



THE FORECASTING POTENTIAL OF ADAPTIVE MODELS IN TOURISM

Anna Aleksandrova^a , Ekaterina Aigina^b , Veronika Dombrovskaya^c ^a Lomonosov Moscow State University, Faculty of Geography; <http://orcid.org/0000-0002-1772-8431>; e-mail: analexan@mail.ru^b Lomonosov Moscow State University, Faculty of Geography; <http://orcid.org/0000-0002-3832-169X>; e-mail: eaigina@yandex.ru^c Tver State University, Department of Tourism and Nature Management; <https://orcid.org/0000-0002-7138-1774>;
e-mail: Dombrovskaya.VE@tversu.ru**How to cite (APA style):** Aleksandrova, A., Aigina, E., Dombrovskaya, V. (2021). The forecasting potential of adaptive models in tourism. *Turyzm/Tourism*, 31 (2), 181–196. <https://doi.org/10.18778/0867-5856.31.2.10>**ABSTRACT**

The article discusses forecasting as one of the special scientific research areas which contribute to the assessment of tourist activity development prospects, the identification of key tourism development factors and effective management decision criteria. The study provides an overview of modern research methods used in Russia and other countries for making forecasts in the field of tourism. It aims at assessing the predictive capabilities of adaptive modeling, not frequently used currently in tourism research, for the quantitative analysis of tourist flows using the example of Barcelona, a major urban tourist destination in the pre-pandemic period. An example of a forecast for tourist numbers based on adaptive models is proposed, one of the key indicators showing tourist region success which have proven successful in the study of processes with a dynamic but unstable character.

KEYWORDS

forecasting, adaptive modeling, tourist flows, Holt-Winters method

ARTICLE INFORMATION DETAILSReceived:
17 September 2021
Accepted:
8 November 2021
Published:
31 December 2021**1. INTRODUCTION**

There are hardly any other social phenomena that can be compared with tourism in regards to its complex multifaceted nature, impacting almost all aspects of human life and, at the same time, its strong sensitivity to various challenges from outside. This is one of the reasons, perhaps the most important, for its lack of investigation despite its relatively long history, scale and significance. Tourism research methodology is still in the formation stage, however government authorities, businesses and consumers interested in tourism development need timely, complete and reliable information to assess the effectiveness of decisions taken, as well as to subjectively evaluate the prospects of tourism industry growth today. Progressive changes in tourism and the strengthening of tourist destination competitiveness are not viable

without a long-term strategy containing a system of prospective priorities, goals and subjectives for public administration (Aleksandrova, Aigina, 2020).

High efficiency in management is inconceivable without forecasting, the special scientific research on particular development prospects of a phenomenon. The use of forecasting methods in tourism is one of the most relevant and important areas since strategic and tactical plans for tourism industry development need to be based on trend analyses and objectively demonstrated prospects. Ongoing research shows that the 'mathematical apparatus' used in tourism studies is quite diverse and is actively used to solve multifaceted problems. However, the accuracy of the forecasts obtained by using different methods varies significantly depending on time series length, the research subject (tourism experience, study area, etc.) and requires further clarification.

The study aims at assessing the predictive capabilities of adaptive modeling for the quantitative analysis of tourist numbers using the example of Barcelona in the pre-pandemic period. The choice of Barcelona as the research subject is due to its leading position in 'World's Best Cities' rankings and, as a result, gives a dramatic picture of the manifestations of overtourism (congestion, increasing prices and social tension), confirmed by the authors' personal experience during field studies there. The COVID-19 pandemic and related restrictions have led to a deep recession in its tourism market as tourism has become one of the worst-hit areas of the economy. However, the problem of overtourism has not lost its relevance. Moreover, the problem of tourist management has become vigorously evident at all levels, so new methods for tourist activity forecasting will possibly contribute to management decision optimization.

The research methodology gives substantiated arguments for adaptive modeling selection and, in particular, the Holt-Winters method for predicting quantitative tourist flow indicators based on Barcelona-El Prat Airport data. The research results give a step-by-step solution to the problem with a detailed explanation of the methods used.

2. LITERATURE REVIEW

In the greater part of Russian and foreign research, tourist demand acts as the research subject both for domestic and inbound tourism markets. Several review articles note that it is the variable of tourist arrivals and/or overnight stays in accommodation facilities that is still the most popular indicator (Khaidi, Noratikah, Noryanti, 2019; Song, Li, 2008). In this case, forecasting schemes either follow the path of identifying patterns for time series development, with seasonality taken into account, or build multidimensional models that make the tourist demand factor dependent on a set of predictors forming cause-and-effect relationships. Multiple regression analysis which makes a specific indicator dependent on many influencing factors has become a fairly common method for solving this problem.

The selection of predictors, which can vary significantly, is of particular interest (Gladilin, Gladilin, 2016; Nikolaeva, Oreshkina, 2016). As the number of influencing factors included in a study is not limited, it automatically leads to the emergence of problems of selection. Its solution is possible through the use of the Least Angle Regression (LARS) algorithm, which sequentially changes predictors' values in terms of significance within the model (Lourenço, Gouveia, Rua, 2021). In addition to regression models, a forecast can

be constructed with the help of logistic equations of tourist flow dynamics (Verhulst logistic equation) based on an assumption regarding the presence of a hidden mechanism of influence in the internal properties of the flow and the role of the external environment. According to some researchers, a growth rate correlates with first impressions of a tourist area, and a slowdown is associated with competition between visitors under overtourism conditions (Chu, 2014; Zhagina, Nizovtsev, Svetlosanov, Pakhomova 2019).

The role of the host geographic location in global Spatiotemporal Autoregressive model construction prevails. This uses Spatial Weights matrices to assess the territorial proximity/remoteness of countries supplying and receiving tourists (Jiao, Li, Chen, 2020). Dynamic Spatial Panel models also assess the impact of region (province) location, including distance from capital cities (Yang, Zhang, 2019).

When making forecasts in tourism, special attention is paid to seasonality. The seasonal component is included in the dynamic structure along with the trend (general development trend) and the random component (unpredictable value, often called a random error). In some cases, cyclical factors that, in comparison with seasonality, have a greater effect and no strictly defined duration period are also highlighted and these components (time series decomposition) become one of the first stages of the analysis process. Further research usually includes measuring seasonal fluctuations (for example, using the absolute differences method, taking into account deviation of actual levels from those found when detecting a trend), mathematical modeling of seasonality and forecast construction. If the forecasting methodology is based on the assumption that seasonal fluctuation parameters remain relatively unchanged, they can be described using seasonal indices (Barchukov, 2008).

The application of Hodrick-Prescott (HP) and Hamilton statistical filters to highlight trends and seasonal components in time series is also interesting. Both methods use mathematical time series smoothing but are different in their basic approaches. The HP filter is symmetric (i.e. two-sided), as a result there is a shift in the initial and final smoothed values. The Hamilton filter eliminates some of the HP method disadvantages as it is asymmetric (one-sided) helping to avoid the problem of shifting final values. The application of these methods using specific data on tourist numbers in Australia (1991–2019) shows that it makes sense to combine these approaches to solve various tasks (Bosupeng, 2019).

Smoothing methods help to 'suppress' short-period components to identify the most characteristic features of the time series and to make them 'free' from a random component. For this purpose, the ETS (Error, Trend, Seasonal) model, as well as ARIMA (Autoregressive Integrated Moving Average) model which, unlike its

predecessor ARMA (Autoregressive Moving Average) works both with stationary (unchanging characteristics over time) and with non-stationary time series, are actively used in forecasting.

In academic research, a combined approach to making a forecast for a given subject using several statistical methods with comparison of subsequent results is very popular. For example 1) ARIMA; 2) UCM (Unobserved Components Model); 3) building a dynamic spatial panel model; 4) STARMA (Space-Time AutoRegressive Moving Average) are used by researchers from China to predict provincial tourist numbers. In addition to traditional statistical significance criteria, AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are used for choosing models (Yang, Zhang, 2019).

It should be noted that modeling based on Artificial Neural Networks (ANN) has become one of the most popular structural methods. A neural network is a mathematical tool with an important advantage – the ability to identify and generalize complex dependencies between the input and output parameters on which the possibility of making forecasts is based. Recently, these models have found an application in tourism studies. At the same time, many specialists working with neural networks argue that ANNs' potential in forecasting is huge, but careful selection of input parameters and large arrays of training samples are required to obtain high-quality results. The disadvantages of ANN are also often referred to as the 'black box effect' – neural network opacity expressing a lack of understanding of the processes taking place 'inside'. Hybrid models appear to eliminate these problems in forecasting. The writers of one of these works have suggested a synthesis of ANN and fuzzy logic methods creating an intelligent ANFIS system (Adaptive Neuro-Fuzzy Inferences Systems). This combination allows increasing ANN self-learning capabilities with the predictive capabilities of system improvement due to fuzzy logic methodological tools, well-proven for multiple but not always obvious relationships between dissimilar components (Atsalakis, Zopounidis, 2018).

The review given above shows that forecasting methods used in tourism research are fairly diverse and include different techniques and procedures to build forecasts based on data analysis (including retrospective) as well as consideration and measurement of external and internal relations of the forecast subject. However, the accuracy of the forecasts varies significantly depending on such aspects as time series length, the subject of study (tourism phenomena and study area). As many scholars argue, there is no single forecasting method that would give the best forecast for any problem (Ghalekhondabi, Ardjmand, Young, Weckman, 2019). The choice of method for each case should be individual.

3. METHODOLOGY

This study focuses on adaptive forecasting which is not frequently used in modern tourism research. The concept of adaptation is interpreted as the ability of a system to demonstrate purposeful adaptive behavior in conditions of external environmental fluctuation. The use of the term 'adaptive control' concerning a system's ability to change and achieve a certain optimal performance has become quite common in this context. Various individual enterprises, industries and economic processes can be regarded as adaptive even though many economic phenomena in the modern world are unique by nature, and this makes it difficult to apply the experience gained in practice. This aspect shows the particular importance of adaptive management which is aimed at organization and management. In fact, adaptive management is a set of actions and methods that facilitate change in a controlled subsystem depending on external environmental disturbances (Tikhomirova, 2021). Responsiveness to changing conditions acts as the main characteristic in this case.

