EPT 1/2019 (45) | ISSN: 1644-0501 | www.wnus.edu.pl/ept | DOI: 10.18276/ept.2019.1.45-01 | 7–16

ASSESSMENT OF APPLICABILITY OF FORECASTING METHODS OF PASSENGER TRAFFIC BY AIR AS A KEY FACTOR IN TOURISM DEVELOPMENT IN POLAND

Kamila Bednarz-Okrzyńska

University of Szczecin
Faculty of Management and Economics of Services
Department of Quantitative Methods e-mail: kamila.bednarz@wzieu.pl

JEL Codes C01, C22, C53

Keywords tourism, airports, passenger traffic, forecast, seasonality

Abstract The author of the paper focuses on the key factor in tourism development which is air passenger transport. It considers the forecast for five airports which in the first half year attended to the largest number of passengers as well the overall passenger traffic in Poland. The forecast was made for the first two quarters of 2018 on the basis of the quarterly data from the period of 2011–2018. The method applied in the paper consisted in making calculations for varying number L of recent years taken into account in the model. The model of additive and multiplicative seasonality was considered. The research problem was formulated as the question whether the smallest value of residual standard deviation corresponds to the smallest value of error MAPE (Mean Absolute Percentage Error). By applying the method of trends of univariate periods the following were determined: residual standard deviation, relative error *ex ante* VA and forecast error MAPE. The purpose was to answer the question whether the minimum value of MAPE corresponds to the minimum value of standard deviation S, or to the smallest value of error *ex ante.* In the case of the linear trend, a simplified method of "three points" was proposed, whereas for the parabolic trend, a method of "four points" was proposed. It was found that the method of trends of univariate periods yields the forecast error *ex post* comparable to the one of classic methods of seasonality analysis (including the Klein method).

Introduction

An extremely important determinant of tourism development is the development of transport. Assuming that the most typical feature of tourism is tourists' change of location, the transport that enables such a movement constitutes the necessary condition for tourist activity (Milewski, 2010a, p. 219).

One of the crucial factors in tourist movement development these days has been the expansion and application of air traffic to large-scale passenger traffic. The passenger air traffic has enabled travelling far longer distances in a far shorter time, hence has become the main means of transport in long distance foreign travel, including intercontinental travel. The share of air transport in passenger transport increases with the travel distance. In the case of shorter distances the role of air traffic is important if the air routes run over hardly accessible areas (such as islands or bays) (Milewski, 2010b, p. 117).

The topic of interrelation between air passenger transport and tourism has been thoroughly covered in numerous papers (Briggs, 2004, pp. 117–131; Chung, Whang, 2011, pp. 1335–1340; Turton, 2004, pp. 69–78; Wheatcroft, 1998, pp. 159–179). The impact of air transport development on tourism development can be analyzed from various viewpoints, including the emergence of new tourist attractions, development of new tourist facilities (especially tourist accommodation), tourist services and tourist traffic. It may also raise the degree of transport accessibility of a given region or town, which plays an important role not only in tourist development but also in a general socio-economic development (Kowalczyk, 2013, p. 61).

According to United Nations World Tourism Organization assessments, in 2017, 57% of tourists included in the reports on the world tourist traffic used air transport (UNWTO, 2018). With respect to island countries (located mainly in the Pacific and Indian Ocean as well as in the region covering the Caribbean Sea and the Gulf of Mexico), air traffic constituted almost 100%. And in the case of continental countries, the importance of air traffic for tourist development is becoming more and more vital.

The dynamic expansion of air sector and its dramatic impact on tourist development constitutes a serious challenge for accurate passenger traffic forecasting. And this is where the selection of a proper method of forecasting gains a special significance. Therefore, the author's main objective is a proper assessment of the methods of passenger traffic forecasting for air transport.

