Forecasting US Tourists' inflow to Slovenia by modified Holt-Winters Damped model: A case in the Tourism industry logistics and supply chains

Dejan Dragan 1, Abolfazl Keshavarzsaleh2, Tomaž Kramberger1, Borut Jereb1, Maja Rosi1
1 University of Maribor/Faculty of Logistics, Celje, Slovenia
2 University of Malaya (UM) and International University of Malaya-Wales (IUMW), Kuala Lumpur, Malaysia

Abstract—Forecasting is important in many branches of logistics, including the logistics related to Tourism supply chains. With an increasing inflow of American tourists, planning and forecasting the US tourists inflow to Slovenia have gained far more importance attention amongst scholars and practitioners. This study, therefore, was conducted to forecast the American tourists inflow to Slovenia using one of the predictive models based on the exponential smoothing approach, namely Holt-Winters damped additive (HWDA) exponential smoothing method. The model was modified by several improvements, while the obtained results were generalized to other supply chain components. The results show that the forecasting system can predict well the observed inflow, while the methodology used to derive the model might have enriched the plethora of existing practical forecasting approaches in the tourism domain. Benchmarking demonstrates that the proposed model outperforms a competitive ARIMA model and official forecasts. The practical implications are also discussed in this paper.

Key words—Tourism industry logistics and supply chains, Tourism forecasting, US tourists and Slovenian Tourism, Damped Holt-Winters additive model, Time series analysis and prediction.

I. INTRODUCTION

The United States are an important global player in the worldwide tourism industry; not only have many tourists traveled to the United States but also many tourists from the United States make a journey to other countries. Indeed, advancements in Information Technology, mass media, transportation infrastructure along with globalization create both interconnected worldwide economic ecosystems and interlinked nexus of tourists who are being driven by many driving factors to choose the next destination. For example, tourists are drawn to Slovenia for its natural, cultural and historical attractions. A promotional Slovenian tourism destination identity campaign, so-called “I feel Slovenia” in April 2016 [1] has proven to be substantially successful in its mission. As a consequence, the Slovenian tourism industry started to boost up.

As globalization takes place, periodical socio-political alternates on the other side of the globe may lead to implicit or explicit progress or regress of another nation which is located on another continent. For example, two years ago, Melanie Trump has become a first “first lady” in the US with roots from Slovenia. Ever since then, perhaps completely coincidentally, already increasingly accelerated-growth of visits of American tourists has been even multiplied, presumably also due to Melanie’s popularity among the American people. Whether this progression in Slovenia’s tourism inflow is implicitly or explicitly linked to Melanie Trump or not, is beyond this study’s scope. However, the significance of planning and forecasting within Slovenia’s tourism context due to the increasing growth of US tourists need to be substantially addressed.

The proper planning and forecasting of the tourists inflow in a tourism supply chain are of the utmost importance research area. The higher the number of the tourists’ inflow, the more precise the planning and forecasting ought to be. Indeed, an ineffective tourism supply chain today will lead to various losses tomorrow [2]. Tomorrow’s losses might not only be appeared as a loss in revenue but also might be seen as the worst types of losses ever such as the emergence of...
inefficient tourism products/services and a damaged destination identity. The tourism industry is a complex industry; thereby, there are many factors and indicators linked to its supply chain components which cannot be simply ignored. To avoid, there is a need for being proactively productive when it comes to planning and forecasting.

This paper as a forecasting study aims at highlighting the importance of the role of increasing influx of US tourists into Slovenia. Perhaps the role of Slovenian first lady is also essentially important for such increased influx if presumed on the basis of different official and unofficial sources. However, unfortunately, the fact is that no quantitative study (e.g., grounded on a questionnaire-based survey among the US tourists in Slovenia) has been conducted yet until now to identify the main reasons for the coming of US tourists into Slovenia. In any case, we should put the spotlight on the fact that the official expectations about the future tourists’ arrivals from the US might be significantly underestimated if the role of Slovenian first lady would be truly confirmed, but deceitfully neglected.

This study was conducted to forecast the American tourists’ inflow to Slovenia using one of the predictive models based on the exponential smoothing, i.e., the Holt-Winters Damped Additive (HWDA) exponential smoothing model [3]. The model was modified by several improvements, while the obtained results were generalized to other supply chain components. This way, our study uniquely deployed a composed modeling framework in regard to forecasting through modeling design heuristic procedure prior to testing and validating the model. It is argued that the findings of the study can be used by practitioners and scholars who are interested in the application of forecasting within specific countries’ tourism industry context.

II. THE TOURISM INDUSTRY IN SLOVENIA AND THE FORECASTING CONTEXT OF THE STUDY

A. Characteristics of Slovenian tourism

Tourism is known to be an important economic activity globally due to its direct and indirect economic impacts. To elaborate, the direct economic impacts can be seen as any impacts on commodities, and tourism sub-industrial sectors while the indirect economic impacts can be seen as travel and tourism investment spending and the emergence of public-private partnership projects which are pertinent to tourism infrastructure and service ecosystem. Slovenia due to its strategic geographical location has always been seen as a great destination for tourists, particularly in the areas such as transit tourism, winter sport and, summer seaside vacations [4, 5].

Tourism in Slovenia is at the pivotal point of its history [1]. The tourism industry in Slovenia has gone through many modifications as times go by, more importantly, privatization. It is argued that this country although it is mature in tourism product development, it lacks a service culture. The privatization of the industry has positively contributed to the nations’ economy because private ownership performed much better regarding sustainability, internationalization, marketing strategies and, tourism product development [6-8].

B. The role of American tourists in Slovenia

The final fruit of all above-said efforts has led to attracting many tourists from all around the world, especially from the United States, particularly for the reason that they are on average the best consumers spending a lot of money. With an increasing number of American tourists’ inflow to Slovenia over the past two decades, this study conducted an improved exponential smoothing model for forecasting of the United States tourists’ inflow to Slovenia.