Thus, the adaptation idea becomes multifaceted: it is both a property and a process taking place within a system that can and should be managed; it is also a method of working with the structure and parameters of models describing this system.

In most cases, official information showing the level of tourism development and the current state of the industry is presented in the form of time and/or spatial series that are neither long nor stationary. Their characteristics can vary significantly over time responding to fluctuations in factors that have a direct or indirect impact on the research subject. In addition to the power of their influence, their quantitative and qualitative composition can also change. However, many writers note that the causal economic processes of such series cannot be considered completely random, since they bear traces of the influence of the previous state (Andrianova, Golovin, Zykov, Lesko, Chukalina, 2020).

Attempts to use classical forecasting models for such complex dynamic processes have many disadvantages. For example, when building forecasts using extrapolation methods based on the principle of transferring trends from the past and present to the future, accuracy is directly proportional to the length of the observation series and also requires its stationarization. It is difficult to comply with these requirements when obtaining data in tourism.

Adaptive modeling principles based on exponential smoothing and adaptive models of seasonal phenomena are not so strictly dependent on the original series under study. However, unlike extrapolation methods, the adaptive model provides a calculation of predicted values following the patterns of mainly recent or the latest observations using special parameters and not all

the previous ones (Chekmarev, 2020). This feature can be attributed as an evident advantage of the method when in the course of time series analysis the trend that exists now is of key importance rather than the trend that developed over the entire observation period. Thus, adaptive forecasting is based on self-adjustable recurrent models. Thanks to their usage subsequent values can be calculated through previous ones with their informative value taken into account. This allows consideration of a high degree of variability in the context of the rapidly changing characteristics of business processes since the idea of a constant revision of predicted values, as actual data becomes available, forms the basis of adaptive models.

Based on the conclusions presented above, this study discusses the following research questions:

- Is the adaptation mechanism (the model coefficient adjustment mechanism) capable of adapting for tourism the specific conditions manifested in an absence of rigid development patterns when constructing a forecast for tourist flow quantitative indicators?
- What accuracy assessment does a forecast based on adaptive modeling have?

In the present research the quantitative indicators of tourist numbers at Barcelona-El Prat Airport taken from Aena, the Spanish airports and air navigation official operator, are used for description (Aena, 2019). One-dimensional monthly data time series from 2010 to 2019, converted into seasonal data time series by summation, were produced to get basic important information. A graphical representation of the resulting series is shown in Figure 1.

As seen on the graph, arrivals are growing and have a pronounced seasonal component. Brown's double exponential smoothing method (working with a linear trend in time series levels and the absence of seasonality) and Holt's method, which also does not focus on seasonal fluctuations, as proposed by some authors for solving similar forecasting problems, are not optimal (Demin, Semenova, 2001). In this case, it is suggested that a more advanced Holt-Winters method (triple exponential smoothing) is applied which uses two adaptation parameters (Holt's method) and takes into account the seasonality indicator (Winter's method) for adjusting the coefficients (Atchade, 2018). Depending on the way the influence of this method is considered, two types of models are distinguished:

- models with a multiplicative coefficient of seasonality (where the deviation linearly increases or decreases with an increase in the value measured);
- models with an additive coefficient of seasonality (where the error is constant at each point) (Davnis, Tinjakova, 2006).

The approach typical to the second type of model is applied to analyze the predictive abilities of adaptive models using the example of Barcelona.

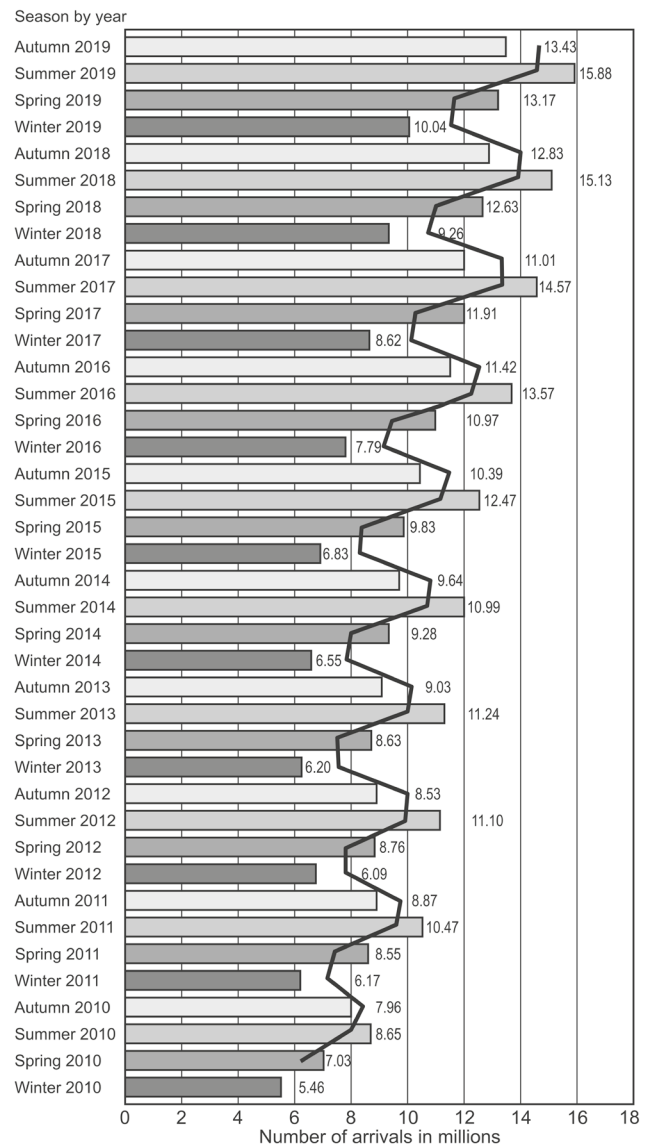


Figure 1. Changes to tourist arrivals at Barcelona-El Prat Airport 2010–2019 (million arrivals)
Source: based on Aena data

- The stages of the problem-solving process include:
- presentation and criteria-based assessment of the initial data;
 - calculation of the initial parameter values;
 - gradual adjustment of parameters using recurrent equations;
 - forecasting;
 - forecast quality assessment.

4. RESULTS

The modeling and forecasting process is preceded by testing the initial time series for stationarity, randomness, presence/absence of autocorrelation and other research that helps to identify the time series structure.

As is known, time series stationarity is manifested through the absence of significant changes in levels over time. This means that times series values show insignificant random fluctuations around the mean. If the process has a well-defined structure, then the series describing it is called non-stationary, and its observation reveals seasonal effects, trends and other components that depend on time. The nonstationarity of the data series under consideration in this research is obvious and is confirmed by classical statistical tests (Magnus, Katyshev, Peresetsky, 2004). In the case of forecasting using classical tools, non-stationary series require stationarization, otherwise the quality of the time series model is questioned. Adaptive modeling application gives the opportunity to work out the forecasts for almost any economic process where stationarity is extremely rare (Davnis, Tinjakova, 2006).

To test the hypothesis about the presence of a non-random component (definite development trend) in time series, the run criterion is used based on the median, and the 'run-up' and 'run-down' criterion. The point of the first method is to compare the series initial values with the calculated median and to specify the series where the sequence of levels is greater or less than the median (equalities are excluded from the analysis). Therefore, two parameters are determined: $\nu(n)$ is the number of series (in this case equal to 16) and $\tau(n)$ is the length of the largest series (for the factual series of observations it is equal to 7). The established parameters are compared with the calculated critical values:

$$\left\{ \begin{array}{l} \nu_1(n) > \left\lceil \frac{1}{2}(n+1 - 1.96\sqrt{n-1}) \right\rceil \\ \tau_1(n) < \lceil 3.3lg(n+1) \rceil \end{array} \right\}$$

Since one of these inequalities does not hold, the hypothesis of series randomness is rejected.

The second method consists of constructing a sequence $\delta_1, \delta_2, \dots, \delta_{n-1}$

$$\delta_t = \begin{cases} +, & \text{if } y_{t+1} - y_t > 0 \\ 0, & \text{if } y_{t+1} - y_t = 0 \\ -, & \text{if } y_{t+1} - y_t < 0, t = \overline{1, n} \end{cases}$$

Similar to the criterion considered above, the number of series $\nu(n)$ (16) and $\tau(n)$ (2) is determined, and then these parameters are compared with their critical values.

$$\left\{ \begin{array}{l} \nu_2(n) > \left\lceil \frac{1}{3}(2n-1) - 1.96\sqrt{\frac{16n-29}{90}} \right\rceil \\ \tau_2(n) < \tau_0 \end{array} \right\}$$

where n is the number of elements in a series,

$$\tau_0 = \begin{cases} 5, & \text{if } n \geq 26 \\ 6, & \text{if } 26 < n \leq 153 \\ 7, & \text{if } 153 < n \ll 1170 \end{cases}$$

For $n = 40$ this system of inequalities takes the following form:

$$\begin{cases} \nu_2(n) > 21 \\ \tau_2(n) < 6 \end{cases}$$

Since the first inequality obviously does not hold, the hypothesis about series randomness is rejected.

To check the series for first-order autocorrelation, which can significantly reduce the predicted values of the future model, the most well-known test Durbin-Watson (DW) is used. The general criterion scheme is as follows: according to the constructed empirical regression equation the deviations of the actual and theoretical values (ε_t) are determined and then using the formula

$$d = \frac{\sum_1^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_1^n \varepsilon_t^2}$$

the DW value is calculated. The obtained value of the DW criterion for the series analyzed turned out to be 1.906. When compared with the tabular values at a 5% level of materiality this value turns out to be greater than the critical value ($1.906 > 1.49$) proving the absence of autocorrelation.

The calculation of first and higher order autocorrelation coefficients is applied for mathematical confirmation of seasonality. The results are shown in Table 1.

Table 1. Calculation of 1–4 orders of autocorrelation coefficients for Barcelona tourist number time series (2010–2019)

Lag	Autocorrelation coefficient
1	0.526
2	-0.005
3	0.479
4	0.986

Source: authors.

