Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 161 (2016) 253 – 258

Procedia Engineering

www.elsevier.com/locate/procedia

World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium 2016. WMCAUS 2016

Advanced Forecasting Methods Based on Spectral Analysis

Dariusz Grzesica[a,*], Paweł Więcek[a]

[a]*Institute of Construction and Transportation Engineering & Management. Transport Section. Cracow University of Technology. Cracow. Poland*

Abstract

This article presents a time series estimation and prediction methods with the use of classic and advanced forecasting tools. In addition, rarely applied in practice approach using spectral analysis to an identification of variation patterns and prediction will be presented. The effect of spectral analysis will be an estimation of prediction model parameters. The main assumption is the model consist of trigonometric functions combination with a certain frequency. The model includes only those frequencies which have greatest influence on process variation. The effectiveness of the method will be examined by a numerical example. The area of the proposed methodology is broad and goes beyond economics.

Peer-review under responsibility of the organizing committee of WMCAUS 2016

Keywords: spectral analysis; forecasting time series;

1. Introduction

Time series analysis allows the detection of the nature described by the line of observation of the phenomenon and the ability to predict future values. Dynamically changing the values of time series are the result of exposure to many factors. which often cannot be identified. To avoid the necessity of determining their quantity. type and quantity of the effect. you can use the strong dependencies that arise between them. registered in the following time units. These variables are often characterized by cyclical variations. and random fluctuations. which can also be changed in accordance with the given pattern. In the time series there is the need to transform the observed values of the variable y_t in a form suitable to determine the dependencies that occur between values. This purpose will be used for spectral

* Corresponding author. Tel.: +48-12-628-30-49; fax: +48-12- 628-30-93.
E-mail address: dgrzesica@pk.edu.pl

Peer-review under responsibility of the organizing committee of WMCAUS 2016
doi:10.1016/j.proeng.2016.08.546

analysis as a tool for transforming time series into a series in the frequency domain. In the analysis of the mass spectra of the time series $\{y_t: t = 1. 2. \ldots. T\}$ is considered as an infinite sum of time series of different frequencies ω. which correspond to the oscillation periods $\tau = 2\pi/\omega$. For the description of the model will be used the trigonometric functions sine and cosine. i.e. a linear combination of elementary periodic signals. In turn. the transformation will be carried out using a discrete Fourier transformation for product demand data within two months. For the analysis of cyclic fluctuations. a time window of Bartlett will be used. It is an example of a filter which will be used to smooth the spectrum. obtained in the process of time series transformation. This approach allows you to identify those frequency components that contribute significantly to the explanation of changes in the researched time series. On this basis. a forecast for future time periods will be constructed.

2. Review of spectral analysis uses

The time series describing the given phenomena in time. in many cases. may take a chaotic process. Except of the number of components. such as a trend or a constant level of variable. there is also the cyclical and random component. The imposition of many cyclic components of different frequencies in combination with random fluctuations leads to the fact that widely used methods of forecasting are not accurate enough (Clements and Hendry. 2001). The considered phenomenon described using cyclical time series changes. is often called spectral analysis and can be applied to predict future sales in a company. It is a popular method. often used in Geophysics to study the physical processes occurring in the Earth (Burkhard. 2000). in astronomy to study the stars (Chattopadhyay and Chattopadhyay. 2014) and meteorology to predict the weather (Ehrendorfer. 2012). It can also be used for a prediction in the field of transportation as real-time traffic flow (Tchrakian et al.. 2012) or short-term traffic flow forecasting as an element of a hybrid method (Zhang et al.. 2014) using spectral analysis. In the literature of the subject. spectral analysis is not widely used. despite the fact that many economic and logistics phenomena include cyclic changes that are repeated at a certain period of time. In the field of management and optimization of goods and cash flows within companies organized in the logistics network it can be found only a slight mention in a scientific publication about the possibility of using spectral models in forecasting. There are no scientific papers that accurately reflect the technique under consideration approach. supported by the analysis of the efficiency by a representative set of real data. In connection with all the foregoing issues. there is a strong prerequisite for this work to increase benefits from the use of the presented approach in the field of demand forecast of the goods flow. in logistics enterprises.