During the last few years, presumably also due to the fact, that the US has for the first time got a Slovenian first lady, the increase in arrivals of US tourists is even much more significant. This fact has been most likely caused by the fact that Slovenia has become widely recognized in the eyes of many US citizens. Also, the tourists who have already been in Slovenia non-stop spread a good opinion about this country all over America.

This increasing tourists’ inflow has led to a continuous research effort regarding the tourism industry in Slovenia and its tourism supply chain specifications. Since the launch of “I feel Slovenia” in which the country’s tourism has set to be a beaten heart of Slovenia’s identity, the Slovenian tourism industry has gained a tremendous boost [1]. Although it is boosting up as time goes by,
intensified competitions have impeded the progress towards reaching the tourism-led economic growth competitiveness. The competitiveness cannot be obtained overnight; thereby, planning and forecasting are golden keys to the future success of the Slovenian tourism industry.

C. The proposed forecasting approach

Having this in mind as the spotlight of our study, we feel responsible to contribute back to this sector through deploying and yielding the current practical forecasting modeling approaches within the tourism sector in Slovenia. To do so, the US tourists' inflow historical cumulative time series data to Slovenia[9] have been analyzed and predicted by using the well-known HWDA exponential smoothing model [3, 10] that was carefully investigated and improved with a special heuristic.

The standard Holt-Winters (HW) method can handle the time series with the trend, level, and a single cycle (i.e., the seasonal pattern). According to [3, 11], Taylor has in his three works [10, 12, 13] developed a multiple cycle version of Holt-Winters model, which is able to deal with the time series with longer forecasting horizons and multiple cycles. Since then, with the exception of work [11], at least to the best of our knowledge, practically none research can be found that would make some further relevant improvements regarding the HWDA model.

Thus, we believe that in this paper, after more than a decade, some essential enhancements have been conducted with respect to the basic version of HWDA model. Some of the ideas for these improvements have been initiated due to inspiration from some of our previous studies [14-16]. At this place, it must be emphasized that the described modeling approach for the HWDA model is generic, which means that at least in principle, it might have been used for any similar type of forecasting in any industry, as well as in any country or cluster of countries.

D. Possible contributions

It is argued that this study contributes to knowledge and practice in the following ways:

(a) After a longer time period of existence of the classical HWDA model, some essential enhancements have been conducted with respect to the basic version of HWDA model,

(b) The model is proven to be more or less effective for the case of the observed country. However, it could also be generalized to cover other supply chain members within a certain industry or industries in general irrespective of their size, type, and location. The latter means that a presented model could be due to its generalization conducted for any similar types of forecasting (e.g., the demand forecasting, sales forecasting, road freight transport forecasting, etc.).

(c) There are practically no similar studies identified in the tourism management context, introducing a specially composed modeling framework for the forecasting considering the proposed modeling design heuristic procedure,

(d) The proposed approach possesses a broad spectrum of the different rigorous criteria for model testing and validation that can rarely be found in other comparable studies, and

(e) Surprisingly, to the best of our knowledge, there were practically none studies detected that would systematically analyze the predictive models in any context regarding the tourism industry in Slovenia.

The derived model was identified within the quarterly based time interval (2003-2016). Afterward, during the benchmarking, the forecasts of our model have been compared with the real data from different sources (the actual data obtained from different tourism facilities), as well as with the competitive ARIMA model and the official forecasts for the year 2017 (such as from Slovenian Statistical Department (SSD) and Department for Macroeconomic Analyses and Development (DMAD)). Surprisingly, our model has on average significantly outperformed ARIMA model and official forecasts for 2017 (app. 85,000 US tourists) and approached much closer to the value of the real data (on average about 101,000 tourists).
E. Main motives for research

There have been two major motives for conducting this study. The first one was conceptual and related to the improved methodology regarding the modeling design of the conducted predictive model. The second one was applicative since we wanted to investigate to what extent the official forecasts are reliable and trustworthy regarding US tourists. Namely, the Slovenian Tourism is very important for the country economy since it represents about 13% of the total GDP.

Thus, any unreliable forecasts regarding any important aspect within the tourism industry can lead to inappropriate planning of future strategical-level tourism-related policies. These facts should also be considered for the forecasting of future inflows of important groups of tourists (from the most relevant countries). Since the US tourists play a more and more important role as “generously-spending” customers, the accuracy in forecasting is quite important, while significant miscalculations from the official institutions should have been strictly forbidden.

Furthermore, our suspicion about the credibility of forecasts of official institutions was grounded on the facts that in the recent past, the official institutions have already made many serious erroneous forecasts about pretty important things, such as inadequate forecasting regarding the GDP, the future liquidity of Slovenian banks and firms, and so on. These mistakes have sometimes caused millions or even dozens of millions of Euros of unnecessary losses.

Presumably, our model has much more accurately recognized a possible impact that popularity of Melanie Trump has had on an increased inflow of US tourists than it was detected by Slovenian official institutions. Naturally, maybe there are also some other additional reasons for such unusually enormous rise of the US tourists' inflow in just one observed year (from the end of 2016 to the end of 2017).

These reasons might be perhaps related to a good state of US economics’ indicators including the increased Consumption Price Index, or the global rise of the tourists’ visits worldwide. Yet, it is still unusual that even a 19.8% increase in tourists happens in just one year. Namely, if we denote a time-dependent inflow with a symbol $y(t)$, we are dealing with the situation:

$y(2016Q4) = 81000 \rightarrow y(2017Q4) = 101000 \rightarrow 19.8\% \text{ increase}$ of US tourists in only one year!

Also, before Trump was elected for the US president, Melanie’s hometown called Sevnica was just an average boring small town. Conversely, after his election, in just a couple of years, it has transformed into an important modern touristic spot. Even more, according to the official records (e.g., SSD: https://www.stat.si/StatWeb/Field/Index/24), the huge tourists' inflow increase has happened in Sevnica, while the majority of new visitors have been Americans!