This allows the conclusion that there are seasonal fluctuations over four seasons confirmed by the graphical method (Figure 1).

As later, when constructing an adaptive model, a linear trend is assumed to be obtained, followed by a seasonal component selection, and is necessary to be sure that the Gauss-Markov criterion is met. This checks the assumptions for applying the least squares method (LSM) when constructing linear regression. The R 'gvlma' package was used to determine the feasibility of Gauss-Markov conditions (Pena, Slate, 2019). The results of the study show that with the assumption of a normal distribution for the residual component, homogeneity of residue dispersion and the absence of correlation between them (heteroscedasticity test), the importance of the proposed model with the linear trend is confirmed.

gvlma (x = model.1)			
	Value	p-value	Decision
Global stat	2.7497	0.6005	assumptions acceptable
Skewness	0.3376	0.5612	assumptions acceptable
Kurtosis	0.9246	0.3363	assumptions acceptable
Link function	0.6771	0.4106	assumptions acceptable
Heteroscedasticity	0.8104	0.380	assumptions acceptable

s1 -2689727.5
 s2 -261705.5
 s3 1522453.4
 s4 -250691.8

The result also has a graphical interpretation (Figure 2).

As a result, the initial series test proved that this data can be used for making the forecast. Using the R software environment for adaptive method realisation is also suggested. In particular, the Holt-Winters function is used to build a trend model in its framework with the parameters taken as initial *a* and *b* values.

The main idea can be described as follows: the initial state of the model is determined at the initial point of the series using the current values of the parameters; based on this, forecasting is carried out one step ahead. After that, the discrepancy between the predicted and actual values – the deviation (the sum of squared errors (SSE) as one possible option) – is measured, which forms the basis for full agreement adjusting the parameters to the series dynamics. Then the process called ‘model training’ is repeated for the next time point, and so on. The numerical parameters alpha and beta for a series with no seasonality, and alpha, beta, gamma for a series with a pronounced seasonal component act as tools for adjusting and ‘training’ the model.

Seasonal component estimates are calculated as averaged values for trend deviation from the actual time series values. Alpha (coefficient for level smoothing), beta (coefficient for smoothing trending), and gamma (coefficient for seasonal smoothing) are calculated as parameters. All the specified parameters lie in the range from 0 to 1; their selection can be carried out at the start and specified in the form of constants or these values are determined automatically during the program operation. The selection scheme in this case is based on minimizing the sum of squared errors (SSE) (Atchade, 2018). The result of the initial data processing with alpha = beta = gamma = 0.1 is presented as:

Holt-Winters (x = data, alpha = 0.1, beta = 0.1, gamma = 0.1)

Smoothing parameters:

- alpha: 0.1
- beta: 0.1
- gamma: 0.1

Coefficients:

- [.1]
- a 13170119.6
- b 166126.0

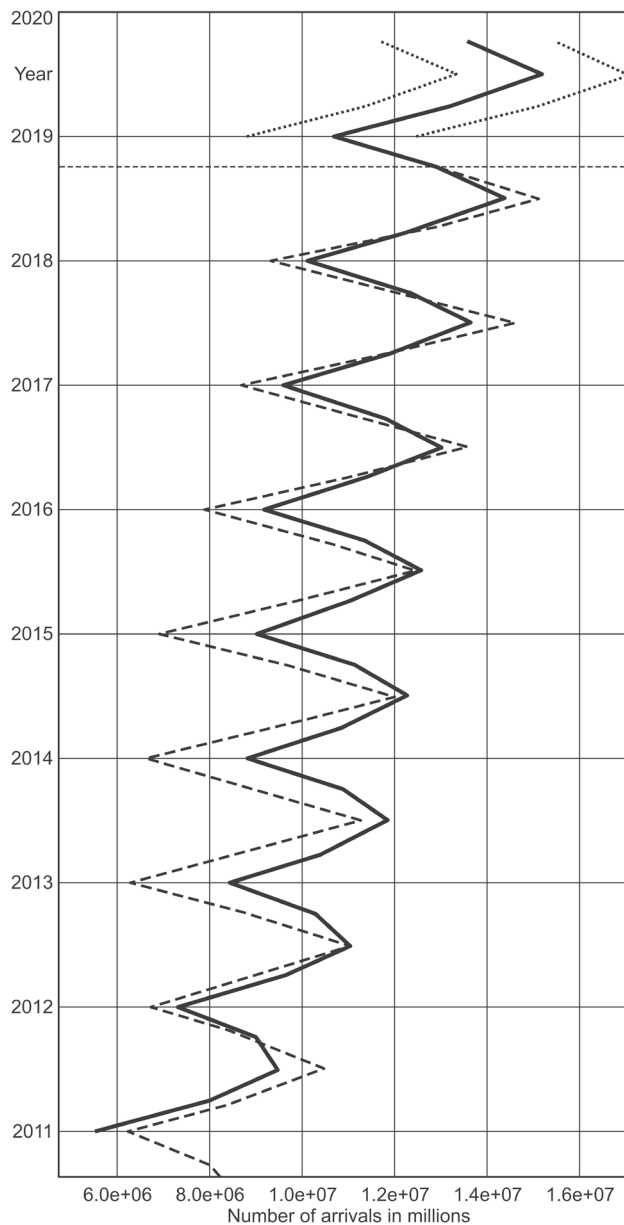


Figure 2. Initial, smoothed and predicted time series levels of tourist arrivals at Barcelona-El Prat Airport with smoothing parameters equal to 0.1
 Source: authors

As seen in Figure 2, the discrepancies between the initial and smoothed levels are quite noticeable which is confirmed by the SSE value = $3.8 \cdot 10^{13}$.

Automatic selection of adaptation parameters values gives considerably better results (Figure 3).

Holt-Winters (x = data)

- Smoothing parameters:
- alpha: 0.4174945

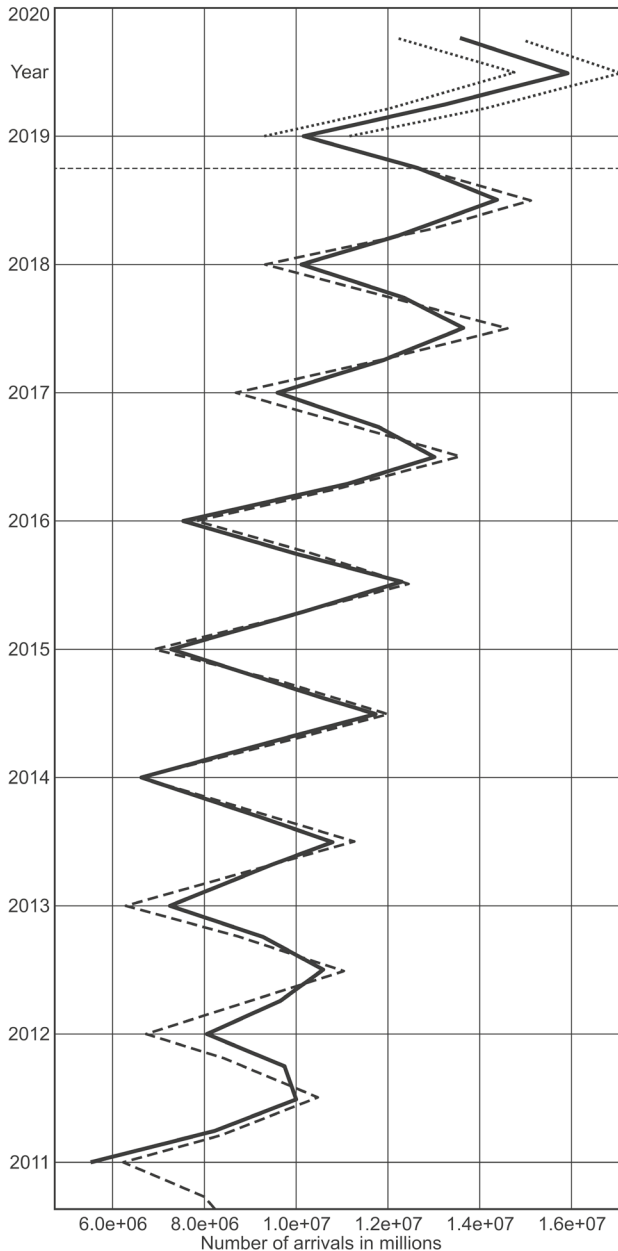


Figure 3. Initial, smoothed and predicted time series levels of tourist arrivals at Barcelona-El Prat Airport with automatically chosen smoothing parameters
Source: authors

beta: 0.2896451
 gamma: 1
 Coefficients:
 [.1]
 a 12967359.68
 b 195701.01
 s1 -3085720.38
 s2 64649.76
 s3 2401235.35
 s4 -136710.68

Minimum deviations of actual and smoothed levels are visually noticeable in Figure 3. In this case, the SSE value is also much lower ($6.5 \cdot 10^{12}$).

5. DISCUSSION

The resulting optimal smoothing parameters make forecast development possible using the predict function. A prediction interval is an important issue when making forecasts. Currently, forecasts with short-term, medium-term and long-term forecasting horizons are differentiated by scientists based on the ‘period of inertia’ (the time the process changes through inertia) concept (Svetunkov, Svetunkov, 2019). It is generally accepted that adaptive forecasting works best over a short-term period. The model proposed was tested by adaptive regulation with predictive calculations for the last four levels (2019 seasons). The assessment of forecast accuracy and reliability is typical of results achieved in this kind of work. The discrepancy magnitude between actual and predicted values can serve as an empirical measure of accuracy; but it is possible only if the lead-time period has already ended and the researcher has the necessary information, as seen in this study.