This paper includes a forecast for five airports which in the first half year attended to the largest number of passengers. These are: Warsaw Chopin Airport, Kraków-Balice, Gdańsk Lech Wałęsa Airport, Katowice-Pyrzowice and Wrocław-Starachowice. Additionally, a prediction of the number of attended passengers by all fifteen Polish airports was made. It was made for the first two quarters of 2018 on the basis of the quarterly data from the period of 2011–2018 (ULC, 2018). In the second section the seasonal fluctuations were dealt with adopting the method of seasonal variation coefficients as well as the harmonic analysis method. In the case of additive seasonal coefficients, the results were compared with the results obtained with the Klein method. While applying the method of multiplicative seasonal coefficients, the harmonic analysis was used. The paper comprises the method which consists in making calculations for varying number L of recent years taken into account in a model. For L falling within interval $<2; 7>$ the residual standard deviation S was determined as well as the value of error *ex post* defined by means of MAPE given by formula:

$$MAPE = \frac{1}{2} \cdot \sum_j \left| \frac{YP_j - yr_j}{yr_j} \right| \cdot 100 \qquad (1)$$

where:

YP – forecast value,

yr – realization value.

The research problem was formulated as the question whether the smallest value of residual standard deviation (the best fit of a model to empirical data) corresponds to the smallest value of error MAPE.

The third section is devoted to the method of trends of univariate periods. In this case, the data included in the study cover the period 2010–2017. As in the previous chapter, calculations were also made for varying number L of recent years. For the determined trend the following were calculated: residual standard deviation, relative error *ex ante* VA and forecast error MAPE. The goal was to answer the question whether the minimum value of MAPE corresponds to the minimum value of the standard deviation S, or to the minimum value of error *ex ante.* In the case of the linear trend, a simplified method of "three points" was proposed, whereas for the parabolic trend, a method of "four points" was proposed.

Analysis of seasonal variability of passenger traffic for selected Polish airports

The starting point in the method of seasonal coefficients is determining the form of a trend. In the paper, a linear trend was assumed, defined by formula:

$$yL_i = \alpha + \beta \cdot t_i \qquad (2)$$

where: $\alpha = ys - ts \cdot \beta$.

As far as parameter beta is considered, its proper form can be found in papers (Maciąg, Stępniak, 2008, pp. 38–47; Zieliński, 2002). This form depends on number L of years of a model, and for quarterly data is given by formulae:

Where:

L = 2

$$\beta = \frac{\sum_{i=5}^{8} y_i - \sum_{i=1}^{4} y_i}{16} \qquad (3a)$$

for L = 3

$$\beta = \frac{\sum_{i=9}^{12} y_i - \sum_{i=1}^{4} y_i}{32} \qquad (3b)$$

for L = 4

$$\beta = \frac{3 \cdot \left(\sum_{i=13}^{16} y_i - \sum_{i=1}^{4} y_i \right) + \sum_{i=9}^{12} y_i - \sum_{i=5}^{8} y_i}{160} \qquad (3c)$$

for L = 5 $$\beta = \frac{2 \cdot \left(\sum_{i=17}^{20} y_i - \sum_{i=1}^{4} y_i \right) + \sum_{i=13}^{16} y_i - \sum_{i=5}^{8} y_i}{160} \quad (3d)$$

The general form (Maciąg, Stępniak, 2008, pp. 38–47; Zieliński, 2002) can be written in the following way:

$$\beta = \frac{\sum_{i=0}^{L-1} (2 \cdot i + 1 - L) \cdot \sum_{k=0}^{K-1} y_{k+K \cdot i}}{\frac{K^2}{6} \cdot L \cdot (L^2 - 1)} \quad (4)$$

where:

n – number of observations,

$L = n/K$ – number of years,

K – number of base periods (for quarterly data $K = 4$).

Starting with the deviation of the value of observations from the trend:

$$z_i = y_i - yL_i \quad (5)$$

the method of additive seasonal coefficients was applied.

Taking into account the model of multiplicative seasonality, the ratio of the value of observations to the value of trend was derived:

$$z_i = \frac{y_i}{yL_i} \quad (6)$$

And subsequently harmonic parameters

$$a_0 = \frac{1}{n} \cdot \left(\sum_i z_i \right) \quad a_1 = \frac{2}{n} \cdot \sum_i z_i \cdot \cos\left(\frac{2 \cdot \pi}{K} \cdot i \right)$$
$$b_1 = \frac{2}{n} \cdot \sum_i z_i \cdot \sin\left(\frac{2 \cdot \pi}{K} \cdot i \right) \quad a_2 = \frac{1}{n} \cdot \left(\sum_i z_i \cdot \cos(\pi \cdot i) \right) \quad (7)$$

where:

$K = 4$,

$i = 1, 2, \ldots, n$.