III. LITERATURE REVIEW

A. The importance of the tourism industry for national economies

The tourism industry is arguably known as one of the main contributors to the countries' national economy. Since introduction of the Tourism and Travel Competitiveness Index (TTCI) by World Economic Forum (WEF) back in 2015, two criteria have been set to be significant in measuring the TTCI [17]; (a) quality of tourism supply/value chain and (b) characteristics of the destination country in itself. It was in this context that the competition occurred, emphasizing on infrastructure, cultural and natural resources indicators of the TTCI, for example, transportation infrastructure, tourism service providing ecosystem, cultural and sports attractions, and world heritage properties[18].

This revolutionary trend, namely Tourism-led Economic growth triggered a heated debate between the scholars whether the relationship between tourism and economic growth is bidirectional and interactive, if so, whether the previous research results are non-conflicting and consistent or no. An in-depth review of previous literature can be divided into two strands; re-looking to this concept from a holistic and detailed-oriented point of view. To elaborate some of the holistic-oriented studies, for example, an empirical study conducted in Spain revealed that
there are a unidirectional cause and effect relationships between tourism and economic growth [19]. In addition, a study was conducted in 2003 in which 13 OECD countries were investigated [20]; the results revealed that tourism and economic growth are positively interconnected. Moreover, another study conducted in Uruguay uncovered a relational attribution of tourism costs on GDP per capita [21]. Dritsakis in 2012 conducted a study covering seven Mediterranean countries. Results of the study provided a solid confirmation on the positive effects of tourism on GDP [22].

It seems that there is a firm consensus among scholars regarding the positive impacts of tourism on nations' economy across 55 various countries [23]. To briefly mention, studies conducted in Turkey [24], Tunisia [25], southern European countries [26], Singapore [27] and South Africa [28] have also shown the same results. However, scholars who used the detailed-oriented perspectives put the emphasis on the role of sub-industries in tourism supply chain and tourism economic growth networks at the sub-industrial level such as travel agents, hotels and airlines [29-32]. Therefore, it can be concluded that planning the infrastructure and deploying the forecasting models in any destination countries will pay off; however, there is a need to put more efforts on the models' efficiency, concomitantly to consider the inclusion of external exogenous effects as well.

B. Planning the infrastructure and the significance of forecasting in tourism

Competitiveness in tourism has its roots in numerous indicators such as destination country's economy, tourism sector infrastructure, tourists' income level, and other economic and political indicators such as buying power, currency rates, social openness, religious and culture, etc. Tourism industry is a complex industry sector that needs to be considered as a monolithic block in which its all supply chain components are matter, particularly the tourism infrastructure. It can be said that there is a robust relationship between tourism development and infrastructure [33].

The infrastructure base of each country can be seen as a determinant of that country's tourism destination attractiveness. Therefore, planning the infrastructure and the significance of forecasting in tourism industry context cannot be ignored anymore. Since the performance of the tourism sector is highly tied to socio-political factors [34], infrastructure development [35], immigration and visa policies [36], and political events [37]; thereby, there is a need for proper forecasting.

The forecasting in tourism research spheres has been highlighted in the previous literature extensively. For example, in regards to destinations' future tourism demand forecasting, many economic factors are considered including both macroeconomic factors and microeconomic ones [38, 39]. Therefore, it can be asserted that in order to obtain effective planning, the necessity of forecasting in tourism need to be intensified.

C. Forecasting approaches in the tourism industry

Forecasting has seen to be beneficial when it comes to the development and investment planning in tourism [40] as well as to demand fluctuation of tourism inflow [41]. There are many studies which deployed various types of forecasting models in tourism inflow forecasting and its other often-related areas. For example, a study conducted by Cho in 2003 aimed at forecasting tourists' arrivals in Hong Kong in which three predictive models were comparatively investigated [42]. Findings of the study revealed that ANN (artificial neural networks) is the most accurate one comparing to Univariate ARIMA and the exponential smoothing method.

Forecasting tourism arrival in Singapore was conducted by the Chu in 2008 in which the fractionally integrated ARIMA models were investigated and ultimately comparatively compared with the traditional ARIMA models [43]. In another study, which was conducted by the same author in 2011 considering Macau as a destination, fractionally integrated ARIMA models, seasonal ARIMA and a piecewise linear model were investigated [44]. The findings of the study introduced a piecewise linear model as the most effective and accurate among others.

The new forecasting model, namely TVP-STSM was developed in a study conducted by Song & Collogues in order to both model and forecast quarterly Chinese, Korean, British and American tourist arrivals to Hong Kong [45]. The empirical results, according to the authors, indicate that the proposed combination of TVP (time-varying parameter) and STSM (structural time series model) works appropriately compared to related traditional ones.
Regarding U.S tourist arrivals, SSA (Singular Spectrum Analysis) using monthly data for tourists inflow to the U.S between the period 1996 and 2012 were investigated in a study conducted by [46]. The results of their study concluded as SSA can be an accurate approach for forecasting tourism demand and it worth to deploy in similar cases.

It can be said that there have been too many forecasting techniques so far. To categorize, quantitative forecasting models which have been developed so far can be seen as Artificial Intelligence-based, time series models and the econometric methodical approaches. Time series models which are dependent to previous data in the series to forecast the future trends, can be divided to some sub-clusters including; SMA (Simple Moving Average), Navi, SES (Single Exponential Smoothing), DES (double exponential smoothing), ARIMA (autoregressive moving average) and BSM (basic structural time series).

Although they are proven to be effective forecasting models when it comes to forecasting tourism inflow, they are substantially limited to non-economic factors [47]. This means the significance of tourists' behavioral attributions is not configured in these models. Some scholar, therefore, argued that the econometric models are performing better, assisting policymakers in assessing the effectiveness of the policies and strategies deployed [48].