One of the ways to assess predictive calculations is to find the Mean Absolute Percentage Error (MAPE) using the formula:

$$\delta = \left(\frac{1}{n} \sum_{t=1}^n \frac{|x_t - \hat{x}_t|}{x_t} \right) \cdot 100\%$$

Table 2. Average approximation error for the predicted values of tourist arrivals to Barcelona calculated in the R environment with different smoothing parameters

Time series level	Actual values	Predicted values with smoothing parameters = 0.1	Predicted values with smoothing parameters automatically selected in the R environment
Winter 2019	10,041,360	10,646,518	10,077,340
Spring 2019	13,171,925	13,240,666	13,424,111
Summer 2019	15,876,816	15,190,951	15,955,689
Fall 2019	13,426,888	13,583,932	13,613,453
		MAPE = 3.01%	MAPE = 1.04%
		correlation coefficient = 0.997	correlation coefficient = 0.999

Source: authors.

If the error value turns out to be less than 10%, then the interpretation of the accuracy estimate can be recorded as 'high'. The correlation coefficient between predicted and actual values can be used as an additional comparative assessment. It is well-known that the closer this indicator is to 1, the higher the relationship between the time series of actual values and the predicted values of the attribute. Table 2 (see p. 187) shows the results of estimating forecast accuracy for various smoothing parameters.

Despite minor MAPE differences in forecast values in both cases, which indicates the successful adaptation of the model, preference is given to the option with the lowest value. That means forecasting with automatically selected smoothing parameters is objectively more accurate. Everything discussed above suggests fairly high

predictive capabilities for the proposed adaptive 'self-adjusting' model which successfully adapts to changing external conditions. However, the final conclusion can be drawn only after comparing other forecasting methods, in particular extrapolation, which is currently used in tourism development strategies and target indicator calculations. But this is a profound task for the next stage of research. Adaptive multivariate modeling including identification and selection of factors that affect tourist flow volume and structure, assessment of their influence on the modeling indicator, identification of the best form for communication and calculation for the desired model parameters are of great interest as further development directions as well. They will contribute to deeper and more detailed research on forecasting of tourism development numerical characteristics.

POTENCJAŁ PROGNOSTYCZNY MODELI ADAPTACYJNYCH W TURYSTYCE

Anna Aleksandrova^a , Ekaterina Aigina^b , Veronika Dombrovskaya^c 

^a Moskiewski Państwowy Uniwersytet Łomonosowa, Katedra Geografii; <http://orcid.org/0000-0002-1772-8431>;
e-mail: analexan@mail.ru

^b Moskiewski Państwowy Uniwersytet Łomonosowa, Katedra Geografii; <http://orcid.org/0000-0002-3832-169X>;
e-mail: eaigina@yandex.ru

^c Państwowy Uniwersytet w Twerze, Wydział Zarządzania Turyzmem i Przyrodą; <https://orcid.org/0000-0002-7138-1774>;
e-mail: Dombrovskaya.VE@tversu.ru

Sposób cytowania (styl APA): Aleksandrova, A., Aigina, E., Dombrovskaya, V. (2021). The forecasting potential of adaptive models in tourism. *Turyzm/Tourism*, 31 (2), 181–196. <https://doi.org/10.18778/0867-5856.31.2.10>

ABSTRAKT

Autorki omawiają prognozowanie jako szczególny obszar badań naukowych, który pomaga ocenić perspektywy rozwoju aktywności turystycznej, zidentyfikować kluczowe czynniki procesu zmian zachodzących w turызmie oraz kryteria efektywnego zarządzania. Artykuł zawiera przegląd nowoczesnych metod badawczych stosowanych w Rosji i w innych krajach do sporządzania prognoz w dziedzinie turystyki. Celem opracowania jest ocena predykcyjnych możliwości modelowania adaptacyjnego, obecnie niezbyt często stosowanego w badaniach turystycznych, do przeprowadzania ilościowej analizy ruchu turystycznego. Na podstawie danych dotyczących Barcelony – dużego miejskiego centrum turystycznego w czasach przed pandemią – opierając się na modelach adaptacyjnych, przedstawiono przykład prognozy liczby podróżnych. Jest to jeden z głównych wskaźników sukcesu turystycznego ujawnionego w trakcie badania procesów o dynamicznym, lecz niestałym charakterze.

SŁOWA KLUCZOWE

prognozowanie, modele adaptacyjne, ruch turystyczny, metoda Holta-Wintersa

INFORMACJE O ARTYKULE

Przyjęto:

17 września 2021 r.

Zaakceptowano:

8 listopada 2021 r.

Opublikowano:

31 grudnia 2021 r.

1. WPROWADZENIE

Trudno znaleźć zjawisko socjologiczne porównywalne do turystyki pod względem złożonego, wielopłaszczyznowego charakteru, oddziałujące na niemal wszystkie

aspekty ludzkiego życia, a jednocześnie tak podatne na różne wpływy zewnętrzne. To jedna z przyczyn, być może najważniejsza, braku opracowań naukowych w tej dziedzinie pomimo stosunkowo długiej historii, dużej skali i znaczenia tego zjawiska. Metodologia

badań nad turystyką wciąż się kształtuje, jednakże władze rządowe, przedsiębiorstwa i konsumenci zainteresowani rozwojem turystyki potrzebują właściwej, pełnej i rzetelnej informacji, aby ocenić skuteczność podejmowanych decyzji oraz subiektywnie oszacować dzisiejsze możliwości rozwoju przemysłu turystycznego. Progresywne zmiany w turystyce oraz wzrost konkurencyjności destynacji podróżniczych nie będą możliwe bez długoterminowej strategii, określającej przyszłe priorytety, cele i stanowisko administracji państwowej (Aleksandrova, Aigina, 2020).

Zarządzanie nie może być efektywne bez predykcji – naukowego badania potencjału danego zjawiska. Zastosowanie metod prognostycznych w turystyce to jeden z najważniejszych sposobów postępowania naukowego, ponieważ strategiczne i taktyczne planowanie rozwoju sektora turystycznego musi być oparte na analizie trendów oraz na obiektywnie prezentowanych perspektywach. Prowadzone obecnie prace pokazują, że aparat matematyczny stosowany w studiach nad turystyką jest dość zróżnicowany i służy praktycznemu rozwiązywaniu wieloaspektowych problemów, ale precyzyjność prognoz uzyskanych różnymi metodami nie jest taka sama, zależy od długości uwzględnianego okresu i przedmiotu badań (doświadczenie turystyczne, obszar badań itd.), a także wymaga dalszych wyjaśnień.

Celem artykułu jest ocena możliwości przeprowadzenia ilościowej analizy liczby podróżnych z zastosowaniem modelowania adaptacyjnego na przykładzie Barcelony z okresu przed pandemią. Wybór stolicy Hiszpanii na przedmiot badań wynika z czołowej pozycji tego miasta w rankingach „World’s Best Cities”. Dzięki rezultatom tych studiów uwiadamiają się dramatyczne skutki nadmiernego rozwoju turystyki (*overtourism*) – zagęszczenie ruchu ulicznego oraz rosnące ceny i napięcia społeczne, potwierdzone osobistym doświadczeniem autorek podczas badań terenowych. Pandemia COVID-19 i związane z nią restrykcje doprowadziły do głębokiej recesji na rynku turystycznym – branża ta jest jednym z najbardziej dotkniętych obszarów gospodarki. Mimo to problem przerosu ruchu turystycznego nie stracił na znaczeniu. Co więcej, trudności w zarządzaniu turystyką uwiłocznily się na wszystkich poziomach, dlatego można przypuszczać, że nowe metody predykcji aktywności podróżniczych przyczynią się do optymalizacji decyzji zarządczych.

Metodologia badań dostarcza mocnych argumentów przemawiających za wyborem modelowania adaptacyjnego, zwłaszcza metodą Holta-Wintersa, do przewidywania ilościowych wskaźników przepływu turystów, w tym przypadku na podstawie danych z portu lotniczego Barcelona-El Prat. Z wyników analizy wyłania się szczegółowe rozwiązanie problemu, ponadto zawierają one dokładne objaśnienie zastosowanych sposobów postępowania.

2. PRZEGLĄD LITERATURY

W większości rosyjskich i zagranicznych opracowań naukowych popyt turystyczny jest przedmiotem badań w kontekście zarówno krajowych, jak i zagranicznych rynków turystycznych. W kilku artykułach przeglądowych zauważono, że najbardziej popularnym wskaźnikiem nadal pozostaje zmienna przyjazdów turystycznych oraz/lub noclegów (Khaidi, Noratikah, Noryanti, 2019; Song, Li, 2008). W tym przypadku techniki prognostyczne albo podążają drogą identyfikowania typowych modeli rozwoju szeregów czasowych z uwzględnieniem sezonowości, albo budują wielowymiarowe modele, które uzależniają czynnik popytu turystycznego od grupy predyktorów (prognostyków) tworzących związki przyczynowo-skutkowe. Analiza regresji wielokrotnej, która stanowi specyficzny wskaźnik zależny od wielu zmiennych, stała się powszechną metodą rozwiązywania tego zagadnienia.

Szczególnie ważny jest dobór prognostyków. Mogą się one znacznie różnić między sobą (Gladilin, Gladilin, 2016; Nikolaeva, Oreshkina, 2016). Liczba czynników wpływających, wziętych pod uwagę w badaniu, nie jest ograniczona, co prowadzi automatycznie do problemu z ich selekcją. Można go rozwiązać poprzez zastosowanie algorytmu LARS (*Least Angle Regression and Shrinkage*), który sekwencyjnie zmienia wartości predyktorów pod względem funkcji pełnionej w modelu (Lourenço, Gouveia, Rua, 2021). Poza modelami regresji prognozę można konstruować za pomocą logistycznych równań dynamiki ruchu turystycznego (równanie logistyczne Verhulsta), na podstawie założenia, że istnieje mechanizm wpływu ukryty w cechach wewnętrznych tego ruchu, a środowisko zewnętrzne odgrywa określoną rolę. Według niektórych badaczy tempo wzrostu skorelowane jest z pierwszymi wrażeniami dotyczącymi obszaru turystycznego, jego spowolnienie zaś – z rywalizacją między gośćmi w warunkach nadmiernego rozwoju turystyki (Chu, 2014; Zhagina, Nizovtsev, Svetlosanov, Pakhomova, 2019).