In the case of the quarterly data (K = 4), the values of four harmonics can be determined (Zieliński, 2002). The multiplicative model of seasonality is given by:

$$ytH_m = yL_m \cdot yH_m \tag{8}$$

where: $yH_m = a_0 + a_1 \cdot \cos\left(\frac{2 \cdot \pi}{K} \cdot m\right) + b_1 \cdot \sin\left(\frac{2 \cdot \pi}{K} \cdot m\right) + a_2 \cdot \cos(\pi \cdot m)$; $m = 1, 2, \ldots, n+2$.

Subsequently, the value of forecast $YP_j = ytH_j$ is determined, where: j = n+1, n+2 and the residual values: $e_i = ytH_i - y_i$.

Figure 1 presents the values of error MAPE (equation (1)) and the values of standard deviation of the model derived from formula:

$$S = \sqrt{\frac{\sum_i (e_i)^2}{L-5}} \tag{9}$$

Figure 1 presents the results of the forecast of the number of attended passengers at fifteen Polish airports as a function of the number of years L included in the model. Solid line with boxes MA represents error MAPE (in %) for the model with the additive seasonality. Solid line with circles MM represents error MAPE for the model with the multiplicative seasonality. Dotted lines represent a rescaled standard deviation: with boxes – the additive model (SA), and with circles – the multiplicative model (SM). Rescaled values S were obtained from the following formulae: $SA' = 0.00001 \cdot SA$; $SM' = 0.00001 \cdot SM$. The standard deviation (dotted line with boxes) takes the smallest value $SA = 261686$ for L = 2, for which MA = 3.82%. For the multiplicative model, the following can be read: L = 2, SM = 74396 and MM = 0.82%. It should be noticed that the proposed criterion determines the minimum value of error MAPE = 0.82%.

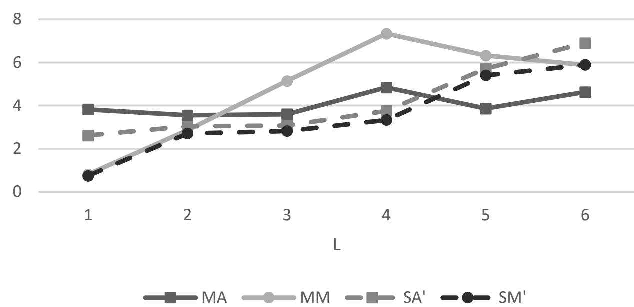

Figure 1. The results of the forecast of passenger traffic at fifteen Polish airports as a function of the number of years L included in the model

Source: author's own study.

A parallel analysis was carried out for the five aforementioned airports. In the case of Warsaw Chopin Airport the proposed criterion yields the value of error MAPE = 2.95% for the additive model (L = 4) and error MAPE = 3.56% for the multiplicative model (L = 2), where the minimum value MAPE = 1.22% (multiplicative model, L = 3). For Kraków-Balice Airport, following the adopted criterion, MAPE = 2.47% (model multiplicative, L = 2) is obtained and MAPE = 1.29% (additive model, L = 2), which corresponds to the minimum value of error. Calculations made for

Gdańsk Airport, following the same criterion, yield MAPE = 2.95% (additive model, L = 4) and MAPE = 3.56% (multiplicative model, L = 2), whereas the minimum value MAPE = 1.22% (multiplicative model, L = 3). In the case of Katowice-Pyrzowice Airport, the criterion yields MAPE = 9.92% (additive model, L = 4) and MAPE = 4.99% (multiplicative model, L = 2), which corresponds to the minimum value of error. For Wrocław-Starachowice Airport, MAPE = 8.02% (additive model, L = 5) and MAPE = 6.77% (multiplicative model, L = 2) are obtained, whereas the minimum value MAPE = 2.40% is obtained for the multiplicative model, L = 5. This means that the assumed criterion of the minimum value of residual standard deviation determined the minimum value of error MAPE for the following airports: Kraków-Balice and Katowice-Pyrzowice. However, for the other airports – Warsaw Chopin Airport, Gdańsk Lech Wałęsa Airport and Wrocław-Starachowice Airport – the proposed criterion does not detect the minimum value of MAPE.

