European Research Studies Journal
Volume XXVI, Issue 4, 2023

pp. 373-381

Analysis and Forecasting of the Number of Overnight Visitors in Poland

Submitted 02/10/23, 1st revision 14/10/23, 2nd revision 16/11/23, accepted 30/11/23

Rafał Klóska[1]

Abstract:

Purpose: *The complexity of tourism development means that its issues can be considered from various points of view. The aim of the article is to utilize suitable statistical methods to depict and forecast the number of tourists (both domestic and foreign) staying overnight in tourist accommodations in Poland.*
Design/Methodology/Approach: *For the purposes of own research to describe changes in the number of overnight visitors in tourist accommodations in Poland, we employed index methods, and classic trend models paired with ex-ante prediction errors for forecasting.*
Findings: *Over the years, countries worldwide, including Poland, have experienced dynamic growth in tourism – a sector increasingly influencing the economy. The COVID-19 pandemic brought unprecedented travel restrictions, leading to the downfall of several travel agencies, accommodations, and restaurants, as well as a significant reduction in the number of tourists staying in accommodations. However, through adept statistical methods, we've depicted the shifts in overnight tourist numbers in Poland and presented a short-term prediction of this trend.*
Practical Implications: *Statistical methods prove beneficial in researching tourism development. The theoretical considerations and the results of analyzes and forecasts presented in this paper allow for a broader characterization of the focal phenomenon, aiding socio-economic decision-making and informing future developmental strategies.*
Originality/Value: *Quantitative methods are now widely used in empirical research, including in the field of tourism, and their usefulness is beyond dispute. They are widely used in the processes of analyses, diagnoses and forecasts, and with their use, the description, assessment and prediction of various socio-economic phenomena become more precise. The article draws attention to the fact that appropriate statistical methods can be successfully used to characterize changes in the number of overnight visitors in Poland.*

Keywords: *Tourism, statistical methods, forecasting.*

JEL codes: *O38, C43, L83.*
Paper type: *Research article.*

Acknowledgment: *This project is financed within the framework of the program of the Minister of Science and Higher Education under the name "Regional Excellence Initiative" in the years 2019 - 2022; project number 001/RID/2018/19; the amount of financing PLN 10,684,000.00.*

[1]*Asocc. Prof., University of Szczecin, Institute of Spatial Management and Socio-Economic Geography, rafal.kloska@usz.edu.pl;*

1. Introduction

Tourism significantly bolsters the economy, influencing it on both macro and regional scales (Kurek, 2007; Meyer, 2010; Kaczmarek, Stasiak, and Włodarczyk, 2010; Panasiuk and Sidorkiewcz, 2017; Uysal, Sirgy, and Kruger, 2018; Panasiuk, 2019; Adamopoulos and Thalassinos, 2020; Szmytke, 2022).

Poland boasts numerous attractions, including the Baltic Sea beaches, Masurian lakes and three mountain ranges like the Świetokrzystkie Mountains, Sudetes, and Carpathians which attract and encourage many tourists, both domestic and foreign, to relax every year.

Additionally, enhanced infrastructure, improved accomodation quality and their availability, including amenities resulting from universal design, are just some of the factors that have a positive impact on the broadly understood development of tourist services, tourist potential and an increase in the use of offered accommodations.

Unfortunately, the COVID-19 pandemic adversely impacted the functioning and proper development of many areas of socio-economic life, including tourism. This article attempts to describe changes in the number of tourists (both domestic and foreign) staying overnight in tourist accommodations over the last dozen or so years in Poland. Appropriate statistical methods were used to enable analysis and short-term forecast of the phenomenon named in the title.

2. Literature Review

The described research methodology, the so-called initial analysis of the dynamics of phenomena is often a starting point for further considerations related to forecasting (Markidakis, Wheelwright, and Hyndman, 1998; Hyndman and Athanasopoulos, 2018). Both Cieślak (2001, p. 18) and Zeliaś (1997, p. 16) understand forecasting as a rational and at the same time scientific prediction of future events.

The specific result (usually numerical) of this process is a forecast for a given period (moment) in the future, called the forecast horizon. This is how the forecast is defined, among others, by Pawłowski (1973, p. 14-15). For Dittmann (2000, p. 19), a forecast "is a statement referring to a specific future, formulated using the achievements of science, empirically verifiable, uncertain, but accepted."

Nowak (1998, p. 10), Hanias *et al.* (2007a; 2007b) understand it as a judgment about the expected development of phenomena and processes in the future based on scientific foundations. Cieślak (2001, p. 18-20), Czerwiński and Guzik (1980, p. 19) and Gajda (2001, p. 135) assumed that in the process of prediction (also called forecasting), a judgment about the future state of the phenomenon under study will be formulated in a rational manner based on scientific foundations.