An in-depth review of the literature indicates that there is various way to forecast tourism demand/inflow or other often-related areas. Some models outweigh others in some cases, depending on various indicators either endogenous or exogenous. For example, the standard ARIMA aims at forecasting according to the past values of the forecast variable while its extended version, known as ARMA(X) includes another predictor (independent) variables in order to enhance the accuracy of a forecast in itself.

The ARIMA(X) is deployed in many studies such as forecasting Turkey's tourism-led revenue [49] forecasting the international tourism demand in Japan [50] and so on. Moreover, Holt-Winters damped additive exponential smoothing model could have positioned itself among other forecasting models due to its robustness and enhanced accuracy. This approach deployed in many studies which address long-lead times such as long-lead time forecasting of UK air passengers [51] forecasting the international tourist arrivals to New Zealand and Australia from 11 destinations [52] monthly volume of tourism inflow into Bulgaria [53] etc.

Somehow, to complement the literature, in this study we aim at introducing our forecasting technique as some counterpart competitive approach to other more common forecasting approaches and screen it against other models. Slovenia is considered as the targeted destination country in this study due to its emergent tourism industry which it attracts many tourists from all around the globe, more importantly, from the U.S. Regarding our modified HWDA forecasting model, contrariwise to its basic version, a number of criteria must be considered such as; (a) the chosen model must be able to provide an accurate forecasting according to the direction on any possible changes occurred and (b) last but not the least, the chosen forecasting model must be able to consider short-run and long-run intervals.

IV. THE HISTORICAL DATA

A. The historical data for US tourists’ inflow to Slovenia

Slovenia has been the destination of many tourists as time goes. The Americans are not an exception, and more importantly, besides the most essential visitors from the EU, they are considered the most significant contributors to the Slovenian tourism industry. Indeed, America is categorized as the distant marketplace for Slovenians and vice versa [54]. Recently, Slovenia expanded and intensified its promotional tourism destination campaign in the USA.

Maja Pak, the Director of the Slovenian Tourist Board, indicated that the media’s interest and the publicity have been a great advantage in order to intensify the Slovenian tourism destination promotional campaigns. At the second quarter of 2016, Slovenia was the designation of 78209 American guests. Compared to the second last quarter of 2015 when Slovenia was visited by 70488 American guests, there is an increase of 10.9 percent.
According to the estimation which was made by the Director of the Slovenian Tourist Board, 2017 has been a witness of 5 to 6 percent increase in American guests' arrivals; statistically, from the end of 2016 which the numbers of guests were about 81,167 to about 85,000 at the end of the year 2017.

Simultaneously, similar estimates (about 85,000) have been made from some other government institutions (SSD, DMAD). According to the governmental reports, the capital city of Ljubljana has been witnessing of many visitors as well. Interestingly, the American tourists were ranked in a fourth high spot regarding the overnight stay's statistics just before visitors from UK, Germany, and Italy.

For the guests from the USA, it is also evident that they conduct on an average far biggest number of overnight stays compared to the tourists from the other distant markets. They also hold the primate about spending the biggest amount of money for direct tourism-related expenses, as well as different indirect expenses (among all tourists in Slovenia as a whole!). Fig. 1 shows a rough estimate of US tourist arrivals, and comparison to their overnight stays for years 2006-2015 [55, 56].

Figure 1. US tourist arrivals and comparison to their overnight stay for years 2006-2015.

Fig. 2 illustrates the quarterly cumulative inflow of US tourists between 2006 and 2016 (SSD: https://www.stat.si/StatWeb/Field/Index/24). As can be seen, both volumes of arrivals and overnights stay in Slovenia has been growing gradually, except a sharp plunge in both volumes of arrivals and overnights stay in Slovenia 2009 due to the eruption of the economic crisis in 2008. However, later it has captured a steady increasing trend again. For the years 2013 to 2016, the inflow has reached cumulative values (at the end of the fourth quarter): 60,224, 66,004, 74,759, and 81,167 tourists, respectively, thus percentage increases in values: 9.5%, 13.26% and 8.5% respectively. The time series shown in Fig. 2 is also the targeted output variable $y(t)$, for which we want to design our forecasting models.
V. METHODOLOGY FOR A FORECASTING MODEL

A. The conceptual modeling framework

Fig. 3 illustrates a conceptual modeling framework for our research. In block B, we can see the available collected data for the tourists’ inflow time series \( y(t) \) that represent a basis for the whole research. Block C refers to the basic version of a treated HWDA model, whose output \( \hat{y}(t) \) is preferred to closely follow the time series \( y(t) \) disturbed with the random noise \( \varepsilon(t) = \varepsilon_y(t) \).

In the modeling process, the basic model enters into the main advanced heuristic framework (block D, stage 1), which is optimized in such a way that can process thousands of model candidates in an acceptable amount of computation time. Further, for a HWDA structure that is always fixed, a wide set of possible parameter sets is settled, from where a different parameter set is assigned to each model candidate in each iteration \( j \) of the procedure (block E, stage 2). In the third stage (block F), the diagnostic checking is conducted, and goodness of fit (GOF) measures are calculated for each model candidate. Here, particularly important is a model candidate’s error \( e(t, j) = y(t) - \hat{y}(t, j) \) that is different in each iteration \( j \) depending on the specific calculated model’s output \( \hat{y}(t, j) \).

A special sub-heuristic is developed during the model selection process which is employed in the fourth stage (block G) in order to obtain and find the best model comparing to many other model candidates. To do so, different pre-defined rules and statistical ones for each candidate have also been deployed. Such heuristic gives us the final most adequate HWDA model, which can be considered as the best one. Further, besides providing the well model’s fit to the real data \( y(i) \), the final obtained model satisfies all other rigorous mathematical and statistical conditions, particularly those related to the residual-based criteria. This way, when the modeling procedure is ended, we obtain the best model with a corresponding output: \( \hat{y}(t, j^*) = \hat{y}_{\text{HWDA}}(t, j^*) = \hat{y}^*, j^* - \text{iteration with best criteria fulfilled} \), for which the optimal error \( e^* = y - \hat{y}^* \) can be obtained (block H in Fig. 3)). Afterward, we can analyze the forecasting
performance of the best model’s output $\hat{y}(t)$ on the basis of different model-error based criteria (block I).