W konstruowaniu globalnego, autoregresyjnego modelu przestrzenno-czasowego najważniejsza jest lokalizacja geograficzna gospodarza. Aby ocenić rolę odległości terytorialnej między krajami wysyłającymi i przyjmującymi podróżujących, zastosowano w nim macierze wag przestrzennych (Jiao, Li, Chen, 2020). Modele dynamicznego panelu przestrzennego określają także znaczenie położenia regionu (prowincji), łącznie z odległością od miast stołecznych (Yang, Zhang, 2019).

Podczas prognozowania w dziedzinie turystyki szczególną uwagę zwraca się na sezonowość. Element ten zawarty jest w strukturze dynamicznej, podobnie jak trend (ogólny trend rozwoju) oraz komponent losowy (wartość nieprzewidywalna, często zwana błędem losowym). W niektórych przypadkach podkreśla się też czynniki cykliczne, w porównaniu z sezonowością

wywołujące silny efekt i niemające ściśle określonego okresu trwania. Te składowe (dekompozycja szeregu czasowego) są jednym z pierwszych etapów procesu analizy. Dalsze badania polegają zazwyczaj na pomiarach fluktuacji sezonowych (np. za pomocą metody różnic bezwzględnych, uwzględniającej odchylenia poziomów aktualnych od tych, które stwierdzono przy odkryciu trendu), matematycznym modelowaniu sezonowości i konstruowaniu prognoz. Jeśli metodologia prognostyczna oparta jest na założeniu, że parametry wahań sezonowych pozostają stosunkowo niezmiennie, można je opisać za pomocą wskaźników sezonowych (Barchukov, 2008).

Równie interesujące jest zastosowanie statystycznych filtrów Hodricka-Prescotta (HP) i Hamiltona w celu uwypuklenia trendów i komponentów sezonowych w szeregach czasowych. W przypadku obu metod wykorzystywane jest matematyczne wygładzanie szeregu czasowego, ale różnią się one zasadniczym podejściem. Filtr HP jest symetryczny (dwustronny), co prowadzi do przesunięcia początkowych i końcowych wartości wygładzonych. Filtr Hamiltona eliminuje niektóre wady metody HP, ponieważ jest asymetryczny (jednostronny) i pozwala uniknąć problemu przesunięcia wartości końcowych. Takie sposoby postępowania z użyciem konkretnych danych dotyczących liczby turystów w Australii (1991–2019) pokazują, że połączenie tych podejść jest zasadne (Bosupeng, 2019).

Metody wygładzania pomagają „wytlumaczyć” komponenty krótkoterminowe, aby zidentyfikować najbardziej charakterystyczne cechy szeregu czasowego i uwolnić je od czynnika losowego. W tym celu w predykcji powszechnie stosowane są modele ETS (*Error, Trend, Seasonal*) i ARIMA (*Autoregressive Integrated Moving Average*), który w przeciwieństwie do swego poprzednika ARMA (*Autoregressive Moving Average*) działa zarówno w przypadku stacjonarnego (cechy niezmieniające się w czasie), jak i niestacjonarnego szeregu czasowego.

W badaniach naukowych bardzo popularne jest mieszanie różnych sposobów sporządzania prognozy na dany temat – wykorzystywanie kilku metod statystycznych i porównywanie kolejnych wyników. Na przykład ARIMA, UCM (*Unobserved Components Model*), dynamiczny model panelu przestrzennego czy STARMA (*Space-Time AutoRegressive Moving Average*) używane są przez chińskich naukowców do przewidywania liczby turystów w poszczególnych prowincjach. Oprócz tradycyjnych kryteriów istotności statystycznej również AIC (*Akaike Information Criterion*) oraz BIC (*Bayesian Information Criterion*) są stosowane przy wyborze modelu (Yang, Zhang, 2019).

Należy zauważyć, że modelowanie oparte na sztucznych sieciach neuronowych (*Artificial Neural Networks*, ANN) stało się jedną z najpopularniejszych metod strukturalnych. Sieć neuronowa jest narzędziem

matematycznym mającym istotną zaletę – pozwala identyfikować i uogólniać złożone zależności między parametrami wejściowymi a wyjściowymi, na których opiera się możliwość tworzenia prognoz. Ostatnio modele te znalazły zastosowanie w badaniach turystyki. Jednocześnie wielu specjalistów pracujących z sieciami neuronowymi twierdzi, że ANN w predykcji ma ogromny potencjał, ale potrzebne są przemyślane wybory parametrów początkowych oraz duża liczba próbek szkoleniowych, aby otrzymać wyniki wysokiej jakości. Wady ANN często nazywane są efektem czarnej skrzynki – nieprzezroczystość sieci neuronowych jest przyczyną braku zrozumienia procesów zachodzących „wewnątrz”. Pojawiają się więc modele hybrydowe, dzięki którym można wyeliminować te problemy. Autorzy jednej z prac zaproponowali połączenie ANN i metod logiki rozmytej, tworząc w ten sposób inteligentny system ANFIS (*Adaptive Neuro-Fuzzy Inferences Systems*). Taka kombinacja pozwala zwiększyć możliwości samodzielnego uczenia się ANN dzięki ulepszonemu systemowi predykcji oraz narzędziom metodologicznym logiki rozmytej, dobrze sprawdzonym w licznych, ale nie zawsze oczywistych związkach pomiędzy odmiennymi komponentami (Atsalakis, Atsalaki, Zopounidis, 2018).

Z zaprezentowanego przeglądu literatury wynika, że istnieje wiele metod prognostycznych stosowanych w badaniach turystycznych, oferujących rozmaite techniki i procedury tworzenia prognoz opartych na analizie danych (również retrospektywnej). W metodach tych zostaje uwzględniony i dokonany pomiar zewnętrznych i wewnętrznych związków przedmiotu predykcji. Trafność przewidywania jednakże jest bardzo zróżnicowana, zależna od takich aspektów, jak długość szeregu czasowego i przedmiot badania (zjawiska turystyczne i obszar badań). Według wielu naukowców nie istnieje jedna metoda, która gwarantowałaby najlepszą prognozę dla jakiegokolwiek zagadnienia (Ghalekhondabi, Ardjmand, Young, Weckman, 2019). Sposób postępowania powinien być dobierany osobno za każdym razem.

3. METODYKA

W niniejszym artykule skupiono się na prognozowaniu adaptacyjnym, które rzadko bywa wykorzystywane we współczesnych badaniach turystycznych. Pojęcie adaptacji interpretowane jest jako zdolność systemu do wykazywania celowego zachowania przystosowanego w warunkach zewnętrznych fluktuacji środowiskowych. Stosowanie terminu „kontrola adaptacyjna”, odnoszącego się do zdolności systemu do zmiany i osiągania pewnych optymalnych wyników, stało się w tym kontekście dość powszechne. Różne

pojedyncze przedsiębiorstwa, gałęzie przemysłu i procesy gospodarcze mogą być uważane za adaptacyjne, nawet jeśli wiele zjawisk ekonomicznych we współczesnym świecie jest z natury unikatowych, co sprawia, że trudno wykorzystać w praktyce zdobyte doświadczenia. Aspekt ten pokazuje szczególne znaczenie zarządzania adaptacyjnego, którego celem jest organizacja i zarządzanie. W istocie to zestaw działań i metod ułatwiających modyfikację w kontrolowanym podsystemie w zależności od zewnętrznych zakłóceń środowiskowych (Tikhomirova, 2021). W tym przypadku główną cechą jest reagowanie na zmieniające się warunki.

Idea adaptacji staje się zatem wieloaspektowa – to zarówno cecha, jak i proces zachodzący w systemie, który może i powinien być kontrolowany. Jest to również metoda pracowania ze strukturą i parametrami modeli opisującymi dany zespół elementów.

W większości przypadków oficjalne informacje na temat poziomu rozwoju turystyki i bieżącego stanu tego sektora zostają zaprezentowane w formie szeregów czasowych i przestrzennych, które nie są ani długie, ani stałe. Ich cechy mogą się znacznie różnicować w czasie, w odpowiedzi na wahania czynników, które mają bezpośredni lub pośredni wpływ na przedmiot badań. Oprócz siły oddziaływania zmienia się również ich ilościowa i jakościowa kompozycja. Wielu autorów zaznacza jednak, że przyczynowe procesy ekonomiczne takich szeregów nie mogą być uważane za całkowicie losowe, ponieważ mają one znamiona wpływu stanu poprzedniego (Andrianova, Golovin, Zykov, Lesko, Chukalina, 2020).

Próby zastosowania klasycznych modeli predykcji dla tak dynamicznych i złożonych procesów mają wiele wad. Na przykład kiedy tworzy się prognozy, stosując metody ekstrapolacyjne, oparte na zasadzie przenoszenia trendów z przeszłości i teraźniejszości do przyszłości, dokładność jest wprost proporcjonalna do długości obserwowanego szeregu i również wymaga standaryzacji. Trudno jest dostosować się do tych wymagań, gdy pozyskuje się dane turystyczne.

Zasady modelowania adaptacyjnego oparte na wygładzaniu wykładniczym i modelach adaptacyjnych zjawisk sezonowych nie są silnie zależne od oryginalnego, badanego szeregu czasowego. Jednakże – w przeciwieństwie do metod ekstrapolacyjnych – model adaptacyjny umożliwia wyliczenie przewidywanych wartości, głównie na podstawie najświeższych obserwacji, z użyciem specjalnych, a nie wszystkich stosowanych wcześniej parametrów (Chekmarev, 2020). Cecha ta może być określona jako oczywista zaleta metody, gdy w trakcie analizy szeregu czasowego kluczowe znaczenie ma trend obecny, a nie trend rozwijający się przez cały okres obserwacji. Prognozowanie adaptacyjne oparte jest zatem na samokorygujących się, modelach rekurencyjnych. Dzięki ich zastosowaniu kolejne wielkości

można obliczyć poprzez poprzednie, biorąc pod uwagę ich wartość informacyjną. Pozwala to na uwzględnienie wysokiego stopnia zmienności w kontekście gwałtownie przeobrażających się cech procesów gospodarczych, ponieważ idea ciągłej rewizji przewidywanych wartości tworzy podstawę modeli adaptacyjnych, w miarę jak aktualne dane stają się dostępne.