The forecast, as a result of prediction, is an estimate of the probable implementation of the forecasted variable based on the knowledge about its current course. The skillful use of an appropriate forecasting method is extremely important in the forecasting process. The choice of a specific method of obtaining forecasts depends on the forecaster, but it should be remembered that it determines the final prediction results.

It should be emphasized that properly constructed forecasts can be useful in making appropriate economic decisions. Most of the cited authors, however, caution against approaching the results of the prediction process uncritically or equating them with a given decision. According to Dittmann (2000, p. 19), "an illusion is the belief that we can completely fathom the secrets of the future."

Czerwiński and Guzik (1980, p. 40-42) state that although "it is tempting (...) to reduce the problem of a "good" forecast to the problem of choosing a rational action (or a rational decision)," "there is no basis for always identifying judgments about the future with actions or decisions resulting from making these judgments." Pawłowski succintly states that forecasts only allow for a rational choice of the best course of action in the future (Pawłowski, 1973, p. 14).

3. Research Methodology

The simplest statistical measures used in the analysis of the dynamics of phenomena include (Klóska and Czyżycki, 2019, p. 14-17):

• absolute increases – illustrating the absolute difference observed in the level of the examined phenomenon in two different periods (moments),
• dynamics indicators (called indices) – understood as the ratio of the magnitude of the same phenomenon in two different periods (moments).

The indicated measures, depending on the type of basis for comparison, may be chain or single-base. While in the first case the basis for comparison is always the period (moment) immediately preceding the examined one, the choice of a constant basis for single-base measures depends on the purpose of the study. Usually, the first period (moment) in the time series is assumed, but if there are justified economic reasons, the basis for comparisons may be different.

If t denotes the numbers of subsequent time units and y_t denotes the size of the studied phenomenon in period (moment) t, then the appropriate formulas used to calculate the above-mentioned measures are as follows:

- chain absolute growth:

$$\Delta y_{t/t-1} = y_t - y_{t-1} \qquad (1)$$

- single-base absolute growth:

$$\Delta y_{t/o} = y_t - y_o \quad (2)$$

- chain index:

$$i_{t/t-1} = \frac{y_t}{y_{t-1}} \quad (3)$$

- single-base index:

$$i_{t/o} = \frac{y_t}{y_o} \quad (4)$$

The most frequently used method of identifying trends – the analytical method – consists in the fact that the development tendency is expressed using a specific mathematical function, in which the dependent variable is the level of the phenomenon observed over time and the independent variable – the time variable. The time series model then has the form:

$$Y_t = f(t) + u_t \quad (5)$$

where:
Y_t - a variable showing the level of the studied phenomenon over time,
$f(t)$ - a specific mathematical form of the function (trend model),
u_t - a random component.

After estimation using the adopted estimation method, the model will take the form:

$$\hat{Y}_t = f(t) \quad (6)$$

When assessing the quality of the built models, appropriate measures and statistical tests are used, i.e. specific methods of model verification. It is worth pointing out here that, among others, (Czyżycki and Klóska, 2011, p. 44-48):

- low values of the convergence coefficient (consistency, mismatch, indetermination) and, therefore, high values of the determination coefficient (fit) are desirable,
- in the case of statistical significance of the multiple correlation coefficient, the degree of fit of the model to the empirical data can be considered sufficiently high,

- the smaller the value of the standard deviation of the residual component Se, the less the theoretical and empirical values of the explained variable differ on average,
- the better the model, the lower the value of the coefficient of random variation VSe, which shows the share of the standard deviation of the random component in the average value of the dependent variable,
- the structural parameters of the estimated model should be statistically significant, but when using the Student's t-test in practice, the significance of a parameter is usually determined by the inequality $t \geq 2$.

Knowledge of the forecasting model is one of the basic assumptions of the prediction theory (Pawłowski, 1973, p. 39-41). Equally important is the need to determine the size of the forecast error. This need arises from the fundamental postulates of prediction theory. The content of the first postulate requires us not only to provide a forecast as a result of the prediction process but also to give an appropriate measure of its accuracy.

Achieving the best possible value of this measure ensures high forecasting efficiency. The pursuit of such efficiency stems from the second postulate of prediction theory.

Forecast accuracy measures are divided into ex-ante and ex-post measures. The former concerns determining the accuracy of the forecast and are estimated at thetime (period) of forecast construction. Ex-post measures aim to determine the accuracy of forecasts because they are prepared after the time covered by the forecast has passed. Forecasts based on classic trend models are made by extrapolating, substituting subsequent values of the time variable t, etc.

