Forecasting Tourism Demand with Denoised Neural Networks

October 26, 2018

Abstract

The automated Neural Network Autoregressive (NNAR) algorithm from the forecast package in R generates sub-optimal forecasts when faced with seasonal tourism demand data. We propose denoising as a means of improving the accuracy of NNAR forecasts via an application into forecasting monthly tourism demand for ten European countries. Initially, we fit NNAR models on both raw and denoised (with Singular Spectrum Analysis) tourism demand series, generate forecasts and compare the results. Thereafter, the denoised NNAR forecasts are also compared with parametric and nonparametric benchmark forecasting models. Contrary to the deseasonalising hypothesis, we find statistically significant evidence which supports the denoising hypothesis for improving the accuracy of NNAR forecasts. Thus, it is noise and not seasonality which hinders NNAR forecasting capabilities.

Keywords: Neural Networks; Singular Spectrum Analysis; Denoising; Signal extraction; Tourism demand; Europe.

1 Introduction

The prediction of tourism demand is one of the most interesting and important areas of research in tourism studies. The existing literature on forecasting tourism demand is wide, ranging in terms of the different countries considered, the various statistical techniques applied and the different set of data employed. A recent and comprehensive review of new developments in tourism demand modelling and forecasting is included in the paper by Wu et al. (2017). The paper considers various data measurement of tourism demand and their determinants. It also discusses the methodological development of three types of forecasting methods, classical times series techniques, artificial intelligence and econometric methods. This article covers 171 studies published from 2007 to 2015 in the field of tourism and hotel demand modelling and forecasting.

As online consumer data through Google and Baidu search engines have become increasingly available and with the fast development of internet and computer science, many researchers have recently started to use big data to improve their forecasts. Yang et al. (2015) used ARMA models to determine whether Google and Baidu search engine data improve tourism demand forecast in China. Bagwayo-Skeete and Skeete (2015) seeks to determine whether Google data can be an indicator for tourism forecasting. They compared Autoregressive Mixed-Data sampling models with Autoregressive models and Seasonal ARIMA models and found that Google trends does in fact help forecast tourism demand. Artola et al. (2015) also evaluated whether internet searches can help predict tourism inflows into Spain. Likewise, Pan et al. (2012) also employed data on five related Google search volume using ARMA with an exogenous model to predict demand for hotel rooms. Li et al. (2017) used the search engine query volumes data to forecast tourism demand for Beijing, using a composite search index. They found that the proposed method outperforms the traditional time series model and a model with an index created by principal component analysis. Liu et al. (2018) applied Vector Autoregressive models to investigate the Granger causality between weather temperature and the web search queries of the destination. They found no significant correlation between weather temperature of the tourism destination and actual arrivals or with its web search queries. Using mixed frequency data, Hirashima et al. (2015) used monthly indicators such as monthly tourist arrivals, monthly passenger counts, and monthly airline passenger outlook to forecast quarterly tourist arrivals to Hawaii. They examined a number of mixed frequency econometric techniques that incorporate high frequency information in the forecasting process and found improvement in forecasting, using high frequency indicator data. Those interested in a recent and comprehensive survey on different types of big data used in tourism research are referred to Li et al. (2018).

Forecasting seasonal variations in time series has been an important, yet difficult task for forecasters both historically (Zhang and Kline, 2007; Hamzacebi, 2008) and in the modern age. Given that seasonality and volatility drives significant movements in monthly tourism demand (Hassani et al., 2015; Claveria et al., 2015), various evaluations have been made using a variety of techniques in the quest for improving tourism demand forecasts. To this end, the emergence of Big Data and the associated interest in Data Mining techniques have increased the importance and applications of Neural Networks (NN) as a forecasting technique.

As evidenced below, as a means of forecasting tourism demand, the NN technique is no stranger to the tourism industry. In brief, the work by Uysal and El Roubi (1999) explored and demonstrated that multiple regression could be replaced with Artificial Neural Networks (ANN) in tourism demand studies and concluded by highlighting the need for further applications and evaluations of ANN in the context of tourism demand analysis. This lead to a surge in research focusing on the application of NN models for forecasting tourism demand (see for example, Law and Au (1999) and Law (2000)) with the trend continuing into the present day.

In this paper, we seek to improve the accuracy of forecasts attainable via NN for the highly seasonal and volatile tourism demand series for ten European countries. The novelty of this paper stems from its proposal to develop a hybrid Singular Spectrum Analysis (SSA) and NN model to improve the accuracy of NN forecasts for seasonal time series. Unlike NN, the SSA technique is a popular denoising and signal extraction approach (Sanei and Hassani, 2015) which has been successfully applied in a variety of fields. See for example, Beneki et al. (2012), Ghodsi et al. (2015), and Hassani et al. (2017a), and references therein.

Developing a hybrid model combining SSA's denoising capabilities with the forecasting power of NN can be productive and useful as both techniques are nonparametric and therefore not restricted by the parametric assumptions of normality, stationarity or linearity. To the best of our knowledge, this paper marks the introductory and successful application of a hybrid model combining SSA and NN for tourism demand forecasting. Here, it is noteworthy that our interest lies in improving the accuracy of Neural Network Autoregression (NNAR) forecasts attainable via the automated *nnetar* algorithm made available through the forecast package in *R* (Hyndman and Athanasopoulos, 2018).

The undertaking of this research was motivated by several factors. First, NN continue to be increasingly applied in tourism demand forecasting, see for example Cang (2014), Claveria and Torra (2014), Claveria et al. (2015), Olmedo (2016), Li and Cao (2018). As such, there is strong justification for seeking to improve the accuracy of NN forecasts.

Secondly, we have previously evaluated forecasting tourism demand with a variety of models, including the NNAR model. The findings which are reported in Hassani et al. (2015) and Hassani et al., (2017b) all point towards the sub-optimal and considerably poor performance of the NNAR algorithm at modelling and forecasting tourist arrivals both in the U.S. and Europe, respectively. Therefore, given the ongoing interest in NN within the industry, we are motivated on a personal level to determine whether there is a possibility of improving the accuracy of forecasts attainable via the automated NNAR algorithm.