It should be noticed that the results obtained for the additive model – both the method of additive coefficients and harmonic analysis – are identical to the results of the Klein method. However, based on our study, the formula for the beta parameter of the linear trend which is often applied in the literature (Cieślak, 2001; Zeliaś, Pawełek, Wanat, 2003) should in fact not be used:

$$\beta = \frac{\sum_{i=1}^{n} y_i \cdot (t_i - ts)}{\sum_{i=1}^{n} (t_i - ts)^2} \qquad \alpha = ys - \beta \cdot ts \tag{10}$$

To take one example, applying equation (10) and the method of additive coefficients, for Kraków-Balice Airport MAPE = 2.52% is obtained, and by applying equations (2) and (3), almost twice as small an error is obtained, MAPE = 1.29%.

Method of trends of univariate periods

A model of trends of univariate periods is one of the simplest methods of forecasting seasonal phenomena (Gierałtowska, 2000, pp. 3–12; Tłoczyński, 2009, pp. 247–261). For instance, if a forecast is related to the first quarter of a given year, we consider the values of observations from the first quarters of previous years, and on this basis we determine a trend. In this paper the following trends are considered: linear, parabolic, logarithmic, power, exponential and hyperbolic. The most useful proved the linear and parabolic trend, for which the analysis was made. Calculations were made for the data gathered in the first and second quarters in years 2010–2017. On the basis of the obtained trends, forecast value YP was determined, as well as the residual standard deviation S, relative error *ex ante* AV and forecast error MAPE. The residual standard deviation was derived from equation:

$$SL = \sqrt{\frac{\sum_i (e_i)^2}{L-2}}; \qquad SP = \sqrt{\frac{\sum_i (e_i)^2}{L-3}} \tag{11}$$

where:

SL – for linear trend,
SP – for parabolic trend,
e_i– residual error of a model,
L – number of considered observations (number of years of the model).

Similarly as described earlier, number L of recent observations was considered. In view of equation (11) for the linear trend, number L fell within interval <3; 8> and for the parabolic trend, L fell within interval <4; 8>. The second column SL in Table 1 includes values MAPE which correspond to the smallest value of the standard deviation for the linear trend. The third column AVL includes the values of error MAPE which correspond to the smallest value of the relative error *ex ante* (linear trend). The fifth and sixth columns (SP, AVP) are analogous to the second and third but have been derived for the parabolic trend. The eighth (ninth) column MA (MM respectively) includes the values of error MAPE determined in second section for the additive (multiplicative respectively) seasonality model. Following equation (11), the smallest number of considered years is L = 3 for the linear trend, and L = 4 for the parabolic trend. Therefore, the following simplified methods can be proposed: for the linear trend – the “three point” method (L = 3), the results of which are presented in Table 1 Column 3P; and for the parabolic trend – the “four point” method (L = 4), the results of which are presented in Column 4P.

Table 1. Values of error MAPE obtained with the method of univariate periods trends

	SL	AVL	3P	SP	AVP	4P	MA	MM
Gdańsk	2.35	2.35	1.40	5.05	2.31	5.05	2.95	3.56
Katowice	15.72	15.72	13.96	10.89	12.32	10.40	9.916	4.99
Kraków	4.09	4.09	4.09	2.12	2.07	1.31	1.29	2.47
Warszawa	5.93	13.45	3.61	7.00	7.00	8.33	6.44	3.76
Wrocław	7.72	7.72	6.23	3.31	3.31	4.88	8.02	6.77
Poland	5.87	5.87	4.71	0.83	0.83	1.07	3.82	0.82

Source: author’s own study.

For method 3P the following equation applies:

$$P3 = \frac{1}{3}\cdot\left(4\cdot y_n + y_{n-1} - 2\cdot y_{n-2}\right) \tag{12}$$

where:

P3 – forecast value,
n – number of all observations.

By analogy, for the "four points" method (parabolic trend) the forecast value is given by:

$$P4 = \frac{1}{4} \cdot (9 \cdot y_n - 3 \cdot y_{n-1} - 5 \cdot y_{n-2} + 3 \cdot y_{n-3}) \tag{13}$$

where:

P4 – forecast value.

For the purpose of this paper, the ranking of the six methods was made (with the exception of the last two columns): for each airport points were awarded in the following way: 6 points to the method yielding the smallest MAPE, 5 points to the method with the second smallest MAPE, and so on. The results of the ranking are: VA – 29, 4P – 24.5, SP – 23.5, 3P – 23, SPL – 15, VAL – 11.

