Analysis of Ukrainian cargo transport industry through Bayesian inference and Kullback-Leibler information divergence

Alexey Pervukhin[1], Victoria Golikova[2]

Bayesian Young Statisticians Meeting (BAYSM), Milan June, 5-6, 2013
Paper no. 31

[1] Sevastopol National Technical University, Department of Technical Cybernetics, Sevastopol, Ukraine
alex@pervuh.sebastopol.ua
[2] Sevastopol National Technical University, Department of Technical Mechanics and Machines, Sevastopol, Ukraine
golikova.victoria@gmail.com

Abstract

The different aspects of an applied problem to increase efficiency of the Ukrainian cargo transport industry are examined. We use an information-theoretical paradigm to analyze observed data on dynamics of its indicators. Empirical models that are based on the data are testified at new data arrival. Bayesian criterion and Kulback-Leibler information divergence are used to verify the models.

Keywords: cargo transport industry, Bayesian inference, Kullback-Leibler information divergence.

Control and planning of freight were and are vital to efficiency in the transport system of Ukraine. Functioning of the different transport modes and transport industry as a whole is dependent on the decisions made here. Moreover, these decisions affect the processes of economic integration of Ukraine and well-developed EU countries.

Increased efficiency of the transport system and making only correct decisions can be reached through analyzing situation, changing trends of cargo transferring growth and developing optimal solutions. Obviously making right decisions under conditions of transition economy is difficult. Analysis and forecasting of indicators of cargo transport system and process of their dynamic interaction is one of the ways to solve the problem. It must be mathematical modeling of stochastic processes of informative parameters' changing on the basis of well-known methods of statistical inference.

There are domestic and foreign research centers where theoretical and applied models have been developed. However, the specific problems require specific models and software, taking into account the features of economic objects under study.

The aim is to continue an analysis of the freight transport system of Ukraine, started in [1]. Econometric models based on the main elements of the system are testified through new data, reflecting and explaining the dynamics of the volume of cargo transferred by main modes of transport from January 2003 to January 2012. Bayesian criterion and Kulback-Leibler information divergence are used to verify the model.

The analyzed data are available at the website of the State Committee of Statistics of Ukraine [2] and are monthly cargo volumes in million tons carried from January 2003 and January 2013 by four main modes of transport: railway (RW), road (RT), water (WT) and pipeline (PT). The role of the air transport is not significant, that is why it isn't considered. The data are non-stationary time series $X_{i,t}, ..., X_{n,t}$, where $i = 1 \div n$ is a number of the transport mode, $t = 1, \ldots, T$ is a number of observation. The values of all cargo volumes, transferred by n different modes of transport at the tome moment t consist $n \times 1$-vector $x_t = (X_{1,t}, ..., X_{n,t})^T = \begin{pmatrix} RT_t & PT_t & WT_t & RW_t \end{pmatrix}^T$. The changing of the vector x_t is multivariate non stationary discrete random process.

The observed data characterizing change of different parameters are discrete non-stationary random processes. Taking that into account the model is

$$X_t = \mu_0 + \mu t + \sum_{j=1}^{p} \varphi_j X_{t-j} + \varepsilon_t \qquad (1)$$

μ_0 is a constant, coefficient μ characterizes trend; φ_j are autoregression coefficients, p is a number of autoregression members, ε_t is normally distributed stationary random process.

The best model is chosen through the Bayesian inference. The novelty of the study is connected with using the information Kullback-Leibler divergence between conceptual reality and approximating model. It is shown that the approaches complete each other. For instance, the formula (2) being result of modeling expresses the relations between cargo volumes, transferred by different modes of transport:

$$RT_t = -0,456PT_t - 0,243WT_t + 0,425RW_t + 4,283 \qquad (2)$$

The relation (2) is graphed in Fig.1.

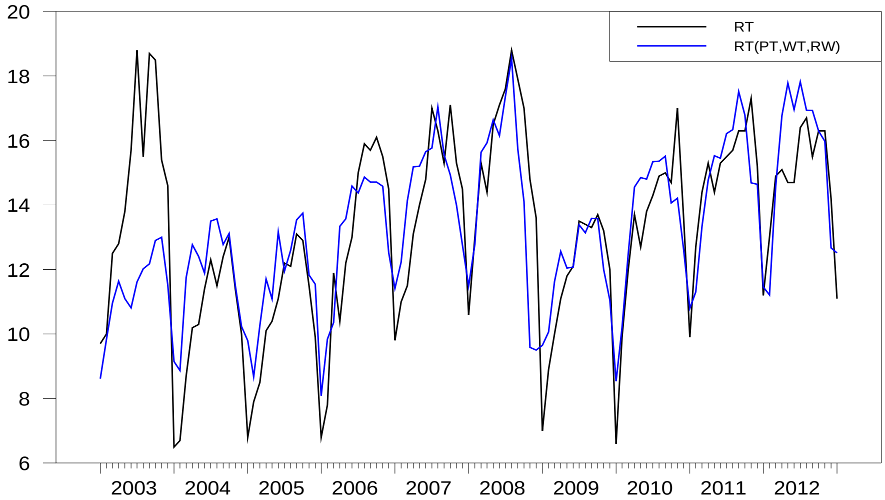

Fig.1. The RT values calculated through (2) and real values

References

[1] Emmenegger, J.-F., Pervukhin, A., Golikova, V. **Modeling of the Freight Transport System of Ukraine**. *Galician Economic Bulletin*; 2011; 4(33); pp. 49-57.

[2] The State Statistic Committee of Ukraine; http://www.ukrstat.gov.ua.