B. Basic Holt-Winters damped additive exponential smoothing model (HWDA)

In general, exponential smoothing (ES) methods refer to a special class of forecasting methods. There exist a whole plethora of methods that belong to the exponential smoothing class of methods. Their major property is that forecasts are weighted combinations of past observations, with more recent measurements being relatively more weighted than the older ones. The title “exponential smoothing” implies the fact that the weights decline exponentially as the measurements get older.

Generally, we are dealing with three basic types of ES methods, i.e., the single ES (Brown’s method), the double ES (Holt’s method), and the triple ES (Holt-Winters method). The latter is suitable when we presume some evident trend and seasonality in the observed time series. According to the empirical evidence, the basic Holt-Winters method has a tendency to make over-or-under forecasts, particularly for longer forecasting horizons. For this reason, Gardner and McKenzie (1989) have applied a new parameter $\phi$ associated to the trend component that damps the trend to a flat line, when the future becomes more distant. This way, we can obtain the Holt-Winters damped additive method (HWDA) in the following form with four parameters $\alpha, \beta, \gamma, \phi$: 

![Diagram of the conceptual modeling framework](image-url)
Level: \[ \ell(t) = \alpha (y(t) - s(t-m)) + (1-\alpha) \left( \ell(t-1) + \phi b(t-1) \right) \]

Growth: \[ b(t) = \beta \left( \ell(t) - \ell(t-1) \right) + (1-\beta) \cdot \phi b(t-1) \]

Seasonality: \[ s(t) = \gamma (y(t) - I(t-1)) - \phi b(t-1) + (1-\gamma) \cdot s(t-m) \]

Forecast: \[ \hat{y}(t+h|t) = \ell(t) + \phi b(t) + s(t-m + h^*_m), \]

where: \[ \phi = \phi + \phi^2 + ... + \phi^h, \]

\[ h^*_m = [(h-1) \mod m] + 1, \quad m \cdot \text{number of seasons} \cdot \text{year} \]

\[ h \cdot \text{time points of the future horizon} \]

Conversely to the other exponential smoothing methods that can be linked with a linear Box-Jenkins methodology and state space approach, the HWDA is nonlinear in its nature [3].

C. A brief discussion of the advanced heuristics conducted in the modeling mechanism

Modeling mechanisms related to the HWDA model were extensively explained in some of the previously conducted studies [14-16]. However, the corresponding heuristics were in previously conducted researches implemented inside the wider structure of the so-called DFM-ARIMAX model (ARIMAX model, whose inputs were the dynamic factors from the dynamic factor analysis). Moreover, since then, many additional novelties have been engaged in this research.

If we carefully observe Fig. 3 and the description of an HDWA model in the previous section, we can see that the HWDA model has a fixed structure with four parameters \( \alpha, \beta, \gamma, \phi \) (see (1)). If we are generating a family of HWDA models by changing the sets of parameters \( \alpha, \beta, \gamma, \phi \in [0 + \varepsilon, \Delta j, 2\Delta j, ... 1 - \varepsilon], \Delta j \rightarrow 0, \varepsilon \rightarrow 0 \), we can obtain a huge group of model candidates.

The more in detail illustrated working mechanism of an applied heuristic for stages 3 and 4 from Fig. 3 is shown in Fig. 4. As can be seen from Fig. 4, three types of tests were first calculated at each iteration of the procedure for each model candidate. Stage 3 has covered the computation of different statistical tests, the residual-based tests (for model’s error \( e \rightarrow e_j(t) = y(t) - \hat{y}_j(t), j-th \text{ iteration} \)), and the future dynamics’ tests (checking of the future out-of-sample (FOS) forecasts: \( \hat{y}_j(t+h) \leq \text{thresholds}, h \cdot \text{FOS horizon} \)).

Also, the model’s error of each candidate was carefully investigated to check whether it holds approximate properties of the normal white noise or not. Immediate exclusion of candidates with inadequate future responses, i.e., \( \hat{y}_j(t+h) \gg \text{thresholds} \) was needed to avoid unusual or impossible future situations compared to the past trend’s dynamics of \( y(t) \). We have also applied so-called Dynamic Time Warping (DTW) giving a DTW value [62, 63] that have measured the “distance” between two signals \( y(t), \hat{y}(t) \) by using the Kullback-Leibler distance:

\[ d[y, \hat{y}] = \sum_{t=1}^{N} \left[ y(t) - \hat{y}(t) \right] \cdot \left[ \log y(t) - \log \hat{y}(t) \right]. \]

The testing of DTW values for different candidates was essentially important since it has in a refined way determined how truly “far” among each other are the signals \( y(t), \hat{y}(t) \) [63].

Let us emphasize that the calculations in stage 3, as well as a reducing procedure for excluding of inappropriate candidates in stage 4, were carried out by very sophisticated heuristics based on the sequence of carefully designed consecutive steps. This way, despite the significantly enormous amount of model candidates, the whole model selection procedure to find the best model candidate was executed in a reasonable computational time. A demand for simultaneously fulfilled all relevant tests in stage 4 was so strict that only a few candidates remained in the reduced set. In the final step, the best candidate was chosen according to the best \%FIT, i.e., the best fit.
\[ \hat{y}_j(t)_{best} = \hat{y}(t, j^*) = \hat{y}_{IWDA}(t, j^*) = \hat{y}^*(j^* - \text{iteration with the best criteria fulfilled}) \]

of the model’s output to the measured time series \( y(t) \).