3. Model formulation

The main objective of spectral analysis is to draw attention to the cyclical processes. It involves the wave structure of considered variables of stochastic processes. allowing us to analyze time series in the frequency domain. This is possible through the use of trigonometric functions as functions sine and cosine. often called harmonics (Anderson. 2011). Each function defined in the interval from 0 to π. The quantity of harmonics for n observations is n/2 (Hearn and Metcalfe. 1995). First harmonic has a period equal to n. the second n/2. one n/3. etc. For the application of this approach the stationarity requirement of the time series should be met [Warner. 1998]. Otherwise. you must delete the trend or to reduce it to stationarity by means of the differentiation operation. (Bisgaard and Kulahci. 2011). Thus. the process can be represented as follows:

$$y_t = f(t) + \sum_{i=1}^{n/2}\left[a_i \sin\left(\frac{2\pi}{n}it\right) + b_i \cos\left(\frac{2\pi}{n}it\right)\right] \tag{1}$$

where: *i* - number of harmonics. a_1. b_1. a_2. b_2.… - constant values. *f(t)* - the function describing the observed trend in the series.

The values of the parameters a_1. b_1. a_2. b_2 are obtained by the method of least squares. using the following structures:

$$a_i = \frac{2}{n}\sum_{i=1}^{n}\left[y_t \sin\left(\frac{2\pi}{n}it\right)\right] \quad dla\ \ i = 1 \ldots \frac{n}{2} - 1 \tag{2}$$

$$b_i = \frac{2}{n}\sum_{i=1}^{n}\left[y_t \cos\left(\frac{2\pi}{n}it\right)\right] \quad dla\ \ i = 1 \ldots \frac{n}{2} - 1 \tag{3}$$

For the last harmonic should be assumed that:

$$a_{n/2} = 0 \tag{4}$$

$$b_{n/2} = \frac{1}{n}\sum_{i=1}^{n}[y_t \cos(\pi t)] \tag{5}$$

Using a discrete Fourier transform for the considered time series a spectrum. which is a function of frequency is obtained. It allows you to determine which of the frequencies are able to explain the variability in greatest extent. that is. those having the greatest influence on the deviation of the forecasted variable. The part of the whole variance of variable through i- harmonic are presented by formulas:

$$\omega_i = \frac{a_i^2 + b_i^2}{2\sigma^2} \quad for\ i = 1 \ldots \frac{n}{2} - 1 \tag{6}$$

$$\omega_i = \frac{a_i^2 + b_i^2}{\sigma^2} \quad for\ i = \frac{n}{2} \tag{7}$$

where: σ^2 deviation of the forecast variable after preliminary deduction of the trend.

4. Results and discussions

The effectiveness of the proposed approach using the spectral analysis was examined on the basis of real plant data. The graph in Fig. 1. shows the daily profile of sales of the product.

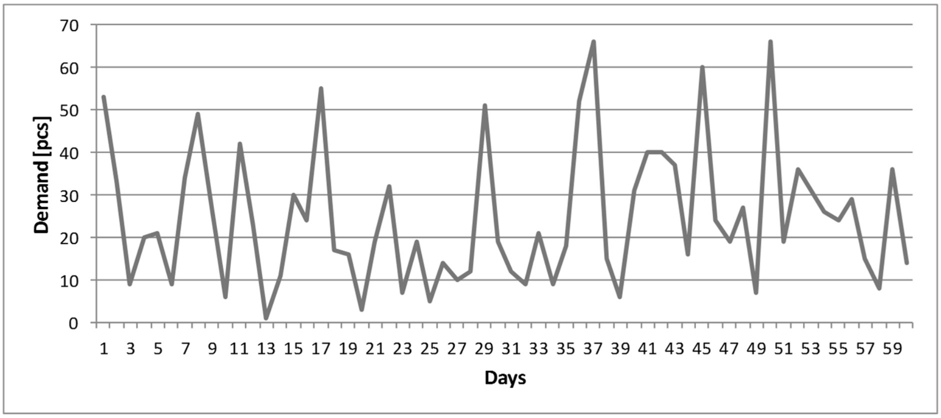

Fig. 1. The profile of demand for the product within two months.

From analysis of charts we can conclude that the pattern of demand for this product. possibly. takes the form over many cycles of various lengths and random fluctuations. To pre-determine the cycles that have the greatest influence on the variable. after preliminary removal of trend. the spectrum of data has been calculated. Then it was subjected to smoothing by spectral Bartlett window with zero values of the edge elements. to extract the frequency ranges in the spectral density of the largest component (Fig. 2). From the diagram it is possible to choose three areas for which the harmonics reach the greatest spectral density. therefore. they have the greatest influence on the shape of a number.