Thirdly, as Curry (2007) points out, the debate on the capacity of feed-forward NN in dealing with seasonality without the need for pre-processing continues. We find this interesting as the historical belief was that NN could concurrently detect both seasonality and nonlinear trend in time series (Gorr, 1994). However, as Zhang and Qi (2005) asserts, forecasting seasonal data with NN results in mixed conclusions as some authors argue that prior deseasonalization is not necessary because NN can model seasonality directly, whilst others conclude the exact opposite.

For example, the analysis conducted by Farway and Chatfield (1995) showed that NN struggle when directly forecasting seasonal time series. Moreover, Nelson et al. (1999) showed that NN can produce significantly better forecasts if trained on deseasonalized data when compared with forecasts from NN trained on seasonally non-adjusted data. Cho (2003) forecasted Hong Kong's tourist arrivals from different countries using Elman's Model of ANN, ARIMA and Exponential Smoothing and found that ANN was best for forecasting tourism demand when the series was without an obvious pattern. Following simulation studies and applications to real data, Zhang and Qi (2005) found evidence to support detrending or deseasonalizing for reducing NN forecasting errors. In fact, this study showed that NN struggle to model and forecast when presented with raw data which has strong seasonality and trends. More recent evidence by Claveria et al. (2017) provides further support for deseasonalizing series when forecasting with NN. Accordingly, it is clear that in contrast to our personal beliefs, many researchers argue in favour of deseasonalisation or detrending as the more suitable approach for improving the accuracy of NN forecasts.

However, it is noteworthy that there have been some authors who reported that NN can model the trend and seasonal effects in time series without the need for denseasonaling (Hamzacebi, 2008). See for example, Frances and Draisma (1997) and Alon et al. (2001). Tseng et al. (2002) sought to create a hybrid model which combined NN back propagation (BP) with a seasonal ARIMA (SARIMA) model and found this hybrid model could outperform two other NN models and the univariate SARIMA model when faced with seasonal time series. Hamzacebi (2008) proposed an ANN model with the Seasonal ARIMA model's seasonality parameter for direct forecasting of seasonal time series with NN.

Nevertheless, in contrast to previous studies, our interest lies in evaluating whether denoising (as opposed to deseaonalising or detrending) can result in significant improvements in the accuracy of forecasts attainable via the automated feed-forward NNAR model when faced with seasonal time series. Moreover, this paper is also the first instance that such an evaluation is conducted on tourism demand data to determine its applicability and validity within the sector.

In the next section, the proposed hybrid model is presented and discussed, followed by an introduction to the benchmark forecasting models. Thereafter, we use Section 3 to introduce the data, forecasting exercise and the metrics used to evaluate forecast accuracy. The main results from the forecasting exercise are presented in Section 4 whilst a comparison of the proposed hybrid model with benchmark forecasting models are considered in Section 5. The paper concludes in Section 6 with some directions for future research and a discussion of the limitations underlying the current study.

2 Methodology

We propose denoising the seasonal European tourist arrivals data with SSA and reconstructing a less noisy series which captures the seasonal variation and trend in tourist arrivals. Thereafter, we seek to generate out-of-sample NNAR forecasts for European tourist arrivals using the reconstructed series and compare our results with those obtained when forecasting the raw data with the NNAR model.

2.1 Denoised Neural Networks Model (DNNAR)

2.1.1 Singular Spectrum Analysis (SSA)

In brief, the SSA technique comprises of two stages known as Decomposition and Reconstruction, each with its own two steps referred to as Embedding and Singular Value Decomposition (SVD), and Grouping and Diagonal Averaging, respectively. The two SSA choices required to decompose and reconstruct a given series are the Window Length L and the number of eigenvalues r (Sanei and Hassani, 2015). In what follows, we present a detailed description of the SSA process, and in doing so, we mainly follow Hassani et al. (2014).

Stage I: Decomposition

Step I: Embedding
Consider a real-valued nonzero time series $Y_N = (y_1, \ldots, y_N)$ of length N. Embedding transfers a one-dimensional time series $Y_N = (y_1, \ldots, y_N)$ into the multi-dimensional series $X_1, \ldots, X_K$ with vectors $X_i = (y_i, \ldots, y_{i+L-1})^T \in \mathbf{R}^L$, where L $(2 \leq L \leq N/2)$ is the Window Length and $K = N - L + 1$. The result of this step is the following trajectory matrix

$$\mathbf{X} = [X_1, \ldots, X_K] = (x_{ij})_{i,j=1}^{L,K}. \tag{1}$$

Step II: SVD
Denote by $\lambda_1, \ldots, \lambda_L$ the eigenvalues of $\mathbf{X}\mathbf{X}^T$ arranged in decreasing order $(\lambda_1 \geq \ldots \lambda_L \geq 0)$ and by $U_1, \ldots, U_L$ the corresponding eigenvectors. The SVD of $\mathbf{X}$ can be written as $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_L$, where $\mathbf{X}_i = \sqrt{\lambda}_i U_i V_i^T$ and $V_i = \mathbf{X}^T U_i / \sqrt{\lambda}_i$ (if $\lambda_i = 0$ we set $\mathbf{X}_i = 0$).

Stage II: Reconstruction

Step I: Grouping
The grouping step corresponds to splitting the elementary matrices $\mathbf{X}_i$ $(i = 1, \ldots, L)$ into several groups and summing the matrices within each group for reconstruction purposes. Denote $I = \{i_1, \ldots, i_p\}$ as a group of indices $i_1, \ldots, i_p$. Then the matrix $\mathbf{X}_I$ corresponding to the group I can be defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_p}$. The split of the set of indices $J = 1, \ldots, d$ into the disjoint subsets $I_1, \ldots, I_m$ corresponds to the representation

$$\mathbf{X} = \mathbf{X}_{I_1} + \cdots + \mathbf{X}_{I_m}. \tag{2}$$

The procedure of choosing the sets $I_1, \ldots, I_m$ is called the eigentriple grouping. For a given group I the contribution of the component $\mathbf{X}_I$ into the expansion (1) is measured by the share of the corresponding eigenvalues: $\sum_{i \in I} \lambda_i / \sum_{i=1}^{d} \lambda_i$.