Odnosząc się do zaprezentowanych wniosków, autorki artykułu próbują odpowiedzieć na następujące pytania badawcze:

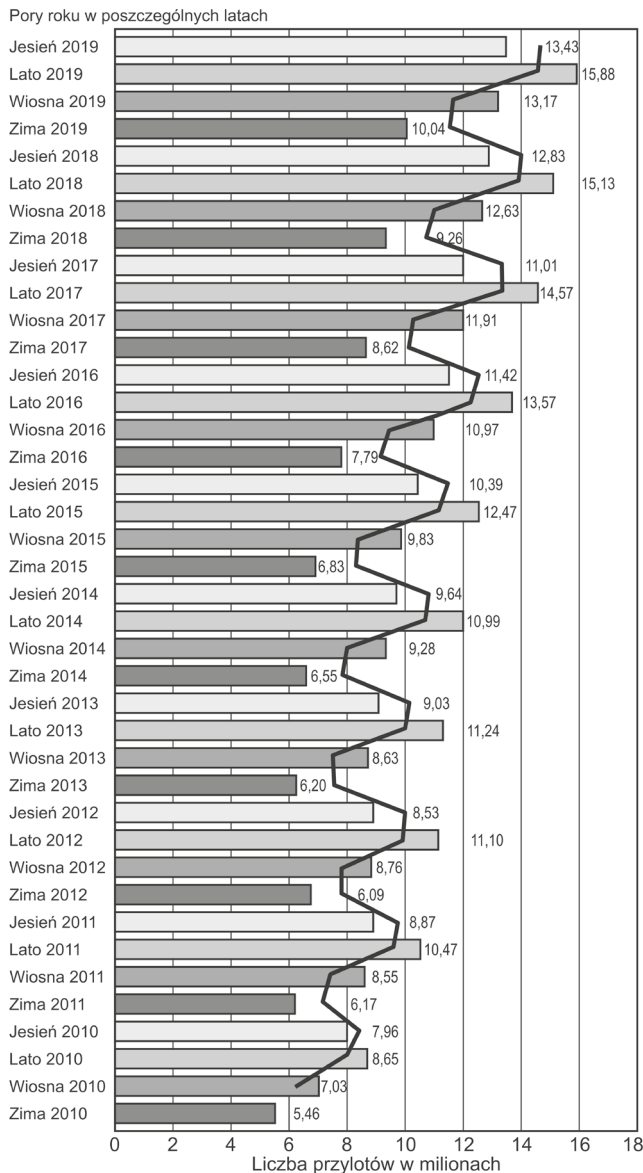
- Czy mechanizm adaptacyjny (mechanizm korekty współczynnika modelu) jest w stanie dopasować do turystyki specyficzne warunki przejawiające się w braku sztywnych wzorców rozwoju podczas prognozowania wskaźników ilościowych ruchu turystycznego?
- Jak dokładna jest prognoza oparta na modelowaniu adaptacyjnym?

W prezentowanym badaniu do opisu wykorzystano ilościowe wskaźniki liczby turystów na lotnisku El Prat w Barcelonie, dostarczone przez Aenę – oficjalnego operatora hiszpańskich portów lotniczych i ruchu powietrznego (Aena, 2019). Dane dotyczące jednowymiarowego, miesięcznego szeregu czasowego w latach 2010–2019 zostały przekonwertowane na dane sezonowego szeregu czasowego poprzez zsumowanie, co było źródłem podstawowych, ważnych informacji. Otrzymany szereg został przedstawiony w formie graficznej na rysunku 1 (zob. s. 192).

Jak widać na wykresie, liczba przylotów rośnie i zawiera wyraźny, sezonowy komponent. Metoda podwójnego wygładzania wykładniczego Browna, działająca w przypadku trendu linearnego na poziomach szeregów czasowych i przy braku sezonowości, oraz metoda Holta, która również nie skupia się na fluktuacjach sezonowych – proponowane przez niektórych autorów w celu rozwiązywania podobnych problemów diagnostycznych – nie są optymalne (Demin, Semenova, 2001). W tym przypadku sugeruje się zastosowanie bardziej zaawansowanej metody Holta-Wintersa (potrójne wygładzanie wykładnicze), w której wykorzystuje się dwa parametry adaptacyjne (metoda Holta) i uwzględnia się wskaźnik sezonowości (metoda Wintersa) w celu dostosowania współczynników (Atchade, 2018). W zależności od sposobu rozpatrywania działania tej metody, można wyróżnić dwa rodzaje analizy:

- modele z mnożnikowym współczynnikiem sezonowości, w których odchylenie linearnie zwiększa się lub zmniejsza wraz ze wzrostem mierzonej wartości;
- modele z addytywnym współczynnikiem sezonowości, w których błąd jest stały w każdym punkcie (Davnis, Tinjakowa, 2006).

Podejście charakterystyczne dla drugiego typu modelu stosowane jest do sprawdzenia zdolności progностycznych modeli adaptacyjnych, takich jak w analizowanym przypadku Barcelony.



Rysunek 1. Przyloty na lotnisko Barcelona-El Prat 2010–2019

Źródło: opracowanie na podstawie danych Aeny

Stadia procesu rozwiązywania problemów to:

- prezentacja i oparta na kryteriach ocena danych początkowych,
- obliczenie początkowych wartości parametrów,
- stopniowe dostosowywanie parametrów za pomocą równań rekurencyjnych,
- prognozowanie,
- ocena jakości prognozy.

4. WYNIKI

Proces modelowania i predykcji jest poprzedzony testowaniem początkowego szeregu czasowego pod względem stacjonarności, przypadkowości, obecności/braku autokorelacji. Przeprowadzane są też inne

badania, które pomagają zidentyfikować strukturę szeregu czasowego.

Stacjonarność (stałość) szeregu czasowego przejawia się w braku istotnych zmian w poziomach w czasie. To oznacza, że wartości szeregu czasowego wykazują nieznaczne, przypadkowe fluktuacje wokół średniej. Jeżeli proces ma ściśle określoną strukturę, wówczas szereg, który go opisuje, zwany jest szeregiem niestacjonarnym, a jego obserwacja ujawnia sezonowe skutki, trendy i inne składniki zależne od czasu. Niestacjonarność danych dotyczących rozważanego w tym artykule szeregu jest oczywista i potwierdzona przez klasyczne testy statystyczne (Magnus, Katyshev, Peresetsky, 2004). W przypadku prognozowania, w którym stosuje się tradycyjne narzędzia, szereg niestacjonarny wymaga standaryzacji. W przeciwnym razie jakość modelu szeregu czasowego jest wątpliwa. Zastosowanie modelowania adaptacyjnego daje możliwość wypracowania prognoz dla niemal każdego procesu gospodarczego, w którym stacjonarność występuje niezwykle rzadko (Davnis, Tinjakova, 2006).

Aby zweryfikować hipotezę dotyczącą obecności nieprzypadkowego komponentu (określonego trendu rozwoju) w szeregu czasowym, stosuje się kryterium *run* oparte na medianie oraz kryteria *run up* i *run down*. Celem pierwszej metody jest porównanie początkowych wartości szeregu z obliczoną medianą oraz określenie szeregu, w którym sekwencja poziomów jest większa lub mniejsza niż mediana (równości są wyłączone z analizy). Zatem określone są dwa parametry, tj. $v(n)$ – liczba szeregów (w tym przypadku równa 16) oraz $\tau(n)$ – długość najdłuższego szeregu (dla obserwacji badanego szeregu to 7). Ustalone parametry porównywane są do obliczanych wartości krytycznych:

$$\begin{cases} v_1(n) > \left[\frac{1}{2}(n+1 - 1,96\sqrt{n-1}) \right] \\ \tau_1(n) < [3,3]g(n+1) \end{cases}$$

Ponieważ jedna z tych nierówności się nie sprawdza, hipoteza o przypadkowości szeregu zostaje odrzucona.

Druga metoda polega na konstruowaniu sekwencji $\delta_1, \delta_2, \dots, \delta_{n-1}$

$$\delta_t = \begin{cases} +, & \text{if } y_{t+1} - y_t > 0 \\ 0, & \text{if } y_{t+1} - y_t = 0 \\ -, & \text{if } y_{t+1} - y_t < 0, t = \overline{1, n} \end{cases}$$

Podobnie jak w przypadku kryterium rozważanego wcześniej, określana jest liczba szeregów $v(n)$ (16) oraz $\tau(n)$ (2), a następnie parametry te są porównywane z ich wartościami krytycznymi.

$$\begin{cases} v_2(n) > \left[\frac{1}{3}(2n-1) - 1,96 \sqrt{\frac{16n-29}{90}} \right] \\ \tau_2(n) < \tau_0, \end{cases}$$

gdzie n to liczba elementów w szeregu,

$$\tau_0 = \begin{cases} 5, & \text{if } n \geq 26 \\ 6, & \text{if } 26 < n \leq 153 \\ 7, & \text{if } 153 < n < 1170 \end{cases}$$

Dla $n = 40$ ten układ nierówności przybiera następującą formę:

$$\begin{cases} v_2(n) > 21 \\ \tau_2(n) < 6 \end{cases}$$

Jako że pierwsza nierówność w sposób oczywisty się nie sprawdza, hipoteza o przypadkowości szeregu zostaje odrzucona.

Aby sprawdzić szereg pod względem autokorelacji pierwszego stopnia, która może znacząco zredukować przewidywane wartości przyszłego modelu, stosuje się najbardziej znany test Durбина-Watsona (DW). Ogólny schemat kryterium wygląda następująco: zgodnie ze skonstruowanym empirycznie równaniem regresji określone są odchylenia rzeczywistych i teoretycznych wartości (ε_t), a następnie obliczana jest wartość DW za pomocą wzoru:

$$d = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}$$

Otrzymana wartość kryterium DW dla analizowanego szeregu wyniosła 1,906. W porównaniu z wartościami tabelarycznymi na poziomie 5% materialności liczba ta okazuje się większa niż wartość krytyczna ($1,906 > 1,49$), co świadczy o braku autokorelacji.