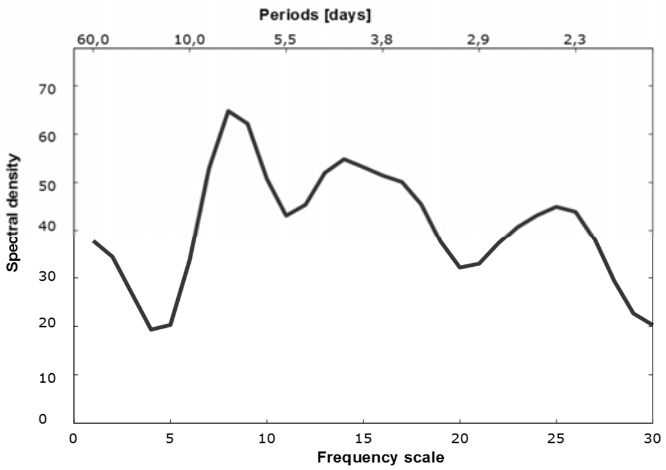

Fig. 2. The spectrum of the analyzed series.

To build the predictive model only the selected harmonics are used. which are given in table 1.

Table 1. Harmonics used to build predictive models.

No. selected harmonics	Designation	Period [days]	No. selected harmonics	Designation	Period [days]
1	n	60	17	n/17	3.53
2	n/2	30	19	n/19	3.16
7	n/7	8.57	20	n/20	3
8	n/8	7.5	21	n/21	2.86
9	n/9	6.67	22	n/22	2.73
12	n/12	5	23	n/23	2.61
13	n/13	4.62	25	n/25	2.4
14	n/14	4.29	26	n/26	2.31
15	n/15	4	28	n/28	2.14

After the estimation of the parameters using the available data samples. the obtained predictive model can be formulated as follow:

$$y_t = 0.0722t + 22.516 - 3.9\sin\left(\frac{2\pi}{60}t\right) + 2.5\cos\left(\frac{2\pi}{60}t\right) + 3.39\sin\left(\frac{2\pi}{60}2t\right) - 3.36\cos\left(\frac{2\pi}{60}2t\right) +$$

$$3.65\sin\left(\frac{2\pi}{60}7t\right) + 4.55\cos\left(\frac{2\pi}{60}7t\right) - 0.51\sin\left(\frac{2\pi}{60}8t\right) + 5.48\cos\left(\frac{2\pi}{60}8t\right) + 4.33\sin\left(\frac{2\pi}{60}9t\right) -$$

$$5.26cos\left(\frac{2\pi}{60}9t\right)+5.78\sin\left(\frac{2\pi}{60}12t\right) - 0.54\cos\left(\frac{2\pi}{60}12t\right) - 2.81\sin\left(\frac{2\pi}{60}13t\right) + 1.68\cos\left(\frac{2\pi}{60}13t\right) -$$

$$4.69\sin\left(\frac{2\pi}{60}14t\right) - 2.28\cos\left(\frac{2\pi}{60}14t\right) + 5.14\sin\left(\frac{2\pi}{60}15t\right) + 1.09\cos\left(\frac{2\pi}{60}15t\right) + 5.45\sin\left(\frac{2\pi}{60}17t\right) -$$

$$0.42\cos\left(\frac{2\pi}{60}17t\right) - 0.94\sin\left(\frac{2\pi}{60}19t\right) - 4.7\cos\left(\frac{2\pi}{60}19t\right) - 3.08\sin\left(\frac{2\pi}{60}20t\right) - 1.74\cos\left(\frac{2\pi}{60}20t\right) -$$

$$3.08\sin\left(\frac{2\pi}{60}21t\right) - 1.14\cos\left(\frac{2\pi}{60}21t\right) + 0.38\sin\left(\frac{2\pi}{60}22t\right) - 4.29\cos\left(\frac{2\pi}{60}22t\right) + 5.22\sin\left(\frac{2\pi}{60}23t\right) -$$

$$1.66\cos\left(\frac{2\pi}{60}23t\right) + 3.02\sin\left(\frac{2\pi}{60}25t\right) + 2.48\cos\left(\frac{2\pi}{60}25t\right) - 1.38\sin\left(\frac{2\pi}{60}26t\right) - 6.19\cos\left(\frac{2\pi}{60}26t\right) +$$

$$3.45\sin\left(\frac{2\pi}{60}28t\right) - 0.52\cos\left(\frac{2\pi}{60}28t\right) \tag{8}$$

The resulting model can be used to forecast future needs for a certain product. The theoretical values from the model along with the forecast is shown in Fig. 3. It can be noted that the designated forecast largely coincides with the actual values of the time series.