The average absolute error of the ex-ante forecast can be calculated based on the formula:

$$V_T = \sqrt{S_e^2 \cdot \left[\frac{(T-\bar{t})^2}{\sum_{t=1}^{n}(t-\bar{t})^2} + \frac{1}{n} + 1 \right]} = S_e \cdot \sqrt{\frac{(T-\bar{t})^2}{\sum_{t=1}^{n}(t-\bar{t})^2} + \frac{1}{n} + 1} \qquad (7)$$

where:

S_e^2 – variance of the random component of the trend model,

S_e – standard deviation of the random component of a linear trend model,

T – value of the time variable in the forecast period.

The average absolute error of the ex-ante forecast is expressed in nominal units of the phenomenon under study. To express the expected forecast error as a percentage, the average relative ex-ante forecast error should be calculated, as expressed by the formula:

$$V_T^* = \frac{V_t}{y_{tP}} \cdot 100\% \qquad (8)$$

4. Results

Using the formulas for the single-base and chain measures described above (see formulas 1-4), Table 1 presents the results of calculations of absolute increases and individual indices of the number of tourists (both domestic and foreign) staying overnight in tourist accommodation facilities in Poland.

Table 1. *Analysis of the dynamics of the number of overnight visitors in Poland*

Year	Y_t	$\Delta y_{t/t-1}$	$\Delta y_{t/o}$	$i_{t/t-1}$	$i_{t/o}$
2011	21 476 616	-	0	-	1
2012	22 635 388	1 158 772	1 158 772	1,0540	1,0540
2013	23 401 138	765 750	1 924 522	1,0338	1,0896
2014	25 083 978	1 682 840	3 607 362	1,0719	1,1680
2015	26 942 056	1 858 078	5 465 440	1,0741	1,2545
2016	30 108 308	3 166 252	8 631 692	1,1175	1,4019
2017	31 989 344	1 881 036	10 512 728	1,0625	1,4895
2018	33 895 930	1 906 586	12 419 314	1,0596	1,5783
2019	35 668 091	1 772 161	14 191 475	1,0523	1,6608
2020	17 878 969	-17 789 122	- 3 597 647	0,5013	0,8325
2021	22 198 972	4 320 003	722 356	1,2416	1,0336
2022	34 249 004	12 050 032	12 772 388	1,5428	1,5947

Source: *Own calculations based on Central Statistical Office data (www.strateg.stat.gov.pl; 29.09.2023).*

A linear trend model was built to describe changes in the number of overnight visitors in Poland in 2011 to 2019 (before the COVID-19 pandemic) (Table 2).

Table 2. *Results of modeling in STATISTICA the linear development tendency of the number of overnight visitors in Poland*

	R= ,99049338 R^2= ,98107713 F(1,7)=362,92 p<,00000 Se=764059,46 Vse=0,027375					
N=9	b*	Bł. std. z b*	b	Bł. std. z b	t(7)	p
alfa0			18515516	555076,3	33,35670	0,000000
alfa1	0,990493	0,051993	1879138	98639,7	19,05053	0,000000

Source: *Own calculations based on Central Statistical Office data (www.strateg.stat.gov.pl; 29.09.2023).*

5. Discussion

Based on the results presented in Table 1, the following conclusions can be drawn:

- in the years 2011-2019, the number of overnight visitors in Poland showed a clear increasing tendency. The collapse caused by travel restrictions and restrictions due to the COVID-19 pandemic resulted in the number of analyzed tourists being lower by as much as 17,789,122 people in 2020, i.e., 49.87% compared to the previous year and by 3,597,647 people, i.e., 16.75% compared to 2011 (the initial research period, adopted as the basis for comparisons of single-basic measures);
- in 2011-2022, compared to the previous year, the largest increase in the number of analyzed tourists occurred in 2022 (by 12,050,032 people, i.e., 54.28%), and the smallest increase was in 2013 (by 765,750 people, i.e., 3.38%);
- in the analyzed period, compared to 2011, the smallest increase occurred in 2021 (by 722,356 people, i.e., 3.36%), and the largest increase was in 2019 (by 14,191,475 people, i.e., 66.08%).

The proposed linear trend model (Table 2) explains the changes in the examined number of tourists by 98.11%, and the degree of fit of the model to the empirical data can be considered sufficiently high. The actual values of the analyzed number of tourists and the corresponding theoretical values obtained on the basis of the built model differ by an average of 764,059 people.