Step II: Diagonal Averaging
The purpose of diagonal averaging is to transform a matrix to the form of a Hankel matrix, which can be subsequently converted to a time series. A Hankel matrix is a matrix where all the elements along the diagonal i + j = const are equal. For example, suppose z_{ij} stands for an element of a matrix $\mathbf{Z}$, then the k-th term of the resulting series is obtained by averaging z_{ij} over all i, j such that $i + j = k + 2$. This procedure is also known as Hankelization of the matrix $\mathbf{Z}$. The output of the Hankelization of a matrix $\mathbf{Z}$ is the Hankel matrix $\mathcal{H}\mathbf{Z}$, which is the trajectory matrix corresponding to the series obtained as a result of diagonal averaging. The Hankel matrix $\mathcal{H}\mathbf{Z}$ uniquely defines the series by relating the value in the diagonals to the values in the series. By applying the Hankelization procedure to all matrix components of (2), we obtain another expansion:

$$\mathbf{X} = \widetilde{\mathbf{X}}_{I_1} + \ldots + \widetilde{\mathbf{X}}_{I_m}, \tag{3}$$

where $\widetilde{\mathbf{X}}_{I_1} = \mathcal{H}\mathbf{X}$. This is equivalent to the decomposition of the initial series $Y_N = (y_1, \ldots, y_N)$ into a sum of m series:

$$y_n = \sum_{k=1}^{m} \widetilde{y}_n^{(k)}, \tag{4}$$

where $\widetilde{Y}_N^{(k)} = (\widetilde{y}_1^{(k)}, \ldots, \widetilde{y}_N^{(k)})$ corresponds to the matrix $\mathbf{X}_{I_k}$.

Once a less noisy series is reconstructed, we use this less noisy series to generate NNAR forecasts via the algorithm introduced in Section 2.1.2.

As this paper exploits SSA mainly for its denoising capabilities, we find it pertinent to report the SSA choices used to decompose these series, plot the signal extractions as an example (Figure 1), and comment on the separation of signal and noise as achieved via SSA.

To this end, the weighted correlation (w-correlation) is a statistic which can be used to present the appropriateness of the various decompositions achieved by SSA. Table 1 reports the SSA choices used to decompose all tourist arrival series and also the corresponding w-correlations (calculated for all SSA decompositions by comparing the two components of signal and noise). In Figure 1, as an example, we show the SSA signal extractions for tourism demand in Germany.

As mentioned in Golyandina et al. (2001), the w-correlation is a statistic which shows the dependence between two time series. It can be calculated as:

$$\rho_{12}^{(w)} = \frac{\left(Y_N^{(1)}, Y_N^{(2)}\right)_w}{\| Y_N^{(1)} \|_w \| Y_N^{(2)} \|_w},$$

where $Y_N^{(1)}$ and $Y_N^{(2)}$ are two time series, $\| Y_N^{(i)} \|_w = \sqrt{\left(Y_N^{(i)}, Y_N^{(i)}\right)_w}$, $\left(Y_N^{(i)}, Y_N^{(j)}\right)_w = \sum_{k=1}^{N} w_k y_k^{(i)} y_k^{(j)} \quad (i, j = 1, 2)$, w_k=min$\{k, L, N-k\}$ (here, assume $L \leq N/2$).

The interpretation of the w-correlation suggests that if the value between two reconstructed components are close to 0, then the corresponding time series are w-orthogonal and are well separable (Sanei and Hassani, 2015), and thus confirms the noise is indeed random even though residual randomness is not an explicit concern for nonparametric models. Here, we use as signal the reconstructed series containing r components and select the remaining r (which does not belong to the reconstruction) as noise. As evident, all w-correlations are close to 0 and this confirms that SSA has successfully achieved a sound separation between signal and noise.

Table 1: SSA decompositions and w-correlations between signal and noise for European tourist arrivals.

Country	**SSA**	**w-cor**
Germany	SSA(60,13)	0.003
Greece	SSA(72, 1-7:10-12)	0.008
Spain	SSA(60,14)	0.005
Italy	SSA(60,14)	0.004
Cyprus	SSA(60,1-10:13-15)	0.003
Netherlands	SSA(60,12)	0.008
Austria	SSA(60,10)	0.006
Portugal	SSA(60,15)	0.001
Sweden	SSA(60,13)	0.002
UK	SSA(48,7)	0.004

Note: $SSA(L,r)$ - Window Length L and the Number of Eigenvalues r. w-correlation refers to the weighted correlation between the signal and noise components, as determined via SSA. For Greece and Cyprus the separation between signal and noise does not follow a binary approach, but instead follows the Colonial Theory based decomposition approach in Hassani et al. (2016). For example, this means that for Greece, eigenvalues 1-7 and 10-12 are grouped together as signal whilst eigenvalues 8-9 and 13-72 are grouped as noisy components. This is because the analysis of eigenfunctions and principal components indicated that a binary decomposition would result in grouping noisy components within the signal and vice versa.

The signal extractions for Germany in Figure 1 gives the reader an indication on the complexities underlying European tourism demand series. The varying amplitudes within the seasonal fluctuations imply that forecasting models with filtering and signal extraction capabilities could benefit.

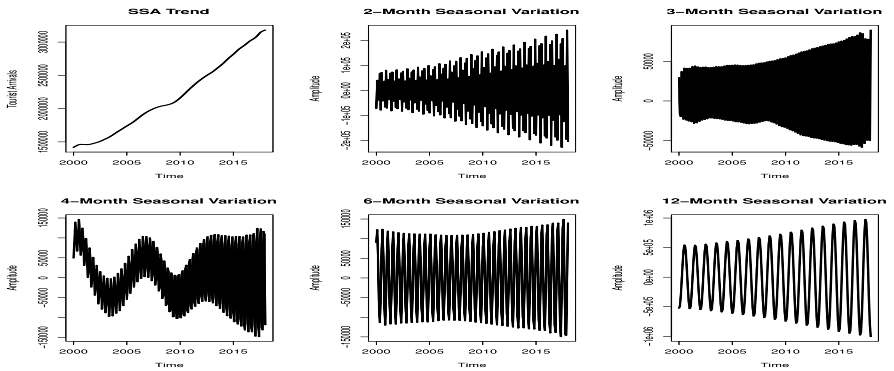

Figure 1: SSA decompositions for tourist arrivals in Germany.