Figure 4. The more in detail illustrated working mechanism of an applied heuristic for stages 3 and 4 from Fig. 3.

Regarding all observed error-based and other criteria, of which some are not depicted in Fig. 4, the following set of expressions can be given for each iteration \( j \) [63-66]:

\[
MSE(j) = \frac{1}{n} \sum_{t=1}^{n} e^2(t, j), \quad RMSE(j) = \sqrt{MSE(j)}, \quad MAE(j) = \frac{1}{n} \sum_{t=1}^{n} |e(t, j)|
\]

\[
MAPE(j) = \frac{1}{n} \sum_{t=1}^{n} \left( \frac{|e(t, j)|}{y(t)} \right) \cdot 100, \quad \max_{err}(j) = \max |e(t, j)|
\]

\[
RMSE_{relative}(j) = \frac{RMSE(j)}{\max(y(t)) - \min(y(t))}, \quad MAE_{relative}(j) = \frac{MAE(j)}{\max(y(t)) - \min(y(t))}
\]

\[
%FIT(j) = 100 \left[ 1 - \frac{\|y(t) - \hat{y}(t, j)\|}{\|y(t) - \text{mean}(y(t))\|} \right]
\]

\[
DTW[d(j)] = DTW\left[ d\left[ y, \hat{y}_j \right] \right] = DTW\left[ \sum_{t=1}^{N} \left[ y(t) - \hat{y}_j(t) \right] \cdot \left[ \log y(t) - \log \hat{y}_j(t) \right] \right]
\]

Here, \( \max_{err} \) refers to the maximum absolute value of the \( e(t, j) \), while

\[
\Delta y = \max(y(t)) - \min(y(t)) = 81167 - 29647 = 51520 \quad (\text{see Fig. 2}) \text{ reflects the dynamic range of}
\]
the \( y(t) \). The interested reader can find the exact form of expressions for all other criteria (e.g., JB test, LB test, etc.) from Fig. 4 in the appropriate statistics based, time-series based, and econometrics literature.

VI. PRACTICAL NUMERICAL RESULTS

MATLAB, a technical computing environment, was deployed in order to calculate all the results. In order to extract the best model, the MATLAB environment was carried out in order to both generate and then identify the HWDA model candidates. Moreover, the Econometrics Toolbox, along with Machine Learning & Statistics Toolbox were considered to carry out the statistical testing and diagnostics of the model candidates. The System identification toolbox was used to derive the benchmarking ARIMA model. In order to merge all parts of the modeling process, MATLAB basic environment was considered.

A. In-Sample estimation and validation results for the best HWDA model

In this section, we discuss about the in-sample estimation and validation results for the best HWDA model, which means the results that are referring to the estimation and test interval. Fig. 5 depicts the prediction results for the best HWDA model on those two intervals, as well as forecasts on the future out-of-sample (prediction) interval. In Fig. 5, the comparison between the observed time series \( y(t) \) and the estimated forecasts \( \hat{y}_j(t)_{best} = \hat{y}(t,j^*) = \hat{y}_{HWDA}(t,j^*) \) for the best model is provided. The separation between the "estimation interval" and the "test interval" is done in order to distinguish between the first 40 observations used to estimate the smoothing parameters \( \alpha, \beta, \gamma, \phi \), and 16 observations used for testing the predictive power of the best HWDA model. Also, the benchmarking with the ARIMA model and official forecast for the year 2017 is shown in Fig. 5 (discussed later in the following two sections). As it turns out, the estimated values for the smoothing parameters of the best HWDA model (see 1) are:

\[
\begin{align*}
\alpha^* &= \alpha(j^*) = 0.701 \\
\beta^* &= \beta(j^*) = 0.922 \\
\gamma^* &= \gamma(j^*) = 0.308 \\
\phi^* &= \phi(j^*) = 0.971
\end{align*}
\]

These parameter values were calculated at the "best" iteration \( j^* \), where the following "best" criteria were simultaneously achieved (see Fig. 5):
The results shown in (4) implicate that a percent of fit 78.66% to the real data was fairly good (c.f. Fig. 5). Also, the output of the best model has never shown some more serious deviations from the real data, since $\text{MAPE}(j^*) = 2.8516\%$ was quite small, i.e., far below 10%, which is usually an acceptable threshold suggested by many researchers, (e.g., [67, 68]). Furthermore, the measures $\text{RMSE}_{\text{relative}}(j^*) = 0.029722$, and $\text{MAE}_{\text{relative}}(j^*) = 0.025711$ were quite low as well. Even more importantly, from the practical engineering point of view, a maximum error has stayed within 5% of the maximum dynamic range of the $y(t)$ (i.e., $\max_{\text{err}}(j^*)/\Delta y = 2671.9/51520 = 0.051$).

Furthermore, as it turned out, all fulfilled diagnostic error-based tests (e.g., J B test, LB test, Q-Q test, ACF and PACF test) has shown that the best HWDA model’s error $e(t,j^*)$ approximately follows the required properties of the normal white noise.

To conclude, our final HWDA model provides a relatively good fit to the real US tourists’ inflow data, although it does not have such sophisticated working mechanism as some more advanced time-series models (e.g., Box-Jenkins family of models [69]. Furthermore, our model does not suffer from any significant overfitting and inadmissible oscillatory behavior at some specific time points as for example the other classical exponential smoothing models, including the basic version of an HWDA model.