The best method proved VA, i.e. the method based on the minimum value of error *ex ante* for the parabolic trend. The worst results have been obtained for the linear trend for the minimum error *ex ante* method and the minimum value of the standard deviation. Similar results have been obtained for the "four point" method, the "three point" method and the method of the minimum value of the standard deviation for the parabolic trend. Taking into account the results of the ranking as well as the simplicity of calculations (equations (12) and (13)), methods 3P and 4P should be recommended.

Conclusions

The role of contemporary tourism in the economic development is substantial and it is still growing. Countries which manage to tap their potential of attractiveness or competently design their tourism product observe a substantial level of revenues from tourism sector contributing to GDP (Gross Domestic Product) level.

Tourism is based on people changing their location in order to reach attractive destinations. Transport that enables such a movement constitutes the necessary condition for tourist activity. Taking into consideration three stages of travel (reaching the destination, stay and return home), the two of them are all about transportation services.

There is a synergy between transport and tourism. On the one hand, the ease of reaching a particular destination – determined by transportation services – is a key factor of the development of any tourist attraction. On the other hand, the transport sector may benefit from tourism, which may generate additional demand on transportation services.

The role of air transport in tourist traffic service has significantly increased since both charter flights and changes in the aviation law were introduced. Liberalization of air transport in Poland has resulted in vital changes in the air transport market, which in turn has had both direct and indirect impact on the whole tourist economy, among others by contributing to development of new trends in tourist travel.

In view of the above observations, it was considered important to make an assessment of methods of forecasting passenger air traffic, which constitutes such a vital factor of tourism development.

In the second chapter the models with the additive and multiplicative seasonality were considered. Application of proper equations defining the linear trend (equations (2) and (3)) in the case of a model with the additive seasonality allowed obtaining the results identical with the results obtained with the Klein method.

As far as the question put forward in the Introduction is considered, the question whether the smallest value of the residual standard deviation corresponds to the smallest value of error MAPE, the conclusions are as follows. The calculations were made for 6 values of number L of years, which implies the probability 1/6. By taking into account two models (additive and multiplicative) the probability 1/3 is obtained. Furthermore, the application of the criterion of the smallest value of standard deviation, a positive result was obtained for three out of six cases, which corresponds to 50% of success. Despite its clear advantage over the derived probability of 33.33%, the criterion cannot be accepted as a method of obtaining the smallest value of forecast error *ex post*.

Summing up the results obtained in section 3, it should be noticed that for three in six considered airports, the method of univariate periods trends (Gdańsk, Warszawa, Wrocław) led to a smaller value of MAPE than the full analysis of seasonality conducted in Chapter 2, and for one airport (Kraków-Balice) it led to similar values of error (1.30% and 1.31%).

It should be stressed that in the process of decision-making, the selection of a proper prediction method is of particular importance. The selected methods of forecasting passenger traffic in air transport which have been described in this paper and the presented results of predictions constitute a useful decision-making tool both in transport and tourist sector.

References

Briggs, D. (2004). Tourism development and airlines in the New Millennium: an operations management perspective. In: L. Lumsdon & S. J. Page (eds.), *Tourism and transport: issues and agenda for the New Millennium* (pp. 117–131). Amsterdam–San Diego–Oxford–London: Elsevier.

Chung, J. Y., Whang, T. (2011). The impact of low-cost carriers on Korean Island tourism. *Journal of Transport Geography*, *6* (19), 1335–1340.

Cieślak, M. (ed.) (2001). *Prognozowanie gospodarcze. Metody i zastosowanie.* Warsaw: Wydawnictwo Naukowe PWN.

Gierałtowska, U. (2000). Metoda trendów jednoimiennych okresów a prognozowanie brakujących danych. *Wiadomości Statystyczne*, *9*, 3–12.

Kowalczyk, A. (2013). Relacje zachodzące między rozwojem transportu lotniczego a rozwojem turystyki. In: R. Pawlusiński (ed.), *Współczesne uwarunkowania i problemy rozwoju turystyki* (pp. 61–72). Cracow: IGiGP UJ.

Maciąg, K., Stępniak, C. (2008). Uproszczone wzory na estymację trendu i efektów sezonowych. *Przegląd Statystyczny*, *2* (55), 38–47.