2.1.2 Neural Network Autoregression (NNAR)

In this paper, in the first instance we consider lagged values of raw European tourist arrivals as inputs to a NNAR model and in the second instance, consider lagged values of reconstructed European tourist arrivals as inputs. The model being developed is known as $NNAR(p, P, k)_m$ model (where p refers to lagged inputs, P takes a default value of 1 for seasonal data, k refers to nodes in the hidden layer, and m refers to a monthly frequency) as made available via the

Forecast Package in *R*. Those interested in a detailed description of the NNAR algorithm are referred to Hyndman and Athanasopoulos (2018). Below, we follow Hyndman and Athanasopoulos (2018) to present some brief and relevant details pertaining to the NNAR models used in this paper.

The *nnetar* algorithm automates and speeds up the process for obtaining NNAR forecasts via a feed-forward NN model with one hidden layer. A seasonal NNAR model has the notation NNAR$(p, P, k)_m$, and inputs $(y_{t-1}, y_{t-2}, \cdots, y_{t-p}, y_{t-m}, y_{t-2m}, y_{t-Pm})$ with k neurons in the hidden layer. The algorithm automatically selects the values of p and P when they are left unspecified. For seasonal data, the default value for P is 1 and p is chosen from the optimal linear model fitted to the seasonally adjusted data. In contrast, for non-seasonal series, the default is set to be the optimal number of lags (based on the Akaike Information Criterion) for a linear $AR(p)$ model. At the same time, if k is not specified, it is set to $k = (p + P + 1)/2$ (rounded to the nearest integer).

Then, the forecasting approach is recursive such that within each horizon, the NNAR model re-estimates parameters by re-modelling the historical data with each new observation that is introduced. This process continues until all required forecasts are computed.

2.2 Benchmark Forecasting Models

We compare the results from the DNNAR model with forecasts from two popular univarite parametric and nonparametric models which are introduced below.

2.2.1 Autoregressive Integrated Moving Average (ARIMA)

ARIMA has been recognized as one of the most established time series forecasting models evidenced by its prediction performance and overwhelming implementations in a broad range of subjects. It aims to separate the signal and noise by adopting the past observations and taking into considerations of the degree of differencing, autoregressive and moving average components. Thus, the ARIMA model can be specified by three parameters (p, d, q), where p donates the number of lags, d indicates the degree of differencing and q refers to the number of error terms to be included. To present the detailed formulations of ARIMA model, we mainly follow Hyndman and Khandakar (2008). The non-seasonal ARIMA(p,d,q) model can then be written as:

$$(1 - \sigma_1 B - ... - \sigma_p B^p)(1 - B)^d y_t = c + (1 + \sigma_1 B + ... + \sigma_q B^q)e_t, \quad (5)$$

or

$$(1 - \sigma_1 B - ... - \sigma_p B^p)(1 - B)^d (y_t - \mu t^d/d!) = c + (1 + \sigma_1 B + ... + \sigma_q B^q)e_t, \quad (6)$$

where μ is the mean of $(1-B)^d y_t$, $c = \mu(1-\sigma_1 - ... - \sigma_p)$ and B is the back shift operator. In what follows, we also provide a brief expansion of the seasonal ARIMA model by following Hyndman and Khandakar (2008). The seasonal ARIMA model can be expressed as:

$$\Phi(B^m)\phi(B)(1 - B^m)^D (1 - B)^d y_t = c + \Theta(B^m)\theta(B)\epsilon_t, \quad (7)$$

where $\Phi(z)$ and $\Theta(z)$ are the polynomials of orders P and Q, and ϵ_t is white noise. If, $c \neq 0$, there is an implied polynomial of order $d + D$ in the forecast function.

We use an automated ARIMA model which is commonly known as 'auto.arima' and is accessible via the Forecast Package in *R*. It is of note that a more de-tailed description of this optimized algorithm and examples of application can be found in Hyndman and Khandakar (2008). In brief, the algorithm begins by repeating KPSS tests (Kwiatkowski et al., 1992) to determine the number of differences d. Then, the data is differenced d times to minimise the Akaike Information Criterion (AIC) and determine the values of p (number of au-toregressive terms) and q (number of lagged forecast errors in the forecasting equation). The efficiency of this algorithm is noteworthy as instead of consid-ering every possible combination of p and q, it opts to traverse the model space via a stepwise search. Also, 'auto.arima' relies on a corrected version of the AIC (AIC$_c$) as indicated below:

$$AIC = -2log(L) + 2(p + q + P + Q + k), \tag{8}$$

$$AIC_c = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2}. \tag{9}$$

where $k = 1$ if c $\neq 0$ and 0 otherwise, L is the maximum likelihood of the data and the last term in parentheses is the number of parameters in the model (this includes σ^2 which is the variance of the residuals).

Thereafter, the algorithm seeks to determine the 'current model' by search-ing for the following four ARIMA models: ARIMA*(2,d,2)*, ARIMA*(0,d,0)*, ARIMA*(1,d,0)* and ARIMA*(0,d,1)* for the one which minimises the AIC$_c$. If $d =$ 0 then the constant c is included; if $d \geq 1$, then the constant c is set to zero. In addition, the model also evaluates variations on the current model by varying p and q by +/-1 and including/excluding c. The steps following on from the minimisation of the AIC are repeated until no lower AIC$_c$ can be found.

2.2.2 Exponential Smoothing (ETS)

ETS was initially introduced in the 1950s and has been playing an important role in many established forecasting methods. The predictions are weighted averages based on the existing observations where the higher weight is assigned to the more recent record with overall exponentially decaying weights. To present the ETS model with more details, we mainly follow Hyndman and Athanasopoulos (2018).