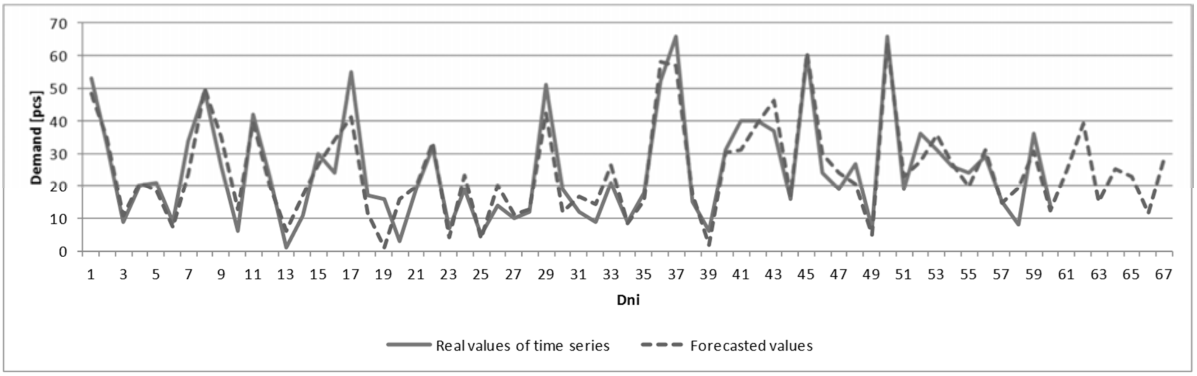

Fig. 3. Theoretical model values of r with the forecast.

On the background of the presented arguments a verification of the effectiveness of the proposed method will be made. by the use of the following evaluation measures:

Mean Absolute Error:

$$MAE = \frac{\sum_1^n |P_i - P_{ri}|}{n} \tag{9}$$

Mean Absolute Percentage Error:

$$MAPE = \frac{\sum_1^n |\frac{P_i - P_{ri}}{P_i}|}{n} \tag{10}$$

where: P_i - actual sales for the period "i". P_{ri} - forecast for the period "i". n - the sum of the forecasts.

The obtained forecast errors (the average relative and absolute error of prediction) for the spectral approach have been compared with the result received by Brown and ARMA models. The results are summarized in table 2.

Table 2. A comparison of the forecast errors for the selected models.

Forecast error	Forecasting models		
	Brown model *	ARMA model	Spectral analysis
MAE	17	12.22	4.62
MAPE	1.36	1.16	0.39

/* the smoothing constant is 0.9

5. Conclusions

Requirements for prognostic models based on time series largely based on the dependencies that occur between considered time series values. Currently. computer technology can be used to predict future events using more sophisticated techniques.

From the analysis of the data. based on the calculated absolute and relative errors of the forecast it can be concluded that the Brown model is currently one of the least accurate methods of prediction. More advanced methods like ARMA models show higher accuracy compared to models using exponential smoothing method. One of the most accurate are the methods that use spectral analysis to analyze time series in the frequency domain. Based on the results of average forecast errors. we can conclude that the prediction based on the spectral analysis approach is 3 times more accurate than the ARMA model. and 3 to 4 times compared to the Brown model. The results of this analysis indicate a higher accuracy of spectral analysis. Presented observations can lead to the creation of new. even more qualitative methods of forecasting.

References

[1] Anderson T. W.. 1971. The Statistical Analysis of Time Series. John Wiley & Son. Canada. p. 386.
[2] Bisgaard S.. Kulahci M.. 2011. Time Series Analysis and Forecasting by Example. A John Wiley & SonsNew Jersey. p. 149.
[3] Burkhard B.. 2000. Spectral Analysis and Filter Theory in Applied Geophysics. Springer-Verlag Berlin Heidelberg. p. 101.
[4] Chattopadhyay A. K.. Chattopadhyay T.. 2014. Statistical Methods for Astronomical Data Analysis. Springer Science+Business. New York. p. 232.
[5] Clements M. P.. Hendry D. F.. 2001. Forecasting Non-stationary Economic Time Series. The MIT Press. London. England.
[6] Ehrendorfer M.. 2012. Spectral Numerical Weather Prediction Models. Society for Industrial and Applied Mathematics. Philadelphia.
[7] Hearn G.. Metcalfe A.. 1995. Spectral Analysis in Engineering: Concepts and Case Studies. Butterworth-Heinemann. London.
[8] Warner R. M.. 1998. Spectral Analysis of Time-series Data. Guilford Press. New York London. p. 39.
[9] Tchrakian T. T.. Basu B.. O'Mahony M.. 2012. Real-Time Traffic Flow Forecasting Using Spectral Analysis. IEEE Transactions on Intelligent Transportation Systems. Vol. 13. Issue 2. p. 519-526.
[10] Zhang Y.. Zhang Y.. Haghani A.. 2014. A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model. Transportation Research Part C-Emerging Technologies. Special Issue. 43. p. 65-78.