It is true that there are some more sophisticated details of the real data dynamics detected that are not covered by our model. We believe that two major reasons are responsible for this fact:

- Firstly, the observed time series seems to contain a quite complex nature of its dynamics, which could not be captured with our model possessing a relatively simple structure. The stochastic mechanism that generates the time series most likely contains certain significant nonlinearities, while there might also be some other important external effects that have not been modeled. Although such phenomena remain un-modeled, our model still offers a useful tool for forecasting of the major variations in the time series dynamics.
- Secondly, the compendium of all required criteria for the selection of our final model was extremely rigorous, particularly concerning the requirements about model’s error to be white noise, so the best fit to the data was not the only criterion. Nevertheless, despite this, the final model provides a satisfactorily well fit, particularly concerning the main trend’s movements.
Benchmarking with the ARIMA model and official forecasts on the prediction interval

Table 1 shows the benchmarking results on the prediction interval, i.e., the comparison of our HWDA model with competitive ARIMA model and official forecasts (c.f. also Fig. 5, the year 2017). Here, ten future predicted quarterly values in the period from 2017Q1 to 2019Q2 are given for the HWDA model and for the ARIMA model. As can be seen from table 1, our HWDA model (value 92451) has at the end of the last quarter of the year 2017 (Q4) by far outperformed the competitive ARIMA model (value 85434) and official forecasts (value approx. 85000) and approached much closer to the real value of approximately 101.100 of US tourists. On the other side, it can be noticed that the forecasts of the ARIMA model possessed a much slower prediction dynamics, i.e., their increasing was far too sluggish.

Table 1. The benchmarking results on the prediction interval (the comparison of the HWDA model with competitive ARIMA model and official forecasts).

<table>
<thead>
<tr>
<th>Year</th>
<th>Quarter</th>
<th>HWDA</th>
<th>ARIMA</th>
<th>Official Forecasts</th>
<th>Real value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>Q1</td>
<td>84735</td>
<td>82372</td>
<td>nr</td>
<td>nr</td>
</tr>
<tr>
<td>2017</td>
<td>Q2</td>
<td>86277</td>
<td>83511</td>
<td>nr</td>
<td>nr</td>
</tr>
<tr>
<td>2017</td>
<td>Q3</td>
<td>89642</td>
<td>84528</td>
<td>nr</td>
<td>nr</td>
</tr>
<tr>
<td>2017</td>
<td>Q4</td>
<td>92451</td>
<td>85434</td>
<td>Approx. 85.000</td>
<td>Approx. 101.100</td>
</tr>
<tr>
<td>2018</td>
<td>Q1</td>
<td>93133</td>
<td>86243</td>
<td>nr</td>
<td>nr</td>
</tr>
<tr>
<td>2018</td>
<td>Q2</td>
<td>93810</td>
<td>86964</td>
<td>nr</td>
<td>nr</td>
</tr>
<tr>
<td>2018</td>
<td>Q3</td>
<td>96544</td>
<td>87608</td>
<td>nr</td>
<td>nr</td>
</tr>
<tr>
<td>2018</td>
<td>Q4</td>
<td>99629</td>
<td>88182</td>
<td>nr</td>
<td>nr</td>
</tr>
<tr>
<td>2019</td>
<td>Q1</td>
<td>$1.0062 \times 10^5$</td>
<td>88694</td>
<td>nr</td>
<td>nr</td>
</tr>
</tbody>
</table>
Regarding the developed ARIMA model, the latter contained the following estimated structure in the form of the transfer function (with the backshift operator $q^{-1}$ and noise $\varepsilon_y(t)$) [70, 71]:

$$A(q) \cdot A_{\text{ARIMA}}(q) = A(q) \cdot \left[ \hat{y}_{\text{ARIMA}}(t) - \hat{y}_{\text{ARIMA}}(t-1) \right] = C(q) \cdot \varepsilon_y(t)$$

$$A(q) = 1 - 0.89205 q^{-1}$$

$$C(q) = 1 + 0.97 q^{-1} + 0.9654 q^{-2}$$

$$| t(a_1) |, | t(c_1) |, | t(c_2) | > t_{\text{krit}} \approx 1.97$$

- Residual tests confirm the approx. white noise of model's error
- All zeros and poles of transfer function inside the unit circle
- All t values (in parenthesis) statistically significant

Concerning the methodologies of official forecasts, they are usually never revealed to the public. However, based on source [72], the forecasts in the tourism sector are usually based on Box-Jenkins family of models, with particular emphasis on the ARIMA and similar models.

C. Out-of-Sample prediction results for the best HWDA model

Fig. 5 also shows a predictive performance of our model on the future prediction interval from the first quarter of 2017 to the end of the second quarter of 2019, i.e., forecasts $\hat{y}_{j} (N + h), N = 56, h = 1,2,\ldots,10$ quarters for ten quarters ahead (from the end of 2016). The more precise details of this predictive behavior are shown in Fig. 6, where the prediction results for the best HWDA model are enlarged focusing exclusively on a prediction interval.

As we can see from Fig. 5 and 6, our model gives quarterly forecasts $\{84735, 86277, 89642, 92451\}$ tourists for the year 2017, quarterly forecasts $\{93133, 93810, 96544, 99629\}$ tourists for the year 2018, and two forecasts $\{1.0062 \cdot 10^5, 1.0076 \cdot 10^5\}$ tourists for the first two quarters of the year 2019.

If we are now firstly focused on the (end of the) year 2017 (c.f. Fig. 5), we can clearly see that our model has predicted the value $92451$, while we have seen before that the competitive ARIMA model and the official institutions have predicted significantly under-estimated values. Thus, our model was able to much more precisely identify the amplified rising trend of the influx’s dynamics than it was predicted from the side of the Slovenian official institutions.

However, it has in rough terms predicted the true value with a time delay of about one year (c.f. the last two forecasts $\{1.0062 \cdot 10^5, 1.0076 \cdot 10^5\}$). Furthermore, the role of the possible future new crisis cannot be ignored and there might be a declining trend as well. Namely, surprisingly, the scope of the predicted trend seems to decline if looking at the last three forecasts $\{99629, 1.0062 \cdot 10^5, 1.0076 \cdot 10^5\}$.