There are many ETS methods differentiated by the approaches of handling smoothing parameters and different time series components. Consider a time series $Y_T = (y_1, y_2, ..., y_T)$ of length T, $\hat{y}_T$ indicates the forecast of y_T given y_1 to y_{T-1}. We firstly start from the simple ETS for a one step ahead forecast of y_{T+1} given Y_T:

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + ..., \tag{10}$$

where $0 \leqslant \alpha \leqslant$ is the smoothing parameter and it controls the exponentially weight distribution as observations are further away in the past.

Donate h as the number of steps ahead, l_t is the smoothed value or level of the series at time t, b_t refers to the trend component, and α, β, γ and δ are representing different smoothing parameters. The trend ETS method by Holt can then be summarized as:

$$\hat{y}_{t+h|t} = l_t + hb_t, \tag{11}$$

$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}), \tag{12}$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}, \tag{13}$$

where b_t indicates an estimate of the slop of the series at time t, α and β are the smoothing parameters for the level and trend, respectively.

Furthermore, the Holt-Winters seasonal method is developed to capture seasonality as well. Donate s_t as the seasonal component, m indicates the number of seasons in a year, we get

$$s_t = \gamma(1-\alpha)(y_t - l_{t-1} - b_{t-1}) + [1-\gamma(1-\alpha)]s_{t-m}, \tag{14}$$

so this multiplicative method and its component form are:

$$\hat{y}_{t+h|t} = (l_t + hb_t)s_{t+h-m(k+1)}, \tag{15}$$

$$l_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(l_{t-1} + b_{t-1}), \tag{16}$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}, \tag{17}$$

$$s_t = \gamma \frac{y_t}{(l_{t-1} + b_{t-1}} + (1-\gamma)s_{t-m}, \tag{18}$$

The ETS methods are further developed to more variations that consider different combinations of the time series components and its diversified refinements. A more detailed description of over 30 different ETS methods and examples of applications are available in Hyndman and Athanasopoulos (2018), which we do not reproduce here.

The Forecast Package in *R* also automates the nonparametric ETS model. In brief, the ETS model is automated to consider the error, trend and seasonal components in choosing the best exponential smoothing model from 30 possible options. This is achieved via the optimization of initial values and parameters using the Maximum Likelihood Estimator and selecting the best model based on the AIC. The algorithms that generate point forecasts for European tourist arrivals[1] and state space equations for each of the models in the ETS framework[2] can be found in Hyndman and Athanasopoulos (2018).

3 Data

The data set used in this study relates to international tourism demand (January 2000 - January 2018) for ten European countries and was obtained via the

[1] https://www.otexts.org/sites/default/files/fpp/images/Table7-8.png
[2] https://www.otexts.org/sites/default/files/fpp/images/Table7-10.png

Eurostat database. This same data set (up until December 2013) was recently used in Hassani et al. (2017) and Silva et al. (2017). However, prior to summarising the data using descriptive statistics, we find it pertinent to further justify the choice of countries considered in this paper.

In Table 2, we highlight some key statistics relating tourism to economic growth in the selected countries. It is evident that tourism is an important contributor to the economic activities of the selected countries and the World Travel & Tourism Council's Travel & Tourism Economic Impact 2018 reports forecast all of these contributions to rise not only in 2018 but also considerably by 2022. Also Seasonality is the dominant feature in all these time series data. In fact Silva (2017) reported that seasonality accounts for more than 90% of the variations in most of these series.

Table 2: Total contribution of travel and tourism to GDP and Employment in 2017.

	GDP	Employment
Germany	10.7%	13.8%
Greece	19.7%	24.8%
Spain	14.9%	15.1%
Italy	13.0%	14.7%
Cyprus	22.3%	22.7%
Netherlands	5.2%	8.9%
Austria	14.8%	16.1%
Portugal	17.3%	20.4%
Sweden	9.5%	11.1%
United Kingdom	10.5%	11.6%

Note: The data have been compiled via various Travel & Tourism Economic Impact 2018 reports published by the World Travel & Tourism Council (https://www.wttc.org/). Percentages reported under GDP should be interpreted in relation to total GDP in the respective country. Percentages reported under employment should be interpreted in relation to total employment in the respective country.

Table 3 reports some descriptive statistics for the data. The Shapiro-Wilk (SW) test for normality indicates that none of the tourist arrivals data are normally distributed and therefore, those interested should consider the Median as a measure of central tendency and the Interquartile Range (IQR) as a measure of variation in this data. We also test the series for seasonal unit roots using the OCSB (Osborn et al., 1988) test for seasonal unit roots. This indicates that all series have seasonal unit roots. In comparison to the tourist arrivals for these same countries which are summarised up until December 2013 in Hassani et al. (2017b), we notice that the structure of the data has changed significantly for two countries. Tourist arrivals in the Netherlands did not have a seasonal unit root up until December 2013 (Hassani et al., 2017b), however the changes in tourism demand post 2013 have created a seasonal unit root. On the other hand, tourist arrivals in Austria which were normally distributed up until 2013 (Hassani et al., 2017b) are now skewed according to the latest tourism demand data. Furthermore, a comparison of Median tourism demand values in Hassani et al. (2017b) against the Median values in Table 3 further justify the selection of these countries in this study as all Median tourist arrivals have increased over time.

Table 3: Descriptive statistics for European tourist arrivals (Jan. 2000 - Dec. 2013).

	Mean	Med.	IQR	SD	SW (p)	OCSB
Germany	2169425	2084257	1044178	774706	<0.01	1
Greece	874189	591676	1170788	845796	<0.01	1
Spain	3763340	3654786	2454851	1622601	<0.01	1
Italy	3685657	3603542	3025709	1893910	<0.01	1
Cyprus	162592	190296	162302	91185	<0.01	1
Netherlands	970665	946600	438000	322085	<0.01	1
Austria	1616651	1588805	742372	550774	<0.01	1
Portugal	629156	588378	431901	333551	<0.01	1
Sweden	413309	272223	297635	316492	<0.01	1
United Kingdom	2065921	1836995	1292000	1035709	<0.01	1

Note: 1 indicates there is a seasonal unit root based on the OCSB test at p=0.05.