The coefficient of variation is at a satisfactorily low level, and the structural parameters are statistically significant. None of the methods used to assess the quality of the model are questionable, which suggests that the estimated trend model has been successfully verified.

Assuming that the linear development trend observed in 2011-2019 (excluding the pandemic period) remains consistent, it would be predicted, based on the built model, that the number of overnight visitors in Poland in 2023 will amount to 42,944,308 people. Given this forecast, there can be an average error of about 1,127,544 people, or 2.63%.

6. Conclusions

Using quantitative methods, the description, assessment, and prediction of various variables in time or space become more precise. The article draws attention to the fact that after the COVID-19 pandemic, which negatively affected the proper functioning and development of many sectors of the economy, the number of overnight visitors in Poland quickly reached the pre-pandemic level, and further increases are expected.

The views presented in this study and the results of our own research made it possible to achieve the goal set at the beginning of the work.

Using index methods, which are known and commonly used in research on the dynamics of phenomena, a description of changes in the number of tourists (both domestic and foreign) staying overnight in tourist accommodation facilities over the past dozen or so years in Poland was made.

Additionally, the observed developmental trend allowed the construction of a trend model, which was used for short-term prediction along with ex-ante assessment.

References:

Adamopoulos, A., Thalassinos, E.I. 2020. Tourism Development and Economic Growth: A Comparative Study for the G-6 Leaders. European Research Studies Journal, 23(1), 368-380.

Cieślak, M. (red.) 2001. Prognozowanie gospodarcze. Metody i zastosowanie. PWN, Warszawa.

Czerwiński, Z., Guzik, B. 1980. Prognozowanie ekonometryczne. Podstawy teoretyczne i metody. PWE, Warszawa.

Czyżycki, R., Klóska, R. 2011. Ekonometria i prognozowanie zjawisk ekonomicznych w przykładach i zadaniach. Economicus, Szczecin.

Czyżycki, R., Klóska, R. 2019. Wybrane zagadnienia z prognozowania. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Dittmann, P. 2000. Metody prognozowania sprzedaży w przedsiębiorstwie. Akademia Ekonomiczna, Wrocław.

Gajda, J.B. 2001. Prognozowanie i symulacja a decyzje gospodarcze. Wydawnictwo C.H. Beck, Warszawa.

Hanias, M.P., Curtis, P.G., Thalassinos, J.E. 2007a. Non-Linear Dynamics and Chaos: The Case of the Price Indicator at the Athens Stock Exchange. International Research Journal of Finance and Economics, (11).

Hanias, M.P., Curtis, P.G., Thalassinos, J.E. 2007b. Prediction with neural networks: the Athens stock exchange price indicator. European Journal of Economics, Finance and Administrative Sciences, 9, 21-27.

Hyndman, R., Athanasopoulos, G. 2018. Forecasting: Principles and Practice. Otexts.

Kaczmarek, J., Stasiak, A., Włodarczyk, B. 2010. Produkt turystyczny. Pomysł. Organizacja. Zarządzanie, Polskie Wydawnictwo Ekonomiczne, Warszawa.

Klóska, R., Czyżycki, R. 2019. Wybrane zagadnienia ze statystyki. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Kurek, W. (red.). 2007. Turystyka. Wydawnictwo Naukowe PWN, Warszawa.

Markidakis, S., Wheelwright, S., Hyndman, R. 1998. Forecasting: methods and applications. John Wiley & Sons.

Meyer, B. (red.). 2010. Potencjał turystyczny, zagadnienia przestrzenne, Zeszyty Naukowe Nr 590, Ekonomiczne Problemy Usług Nr 52. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Nowak, E. (red.). 1998. Prognozowanie gospodarcze. A.W. Placet, Warszawa.

Panasiuk, A. 2019. Rynek turystyczny. Struktura – procesy – tendencje. Difin, Warszawa.

Pawłowski, Z. 1973. Prognozy ekonometryczne. PWN, Warszawa.

Szmytke, R. 2022. Turystyka jako element brandingu narodowego w wybranych krajowych dokumentach strategicznych. Folia Turistica, 59, p. 107-126.

Uysal, M., Sirgy, M.J., Kruger, S. (red.). 2018. Managing Quality of Life in Tourism and Hospitality: Best Practice. MA: CABI, Boston.

Zeliaś, A. 1997. Teoria prognozy. PWE, Warszawa.

Panasiuk, A., Sidorkiewicz, M. 2017. Analiza SWOT wykorzystania funduszy europejskich w regionalnej gospodarce turystycznej w latach 2007-2013. Europa Regionum, Nr. 4, p. 81-92.