Współczynniki autokorelacji pierwszego i wyższego stopnia obliczane są w celu matematycznego potwierdzenia sezonowości. Wyniki zaprezentowano w tabeli 1.

Tabela 1. Współczynniki autokorelacji 1–4 stopnia dla szeregu czasowego liczby turystów przybywających do Barcelony (2010–2019)

Lag	Współczynnik autokorelacji
1	0,526
2	-0,005
3	0,479
4	0,986

Źródło: opracowanie własne.

Pozwala to na konkluzję, że w trakcie czterech pór roku zachodzą fluktuacje sezonowe, potwierdzone metodą graficzną (rysunek 1).

Przy konstruowaniu modelu adaptacyjnego zakłada się otrzymanie trendu linearnego, a następnie wybór komponentu sezonowego. Należy się upewnić, że spełnione jest kryterium Gaussa-Markowa, które weryfikuje założenia metody najmniejszych kwadratów (*Least Squares Method*, LSM) podczas konstruowania regresji liniowej. Aby określić zasadność warunków Gaussa-Markowa, zastosowano pakiet R 'gvlma' (Pena, Slate, 2019). Wyniki badań pokazują, że przy założeniu normalnego rozkładu pozostałych składników

równomierność ich dyspersji i brak korelacji pomiędzy nimi (test heteroscedastyczności) oraz ważność zaproponowanego modelu z trendem linearnym zostają potwierdzone.

gvlma ($x = \text{model.1}$)

	Wartość	Wartość p	Decyzja
Stan ogólny	2,7497	0,6005	założenia akceptowalne
Niesymetryczność	0,3376	0,5612	założenia akceptowalne
Kurtoza	0,9246	0,3363	założenia akceptowalne
Funkcja łączenia	0,6771	0,4106	założenia akceptowalne
Heteroscedastyczność	0,8104	0,380	założenia akceptowalne

Ostatecznie test serii początkowej potwierdził, że te dane mogą być wykorzystane do tworzenia prognozy. Test sugeruje również zastosowanie środowiska oprogramowania R do zrealizowania metod adaptacyjnych. W szczególności używa się funkcji Holta-Wintersa do budowania modelu trendu, biorąc pod uwagę parametry stanowiące wyjściowe wartości a i b .

Główną ideę można przedstawić w następujący sposób: początkowy stan modelu jest określony w początkowym punkcie szeregu czasowego, z uwzględnieniem bieżących wartości parametrów – na tej podstawie stawiany jest pierwszy krok w prognozie. Potem mierzone są rozbieżności pomiędzy wartościami przewidywanymi a rzeczywistymi – odchylenie (suma kwadratów błędów SSE jako jedna z możliwych opcji). Tworzy to podstawę pełnego dostosowania parametrów do dynamiki szeregu. Następnie proces zwany treningiem modelowym powtarzany jest dla kolejnego punktu czasowego. Liczbowe parametry $alpha$ i $beta$ dla szeregu pozbawionego sezonowości oraz $alpha$, $beta$, $gamma$ dla szeregu z uwypuklonym komponentem sezonowym działają jako narzędzia do dostosowywania i „uczenia” modelu.

Element sezonowy jest szacunkowo obliczany jako uśrednione wartości dla odchylenia trendu od rzeczywistych wartości szeregów czasowych. $Alpha$ (współczynnik wygładzania poziomu), $beta$ (współczynnik wygładzania trendu) i $gamma$ (współczynnik wygładzania sezonowego) obliczane są jako parametry. Wszystkie wyszczególnione zmienne mają zakres 0–1. Ich wybór może być dokonany na początku i wyrażony w formie stałych wartości, lub też wartości te ustala się automatycznie podczas działania programu. Plan wyboru w tym przypadku oparty jest na minimalizowaniu sumy kwadratów błędów (SSE) (Atchade, 2018). Wynik przetworzenia danych wstępnych za pomocą $alpha = beta = gamma = 0,1$ prezentuje się następująco:

Holt-Winters ($x = \text{dane}$; $alpha = 0,1$; $beta = 0,1$; $gamma = 0,1$)

Parametry wygładzania:

α : 0,1

β : 0,1

γ : 0,1

Współczynniki:

[1]

a 13 170 119,6

b 166 126,0

s_1 -2 689 727,5

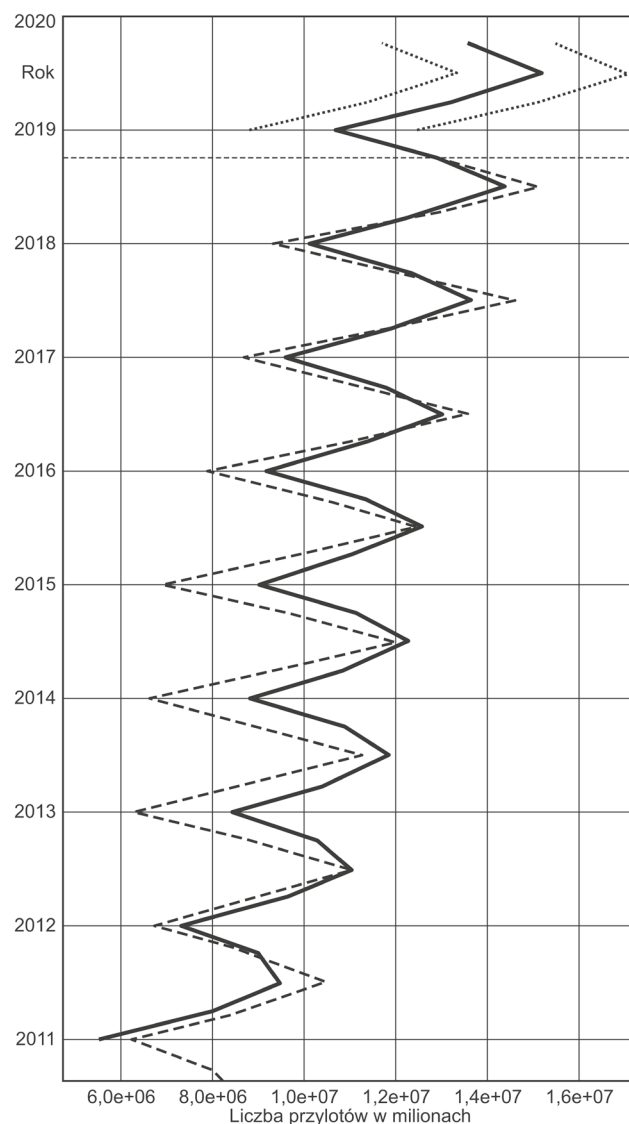
s_2 -261 705,5

s_3 1 522 453,4

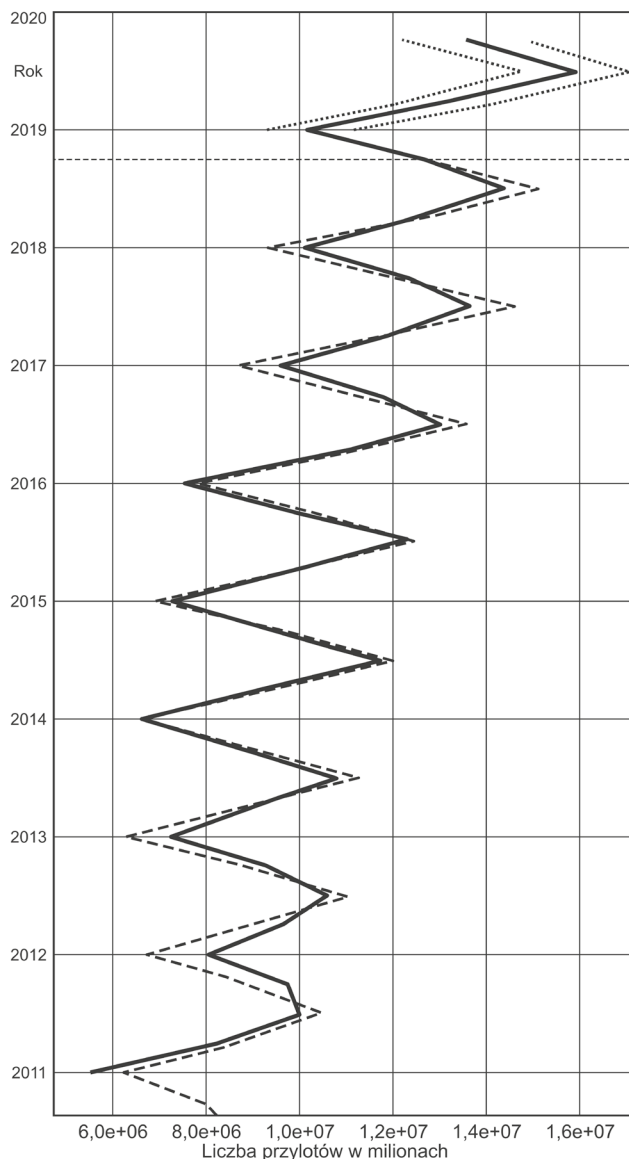
s_4 -250 691,8

Wynik ma również interpretację graficzną (rysunek 2).

Jak widać na rysunku 2, rozbieżności między poziomami początkowymi a wygładzonymi są dość wyraźne, co jest potwierdzone wartością $SSE = 3,8 \cdot 10^{13}$.



Rysunek 2. Początkowe, wygładzone i przewidywane poziomy szeregi czasowe przyjazdów turystycznych na lotnisko Barcelona-El Prat przy parametrach wygładzania równych 0,1
Źródło: opracowanie własne



Rysunek 3. Początkowe, wygładzone i przewidywane poziomy szeregi czasowe przyjazdów turystycznych na lotnisko Barcelona-El Prat przy automatycznie wybranych parametrach wygładzania
Źródło: opracowanie własne

Automatyczny wybór wartości parametrów adaptacyjnych daje znacznie lepsze rezultaty (rysunek 3).
Holt-Winters ($x = \text{dane}$)

Parametry wygładzania:

α : 0,4174945

β : 0,2896451

γ : 1

Współczynniki:

[1]

a 12 967 359,68

b 195 701,01

s_1 -3 085 720,38

s_2 64 649,76

s_3 2 401 235,35

s_4 -136 710,68

Minimalne odchylenia poziomów właściwych i wygładzonych są widoczne na rysunku 3. W tym wypadku wartość SSE jest również o wiele niższa ($6,5 \cdot 10^{12}$).