3.1 Forecasting Exercise

For the forecasting exercise, the data set is separated into an initial in-sample period traversing January 2000 - January 2012 (2/3rd of entire data set) and an out-of-sample period of February 2012 - January 2018. This separation of the period was in line with previous tourist arrivals forecasting studies, see for example Hassani et al. (2015;2017b), Silva et al. (2017) and references therein. Note that the in-sample period expands when producing out-of-sample forecasts for various horizons as we consider a recursive estimation for the NNAR, DNNAR, ARIMA and ETS models. Thus, with each new observation being introduced, the NNAR model updates the data and re-estimates its model parameters. The number of out-of-sample forecasts at each horizon can be calculated using the formula $n - h + 1$ observations where $n = 72$ and h is the forecasting horizon.

3.2 Forecast Accuracy Evaluation Metrics

Root Mean Squared Error (RMSE)

In this paper, RMSE and Ratio of the RMSE (RRMSE) criteria are used as the main metrics for evaluating forecast accuracy. These measures have been successfully adopted as forecast evaluation metrics in many recent papers, see for example Hassani et al. (2018; 2017) and Silva et al. (2018; 2017) and references therein. An example of a RRMSE calculation used in this paper can be defined as:

$$\text{RRMSE} = \frac{DNNAR}{NNAR} = \frac{\left(\sum_{i=1}^{N}(\widehat{y}_{T+h,i} - y_{T+h,i})^2\right)^{1/2}}{\left(\sum_{i=1}^{N}(\widetilde{y}_{T+h,i} - y_{T+h,i})^2\right)^{1/2}},$$

where, $\widehat{y}_{T+h}$ - h-step ahead forecast from the DNNAR model, $\widetilde{y}_{T+h}$ - h-step ahead forecast from the NNAR model, and N - the number of forecasts. If $\frac{DNNAR}{NNAR}$ is less than 1, then the DNNAR forecast outperforms the NNAR forecast by 1-$\frac{DNNAR}{NNAR}$ percent and vice versa.

Mean Absolute Percentage Error (MAPE)

We also consider the MAPE criterion as it is a widely understood measure for evaluating forecast accuracy (Hassani et al., 2015). In brief, the lower the MAPE result, the better the forecast.

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N} |100 \times \frac{y_{T+h} - \widehat{y}_{T+h,i}}{y_{T+h}}|,$$

where y_{T+h} represents the actual data corresponding to the h step-ahead forecast, and $\widehat{y}_{T+h,i}$ is the h step-ahead forecast obtained from a particular forecasting model.

It is noteworthy that the differentiations based on predictive accuracy are further analysed through testing the forecast errors via the modified Diebold-Mariano (DM) test (Harvey et al., 1997) and the Hassani-Silva (HS) test for predictive accuracy (Hassani and Silva, 2015). In each case, we consider the series appearing on the denominator of the RRMSE criterion as the benchmark model.

4 Empirical Results

The findings from the out-of-sample forecasting exercise are reported below via Table 4. We begin our analysis by taking a look at the results at a macro level. Accordingly, based on the RMSE criterion, we can conclude that forecasts from the DNNAR models outperform NNAR forecasts for European tourist arrivals in all cases. The MAPE criterion too supports this conclusion as the DNNAR forecast MAPE values are smaller than those for the NNAR forecasts for all countries across all horizons.

Next, we consider a micro level analysis of the forecasting results by taking into account the performance by country. Here, we test the out-of-sample forecasts for statistically significant differences using both the modified DM test and the HS test. For each country, we also plot the forecasts for the horizon which corresponds to the largest accuracy gain (based on the RRMSE) for the DNNAR model.

In terms of forecasting German tourist arrivals, we find that forecasts from the DNNAR model outperforms NNAR forecasts across all horizons with statistically significant outcomes (based on both the RMSE and MAPE criteria). The RRMSE criterion enables us to conclude that DNNAR forecasts are 45%, 40%, 43% and 43% better than the NNAR forecasts for German tourist arrivals at $h = 1, 3, 6$ and 12 steps-ahead, respectively. In what follows, for each country, we plot the out-of-sample forecasts for all horizons. However, as these series are highly seasonal and show a similar structure, we only explain some interesting elements from the best performing DNNAR forecast. Accordingly, Figure 2 shows the out-of-sample forecasts for German tourist arrivals. Here, we analyse the forecasts at $h = 1$ step-ahead. A close inspection makes it clear that the NNAR forecast based on the raw data over estimates the tourist arrival peaks between 2016 and 2017 whilst the DNNAR forecast which is based on the denoised tourist arrival series is visibly more accurate.

