[ C(q) = 1 + c_1 \cdot q^{-1} + \ldots + c_{n_c} \cdot q^{-n_c} \]  

(3)

The ARIMA transfer function model has a similar structure as the model in expression (3), but the non-stationary output’s time series \( z(t) \) must be substituted with its differentiated equivalent. Usually the first order’s differentiation can achieve the stationarity, i.e.: \( \Delta z(t) = z(t) - z(t-1) = (1 - q^{-1}) \cdot z(t) \). This way, the TF in (3) takes the form:

\[ A(q) \cdot \Delta z(t) = C(q) \cdot \varepsilon_c(t) \]

\[ A(q) \cdot (1 - q^{-1}) \cdot z(t) = C(q) \cdot \varepsilon_c(t) \]  

(4)

5. THE NUMERICAL RESULTS

5.1 The data for road freight transport

Figure 3 shows the 100 quarterly historical time series data for Slovenian freight transport, measured over the period from 1990 up to 2014. From Figure 3 it can be noticed that freight transport significantly increased after the year 2004 when Slovenia entered into the EU. During the time of economic crisis, an oscillatory behavior can be noticed with a significant drop that has resulted in a local minimum reached in the third quarter of the year 2009. Later, the time series was stabilized with a slight drift that was present until the year 2014.
5.2 The estimation of the best ARIMA model

All calculations were conducted in Matlab with its extensions: The Statistics Toolbox, Econometrics Toolbox, and The System Identification Toolbox. The obtained ARIMA(2,1,1) model has the following estimated transfer function:

\[ A(q) \cdot \Delta z(t) = A(q) \cdot [z(t) - z(t-1)] = C(q) \cdot \epsilon_z(t) \]

\[ A(q) = 1 -1.328 \cdot q^{-1} + 0.6303 \cdot q^{-2} \]  
\[ C(q) = 1 + 0.3801 \cdot q^{-1}, \quad \left| \frac{t(a_1)}{t(a_2)} \right|, \left| \frac{t(c_1)}{t(c_2)} \right| > t_{krit} = 1.9966 \]  

(5)

The values in parenthesis represent the t-values of the obtained parameters, which are all statistically significant (\( |t(a_1)|, |t(a_2)|, |t(c_1)| > t_{krit} = 1.9966 \)). For the estimation of parameters, the first 70 time samples were applied, while in the testing of the predictive power of the model, the last 30 time samples were used.

5.3. Model validation

The model (5) has passed all the rigorous tests presented in figure 2. The diagnostic checking has given us the following important results (see equations in figure 2):

\[ AIC = 637.3656, \quad RMSE = 92.1291, \quad MAE = 65.3397 \]
\[ MAPE(\%) = 1.9122, \quad \%FIT = 82.1165, \quad DTW \{d[z, \hat{z}]\} = 78.2293 \]  

(6)

Since the MAPE is quite low, while the %FIT implies the fairly well fit of a model to the real data, we can conclude that the model might be efficiently used for predictions.

5.4. The prediction results

Figure 4 shows the model's output \( \hat{z}(t) \) and fit to the real data \( z(t) \) on the estimation interval: \( t \in \{1990Q_1, ..., 2007Q_4\} \), and on the test interval: \( t \in \{T_n + Q, ..., 2014Q_4\} \). As can be noticed from figure 4, in general, the model fits the real data quite well on the
estimation interval. On the test interval, after the initial oscillatory overshoot, the model gradually manages to capture the real data with a reasonably well fit.

Figure 4 also shows several hundred \(i = 1, ..., M_S\) generated future scenarios \(\tilde{z}(t,i), i = 1, ..., M_S, t \in t_p\) on the prediction interval \(t_p \in \{2015Q_1, ..., 2030Q_4\}\). Moreover, Block A points to the average (mean) values of the interval forecasts \(\bar{z}(t,i) \pm 2 \cdot STD(\bar{z}(t,i)), i = 1, ..., M_S, t \in t_p\), where \(\bar{z}(t,i)\) are the averaged future point forecasts. As it turned out, the normality tests (Jarque Bera test, Ljung-Box test, etc.) has confirmed that the future scenarios \(\tilde{z}(2030,i), i = 1, ..., M_S\) are approximately normally distributed in final time point \(T_p = 2030Q_4\). Consequently, it turns out that the following equations occur after the end of the MCSP procedure (see also equation (1) and Blocks from B to E in figure 4):

\[
\tilde{z}(2030,i) \approx N(18565, 1053), i = 1, ..., M_S, t = 2030
\]

\[
P(16458 \leq \tilde{z}(2030,i) \leq 20671) \approx 95\%
\]

Considering the fact that normality of \(\tilde{z}(2030,i)\) has been achieved, it can be concluded that there is a likelihood with a 95% probability of confidence in the result that the actual value of freight traffic at the beginning of 2030 will reach a certain value at the interval \(16458, 20671\) million Tkm. Furthermore, the most probable results can be expected around a value of 18565 million Tkm of road freight transport that will be most likely achieved in the year 2030.

Figure 4. The prediction results for the road freight transport in Slovenia.
6. CONCLUSION

The combined Monte Carlo Scenario Playing framework with integrated ARIMA model (2,1,2) has been introduced in this paper. The main purpose was to design a forecasting mechanism that would fairly accurately predict the future trends in freight transport in Slovenia until the year 2030. The study presented here was one of the crucial studies carried out during the process of choosing the most suitable technology for a newly designed electronic tolling system. In the future work, it is planned to methodologically upgrade a forecaster by the inclusion of additional modeling of possible seasonal, cyclic, intervention, and other effects. Moreover, it is also planned to enrich a model with influential economic exogenous indicators (such as GDP, import, export, etc.), which also presumably significantly impact on the dynamics of road freight transport.

REFERENCES