5. Dyskusja

Optymalne parametry wygładzania umożliwiają rozwój predykcji za pomocą funkcji przewidywania. Interwał predykcyjny to ważna sprawa przy tworzeniu prognozy. Obecnie prognozy o krótko-, średnio- i długoterminowym horyzoncie różnicowane są przez naukowców na podstawie koncepcji „okresu inercji”, czyli czasu, w którym proces zmienia się poprzez bezwładność (Svetunkov, Svetunkov, 2019). Jest ogólnie przyjęte, że prognozowanie adaptacyjne sprawdza się najlepiej w krótkim okresie. Zaproponowany model został przetestowany przez regulacje adaptacyjne za pomocą kalkulacji predykcyjnych w odniesieniu do ostatnich czterech poziomów (pory roku 2019). Ocena trafności i rzetelności prognozy jest typowa dla wyników uzyskanych w tego rodzaju analizie. Wielkość rozbieżności między wartościami rzeczywistymi a przewidywanymi może posłużyć jako empiryczny miernik dokładności, ale jest to możliwe tylko, jeżeli okres wstępny już się zakończył i badacz posiada niezbędne informacje, jak to widać na przykładzie niniejszej pracy.

Jednym ze sposobów oszacowania obliczeń predykcyjnych jest znalezienie *średniego bezwzględnego błędu procentowego* (Mean Absolute Percentage Error, MAPE) według następującego wzoru:

$$\delta = \left(\frac{1}{n} \sum_{t=1}^n \frac{|x_t - \hat{x}_t|}{x_t} \right) \cdot 100\%$$

Jeśli wartość błędu okaże się mniejsza niż 10%, interpretacja oceny dokładności może być zapisana jako wysoka. Współczynnik korelacji pomiędzy wartościami rzeczywistymi a prognozowanymi może być użyty jako dodatkowa ocena komparatywna. Im wskaźnik ten jest bliższy 1, tym silniejszy związek zachodzi pomiędzy

rzeczywistymi wartościami szeregu czasowego a wartościami przewidywanymi. Tabela 2 pokazuje wyniki oceny dokładności prognozy dla różnych parametrów wygładzania.

Pomimo niewielkich różnic MAPE w prognozowanych wartościach w obu przypadkach, co wskazuje na udaną adaptację modelu, preferowany jest wariant o najniższej wartości. Oznacza to, że predykcja przy automatycznie wybieranych parametrach wygładzania jest obiektywnie bardziej dokładna. Przedstawiona analiza sugeruje dość duże możliwości predykcyjne zaproponowanego, samodostosowującego się modelu adaptacyjnego, który z powodzeniem dostraja się do zmiennych warunków zewnętrznych. Jednakże wnioski końcowe można wyciągnąć tylko po porównaniu innych metod prognostycznych, zwłaszcza ekstrapolacji, która jest obecnie stosowana w strategiach rozwoju turystyki i w kalkulacjach wskaźnika celu. Jest to jednak dalsze zadanie, do podjęcia w następnym etapie badań. Bardzo interesujące jako kierunki przyszłego rozwoju są również: adaptacyjne modelowanie wielowymiarowe, łącznie z identyfikacją i wyborem czynników, oddziałującymi na wielkość i strukturę ruchu turystycznego, ocena ich wpływu na wskaźnik modelowania, a także identyfikacja najlepszej formy komunikacji i obliczania dla pożądaných parametrów modelu. Przyczynią się one do pogłębionych i bardziej szczegółowych badań nad prognozowaniem liczbowych cech rozwoju turystyki.

REFERENCES/BIBLIOGRAFIA

- Aena (2019). *Informes estadísticos*. Aena Estadísticas de tráfico aéreo. Retrieved from: <http://www.aena.es/csee/> (15.04.2021).
- Alexandrova, A., Aigina, E. (2020). *Strategii razvoitija turizma v stranah i regionalah mira*. Moskva: KnoRus.
- Andrianova, E., Golovin, S., Zykov, S., Lesko, S., Chukalina, E. (2020). *Obzor sovremennyh modelei I metodov analiza vremennyh ryadov dinamiki protsessov v sotsialnyh, ekonomicheskikh I sotsiotekhnicheskikh sistemah. Rossiiskii Tekhnologicheskii Zhurnal*, 8 (4), 7–45. <https://doi.org/10.32362/2500-316X-2020-8-4-7-45>

Tabela 2. Średni przybliżony błąd dla przewidywanych wartości przyjazdów turystycznych do Barcelony wyliczony w środowisku R przy różnych parametrach wygładzania

Poziom szeregu czasowego	Wartości rzeczywiste	Przewidywane wartości z parametrami wygładzania = 0,1	Wartości przewidywane przy parametrach wygładzania automatycznie wybranych w środowisku R
Zima 2019	10 041 360	10 646 518	10 077 340
Wiosna 2019	13 171 925	13 240 666	13 424 111
Lato 2019	15 876 816	15 190 951	15 955 689
Jesień 2019	13 426 888	13 583 932	13 613 453
		MAPE = 3,01%	MAPE = 1,04%
		współczynnik korelacji = 0,997	współczynnik korelacji = 0,999

Źródło: opracowanie własne.

- Atchade, M. (2018). *Adaptivnija metody prognozirovaniya: Realizatsija v Excel i programme R*. Saint Petersburg: Saint Petersburg State University.
- Atsalakis, G., Atsalaki, I., Zopounidis, C. (2018). Forecasting the success of a new tourism service by a neuro-fuzzy. *European Journal of Operational Research*, 268 (2), 716–727. <https://doi.org/10.1016/j.ejor.2018.01.044>
- Barchukov, I. (2008). *Metody nauchnykh issledovanij v turizme*. Moskva: Academia.
- Bosupeng, M. (2019). Forecasting tourism demand: The Hamilton filter. *Annals of Tourism Research*, 79, 102823. <https://doi.org/10.1016/j.annals.2019.102823>
- Chekmarev, A.V. (2020). Adaptivnoe modelirovanie intensivnykh i ekstensivnykh sostavlyajuschikh prognoznogo obraza. *Sovremennaya Ekonomika: Problemy I Resheniya*, 7 (127), 59–69. <https://doi.org/10.17308/meps.2020.7/2396>
- Chu, F.L. (2014). Using a logistic growth regression model to forecast the demand for tourism in Las Vegas. *Tourism Management Perspectives*, 12, 62–67. <https://doi.org/10.1016/j.tmp.2014.08.003>
- Davnis, V., Tinjakova, V. (2006). *Adaptivnija modeli: Analiz i prognoz v ekonomicheskikh sistemah*. Voronezh: Voronezh State University.
- Demin, A., Semenova, Y. (2001). Prakticheskoe ispolzovanie adaptivnykh modeley v turizme. *Kultura Narodov Prichernomoriya*, 16, 34–39.
- Ghalekhondabi, I., Ardjmand, E., Young, W.A., Weckman, G.R. (2019). A review of demand forecasting models and methodological developments within tourism and passenger transportation industry. *Journal of Tourism Futures*, 5 (1), 75–93. <https://doi.org/10.1108/JTF-10-2018-0061>
- Gladilin, V., Gladilin, A. (2016). Regressionnoe modelirovanie i prognozirovanie v turistsko-rekreacionnom komplekse regiona. *Innovatsionnaya Nauka*, 4 (1), 117–120.
- Jiao, X., Li, G., Chen, J.L. (2020). Forecasting international tourism demand: A local spatiotemporal model. *Annals of Tourism Research*, 83, 102937. <https://doi.org/10.1016/j.annals.2020.102937>
- Khaidi, S., Noratikah, A., Noryanti, M. (2019). Tourism demand forecasting – a review on the variables and models. *Journal of Physics: Conference Series*, 1366, 012111. <https://doi.org/10.1088/1742-6596/1366/1/012111>
- Lourenço, N., Gouveia, C.M., Rua, M. (2021). Forecasting tourism with targeted predictors in a data-rich environment. *Economic Modelling*, 96, 445–454. <https://doi.org/10.1016/j.econmod.2020.03.030>
- Magnus, Y., Katyshev, P., Peresetsky, A. (2004). *Ekonometrika: Nachalnyi Kurs*. Moskva: Delo.
- Nikolaeva, T., Oreshkina, E. (2016). Determinanty sprosa na vjezdnoi turizm (na primere stran Evropy i SNG). *Servis v Rossii i za Rubezhom*, 10 (8), 17–28.
- Pena, E.A., Slate, E.H. (2019). *Package 'gvlma'*. CRAN package gvlma. Retrieved from: <https://cran.r-project.org/package=gvlma> (25.06.2021).
- Song, H., Li, G. (2008). Tourism demand modelling and forecasting: A review of recent research. *Tourism Management*, 29 (2), 203–220. <https://doi.org/10.1016/j.tourman.2007.07.016>
- Svetunkov, I., Svetunkov, S. (2019). *Metody sotsialno-ekonomicheskogo prognozirovaniya*. Moskva: Urait.
- Tikhomirova, O. (2021). Adaptivnoe upravlenie predprinimatelskimi strukturami kak otkrytymi dinamicheskimi sistemami. *Fundamentalnye Issledovaniya*, 9 (2), 495–499.
- Yang, Y., Zhang, H. (2019). Spatial-temporal forecasting of tourism demand. *Annals of Tourism Research*, 75, 106–119. <https://doi.org/10.1016/j.annals.2018.12.024>
- Zhagina, S., Nizovtsev, V., Svetlosanov, V., Pakhomova, O. (2019). Problemy razvitiya turizma na territorii Evropeiskogo Severa Rossii. *Sbornik statei XII mezhdunarodnoi nauchno-practiceskoi konferentsii 'Innovatsionnoe razvitie sovremennoi nauki: problemy, zakonomernosti, perspektivy'* (pp. 101–104). Penza: MCNS Nauka i Prosvescheniye.