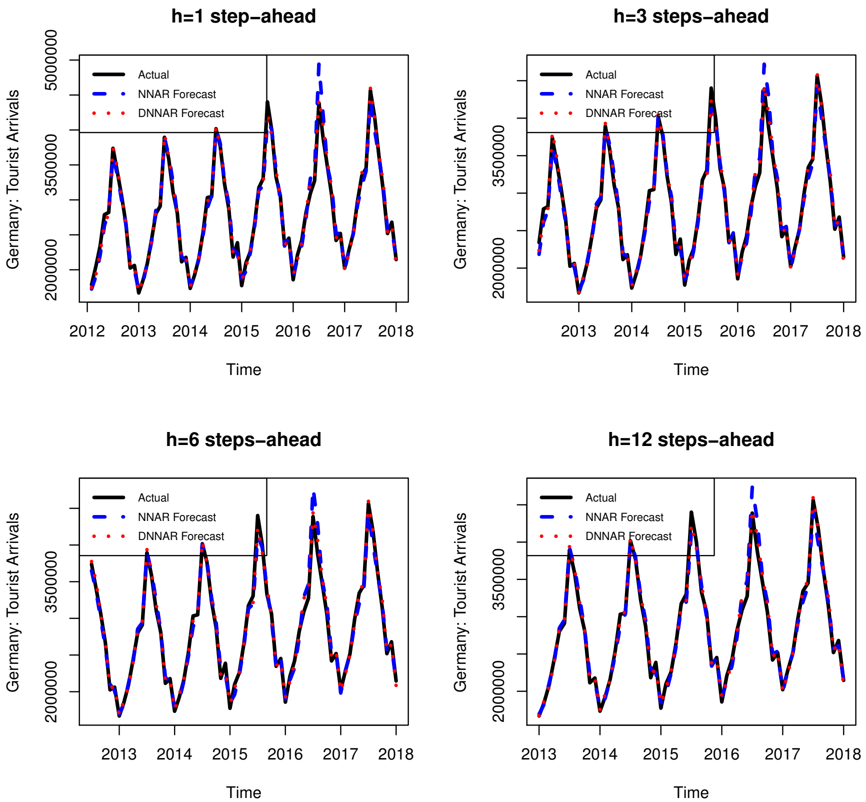

Figure 2: Out-of-sample forecasts for tourist arrivals in Germany.

In terms of forecasting tourism demand for Greece, we find the DNNAR forecasts are statistically significantly better than the NNAR forecasts across all horizons based on both the RMSE and MAPE criteria. The RRMSE criterion shows that DNNAR forecasts are 31%, 43%, 32%, and 52% better than forecasts from the NNAR model at $h = 1, 3, 6$ and 12 steps-ahead, respectively. Once again, we interpret the plot the out-of-sample forecast representing the largest accuracy gain for tourism demand in Greece (Figure 3). In this case, the DNNAR model's capability at providing better forecasts for 4/5 peaks captured in the out-of-sample period is clearly visible. However, it is notable that the NNAR model provides a comparatively better estimate of a single peak which occurs in 2016.

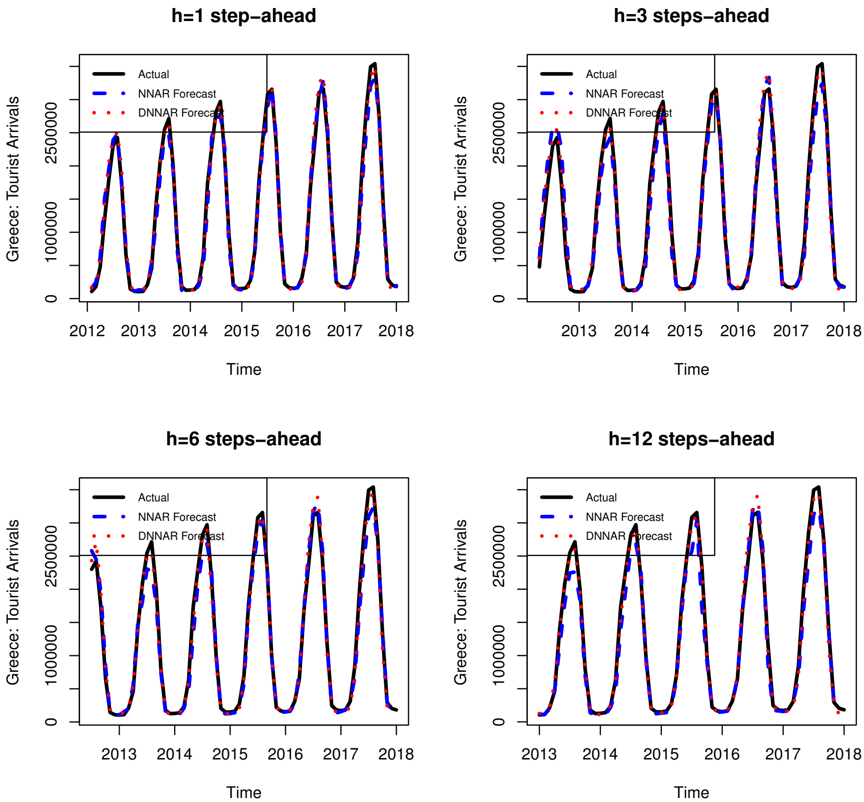

Figure 3: Out-of-sample forecasts for tourist arrivals in Greece.

When it comes to forecasting tourist arrivals in Spain, the RMSE and MAPE criteria show that the DNNAR model once again generates forecasts which are statistically significantly better in relation to those from the NNAR model. Whilst the gains here are comparatively lower, the RRMSE criterion shows that the DNNAR forecasts are 17%, 18%, 22%, and 21% better than forecasts from NNAR at $h = 1, 3, 6$ and 12 steps-ahead, respectively. As the RRMSE criterion suggests the DNNAR model reports the largest accuracy gain at $h = 6$ steps-ahead, and we interpret this plot (Figure 4). Interestingly, when forecasting tourism demand for Spain, we notice that the NNAR model struggles not only with peaks, but also with predicting the troughs. We can see that the NNAR model over estimates the peaks in tourism demand in 2012 and 2017 (the DNNAR model too struggles with estimating this particular peak) and also the troughs from 2014-2017. The DNNAR model is comparatively better at estimating the troughs.

For the case of Italian tourist arrivals, the DNNAR model produces forecasts which are significantly better than those from the NNAR model across all horizons (based on the RMSE and MAPE criteria). The RRMSE indicates that the DNNAR forecasts are 36%, 35%, 31%, and 33% better than forecasts from the NNAR models at $h = 1, 3, 6$ and 12 steps-ahead. For Italy, we interpret the plot for the out-of-sample forecasts at $h = 1$ step-ahead in Figure 5. In this case, we notice both DNNAR and NNAR models struggling to estimate the peak in tourism demand recorded in 2014. However, the DNNAR model successfully picks up the peaks in demand in 2012 and 2015-2017 where the NNAR model overestimates demand.