Thus, the implications of the forecasts for the forthcoming years till the end of the year 2019 or later can seriously worry us. Maybe we can even conclude that there exists certain possibility that some substantial (global?) negative macroeconomic behavior will happen again in the near future, perhaps in the US and/or the EU, or worse, worldwide. These fears can also be detected.
from several other works (such as [73-75]), where scholars emphasize scares about new approaching dangerous economic events.

Namely, in a couple of studies, a new economic stagnation or recession is forecasted somewhere in the period 2019-2020 [74, 76], while some academics even warn about an outbreak of the new global economic crisis [75]. Their expectations are based on various analyses such as for example “what if” scenario playing analysis [73, 76], where China’s declining economic performance represents the biggest concern. The latter is even amplified after the eruption of serious economic war that was initiated by the USA recently against some other countries, where China seems to be a primary target [74].

Figure 6 - The prediction results for the best HWDA model (enlarged details of the prediction interval).

D. Main findings and implications of the study

Based on the results of this study, we can express some main findings and implications. The fact is that we have designed a modified working HWDA model driven by additional specially designed heuristics that helped to overcome the deficiencies of relatively simple exponential smoothing structure.

Our model has significantly outperformed the competitive ARIMA model and official forecasts for the year 2017, although the latter are usually based on much more complex models. We might have expected that this cannot be possible. Namely, due to [77], in times of the last economic crisis, the global nature of the time series has become much more complex with amplified volatility. Consequently, since then, more sophisticated forecasting models are needed to incorporate a bigger complexity of the time series, as models with a simpler structure are not good enough anymore.

Thus, we cannot imagine the reason why the official models have completely failed. Maybe there was some kind of misinterpretation of results or misuse of the models? Nevertheless, it was confirmed again that in general, we should not have blindly followed official prognosis with a hundred percent trust.

Namely, the fact is that our model managed to capture a 19.8% increase in the tourists’ inflow in 2017 much more precisely than the competitive forecasters. Furthermore, we are aware that we have mentioned only indirect assumptions that such large increase of the US tourists’ inflow in such short time has happened also due to an important influence of Melanie Trump on US citizens.

But whatever the reason would be, the contributions in this paper related to the methodological novelties of the designed HWDA model with a fairly good forecasting capability cannot be neglected. Moreover, this study was by our best knowledge the first in-depth attempt of
research after longer time-period that have tried to improve the capabilities of the basic HWDA model. Fairly good forecasting results were achieved despite the fact that the predictions were based only on the historical data of the targeted inflow time series.

Thus, the model was derived without screening the other time series, e.g. those reflecting the US economy (macroeconomic indicators), as the Box-Jenkins ARIMAX and other similar models are capable of administering [14-16, 69]. In the future work, we intend to somehow incorporate more concrete quantification of Melanie’s influence on the US tourists’ behavior, maybe by integrating of additional time series reflecting the time-dependent percentage level of her popularity among US citizens.

Moreover, to make such quantitative research even more credible, it is also planned to apply a questionnaire-based survey among the US tourists in Slovenia in order to identify the truly main reasons for their visiting of Slovenia. Also, we are planning to engage the US macroeconomic indicators’ time series by means of ARIMAX and other more sophisticated models that can involve the exogenous inputs, as well as intervention and other hidden effects in the time series. By simultaneously using additional variables in the model and adopting of similar heuristics as were presented in this paper, we expect to achieve even more adequate predictive behavior.

VII. CONCLUSION

In the paper, we have developed a forecasting module to predict the cumulative inflow of US tourists to Slovenia. The modeling framework was designed in such a way that the classical Holt-Winters Damped model was upgraded with special heuristics to improve the forecasts of its basic exponential smoothing model’s counterpart. In the modeling procedure, a sequence of carefully selected and categorized rigor criteria was adopted to find the best model among the large group of model candidates.

The derived best model has achieved fairly accurate prediction results, particularly regarding the main trend’s movements. As it turned out for In-Sample interval, it achieved even a 78.66% fit to the real-time series data, while the largest model’s error did not exceed the 5% of the dynamic range of the observed time series. Regarding the Out-of-Sample Interval, our model has predicted 92451 tourists for the end of 2017, while the ARIMA model and official forecasts have achieved much worse results, i.e. 85434 and 85000 tourists, respectively. Thus, our model has significantly outperformed the competitive ARIMA model and official forecasts and have got much closer to the real value of about 101000 tourists.

Besides the methodological issues, the particular emphasis was dedicated to the discussion about the American tourists, and presumably the important role Melanie Trump has had regarding an inflow’s increase of even 19.8% of the tourists in just one year. However, the implicit or explicit linkage between growth in tourism inflow and Melanie Trump were emphasized to be outside this paper’s study scope. Indeed, we emphasized that any periodical socio-political alterations across a globe may substantially influence all interconnected globally economic ecosystems including tourism due to advancements in Information Technology, Transportation infrastructure, mass media, etc.

The study has also debated about a growing role that the US tourists have within the scope of Slovenian tourism. Thus, particular attention should be needed to keep persisting doing on enforced tourism marketing to attract even more American tourists. Namely, they are known as greeted tourists and large consumers spending a lot of personal finances to discover all the natural and other beauties that Slovenia can offer to them.

REFERENCES


AUTHORS

A. Dejan Dragan, Ph.D., is the Associate Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: dejan.dragan@um.si).

B. Abolfazl Keshavarzsaleh, M.Sc., M.Sc., is the Teaching and Research Assistant at the University of Malaya (UM) and International University of Malaya-Wales (IUMW), Kuala Lampur, Malešya (e-mail: abolfazl.keshavarzsaleh@gmail.com).

C. Tomaz Kramberger, Ph.D., is the Associate Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: tomat.kramberger@um.si).

D. Boroj Jereb, Ph.D., is the Associate Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: borut.jereb@um.si).

E. Maja Rosi, Ph.D., is the Teaching Assistant at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: maja.rosi@um.si).

Manuscript received by 12 April 2019.

Published as submitted by the author(s).