Milewski, D. (2010a). Rola transportu w kształtowaniu atrakcyjności turystycznej regionu. *Zeszyty Naukowe Uniwersytetu Szczecińskiego*, *582*, *Ekonomiczne Problemy Usług*, *48,* 219–230.

Milewski, D. (2010b). Transport lotniczy w obsłudze ruchu turystycznego – analiza zależności i prognoza zmian na przykładzie województwa zachodniopomorskiego. In: M. Michałowska (ed.), *Efektywność transportu w teorii i w praktyce* (pp. 117–128). Katowice: Wydawnictwo Akademii Ekonomicznej.

Stańko, S. (2013). *Prognozowanie w agrobiznesie. Teoria i przykłady zastosowania.* Warsaw: Wydawnictwo SGGW.

Tłoczyński, D. (2009). Wpływ polityki otwartego nieba na funkcjonowanie regionalnych portów lotniczych w Polsce. In: B. Liberadzki (ed.), *Liberalizacja i deregulacja transportu w UE. Oczekiwania i doświadczenia* (pp. 247–261). Warsaw–Poznań: ILiM.

Turton, B. (2004). Airlines and Tourism Development: the Case of Zimbabwe. In: L. Lumsdon, S.J. Page (eds.), *Tourism and Transport: Issues and Agenda for the New Millennium* (pp. 69–78). Amsterdam–San Diego–Oxford–London: Elsevier.

UNWTO (2018). *Tourism highlights 2018 Edition*. Retrieved from: https://www.e-unwto.org/doi/pdf/10.18111/9789284419876 (15.12.2018).

Urząd Lotnictwa Cywilnego (2018). *Statystyki i analizy rynku transportu lotniczego*. Retrieved from: http://www.ulc.gov.pl/pl/regulacja-rynku/statystyki-i-analizy-rynku-transportu-lotniczego/3724-statystyki-wg-portow-lotniczych (26.10.2018).

Wheatcroft, S. (1998). The Airline Industry and Tourism. In: D. Ioannides, K.G. Debbage (eds.), *The Economic Geography of the Tourist Industry. A Supply-side Analysis* (pp. 159–179). London–New York: Routledge.

Zeliaś, A., Pawełek, B., Wanat S. (2003). *Prognozowanie ekonomiczne. Teoria, przykłady, zadania.* Warsaw: Wydawnictwo Naukowe PWN.

Zieliński, Z. (2002). *Analiza ekonomicznych procesów stochastycznych. Pisma wybrane.* Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Ocena przydatności metod prognozowania lotniczego ruchu pasażerskiego jako kluczowego czynnika rozwoju turystyki na przykładzie Polski

Słowa kluczowe turystyka, lotniska, ruch pasażerski, prognoza, sezonowość

Streszczenie

W niniejszej pracy skupiono się na kluczowym czynniku rozwoju turystyki jakim jest pasażerski transport lotniczy. Rozpatrzono prognozę dla pięciu portów lotniczych, które w pierwszym półroczu wykazały się największą liczbą obsłużonych pasażerów, a także łączny ruch pasażerski w Polsce. Prognozę wykonano dla pierwszych dwóch kwartałów 2018 r. bazując na danych kwartalnych za lata 2011–2018. Zastosowano metodę polegającą na wykonaniu obliczeń dla zmieniającej się liczby L ostatnich lat uwzględnionych w modelu. Rozpatrzono model sezonowości addytywnej oraz sezonowości multiplikatywnej. Postawiony problem badawczy wyrażał się pytaniem, czy najmniejszej wartości odchylenia standardowego reszt odpowiada najmniejsza wartość błędu MAPE. Stosując metodę trendów jednoimiennych okresów wyznaczono: odchylenie standardowe reszt, błąd względny *ex ante* oraz błąd prognozy MAPE. Szukano odpowiedzi na pytanie: czy wartość minimalna MAPE jest zgodna z najmniejszą wartością odchylenia standardowego S, czy też z najmniejszą wartością błędu *ex ante*. W przypadku trendu liniowego zaproponowano uproszczoną metodę „trzech punktów", natomiast dla trendu parabolicznego zaproponowano metodę „czterech punktów". Stwierdzono, że metoda trendów jednoimiennych okresów prowadzi do porównywalnych błędów prognozy *ex post* jak klasyczna metoda analizy sezonowości (w tym metoda Kleina).