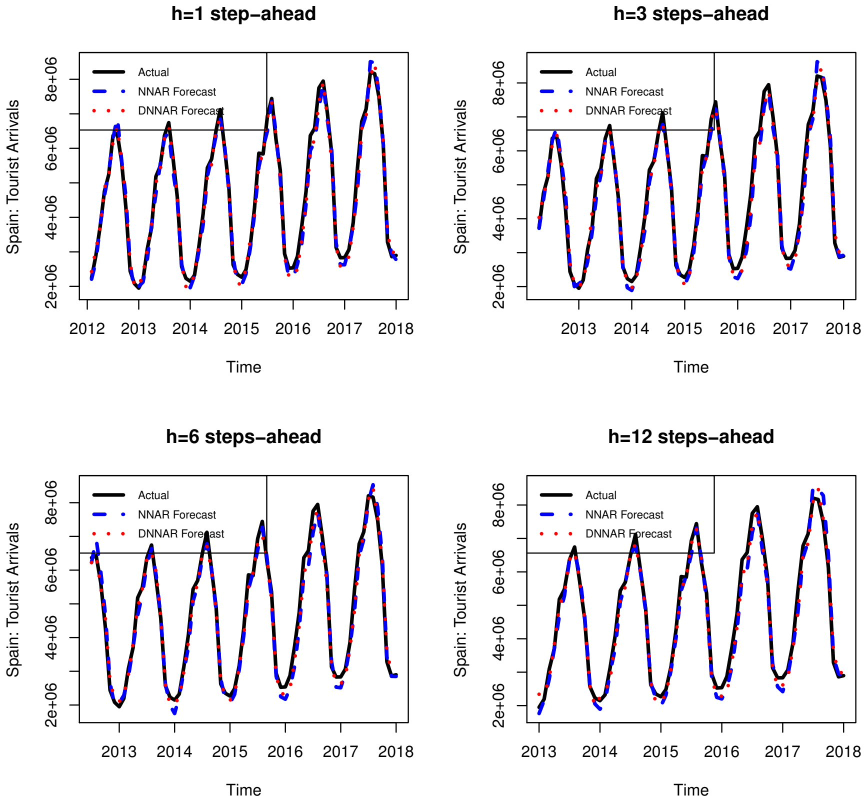

Figure 4: Out-of-sample forecasts for tourist arrivals in Spain.

In terms of the forecasting tourist arrivals in Cyprus, we find the DNNAR models outperforming forecasts from the NNAR models across all horizons based on the RMSE and MAPE criteria. However, we only find evidence of statistically significant differences between forecasts at $h = 6$ steps-ahead. Nevertheless, the RRMSE criterion shows that forecasts from the DNNAR models are 14% better than those from the NNAR models at $h = 1$ and 3 steps-ahead, and 33% and 27% better than forecasts from the NNAR models at $h = 6$ and 12 steps-ahead, respectively. In this case, the forecast which is significantly better is also the one which reports the largest accuracy gain as per the RRMSE, and this plotted is interpreted from Figure 6. Here, we see that the NNAR model struggles to accurately estimate all 6 peaks in tourism demand between 2012 and 2018. In fact, the NNAR model performance is extremely poor at predicting the initial peak in 2012 and the final two peaks within the out-of-sample period. In addition, the NNAR model also struggles at predicting the troughs in 2013 (as is the case with the DNNAR forecast for this point) and 2017. In contrast, the DNNAR model picks up the peaks in 2013 - 2015 more accurately and provides comparatively better forecasts for the final two peaks in 2016 and 2017, though these forecasts are not extremely accurate.

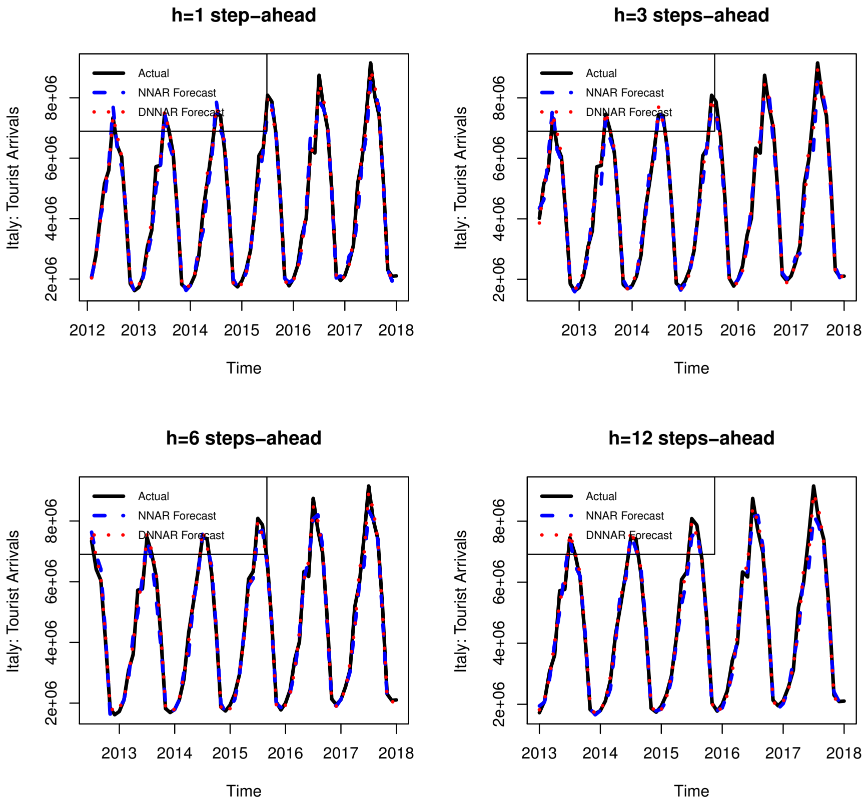

Figure 5: Out-of-sample forecasts for tourist arrivals in Italy.

For tourist arrivals in the Netherlands, the DNNAR models report out-of-sample forecasts which are significantly better than those from the NNAR models across all horizons (as per the RMSE and MAPE criteria). Here, the RRMSE criterion shows that DNNAR forecasts are 37%, 36%, 38%, and 39% better than forecasts from NN at $h = 1, 3, 6$ and 12 steps-ahead. In this case, we analyse tourism demand forecasts for the Netherlands at $h = 12$ steps-ahead. As seen in Figure 7, the NNAR forecast clearly struggles when predicting the seasonal fluctuations for tourism demand in the Netherlands, whereas the DNNAR forecast is comparatively much better at predicting these trends.

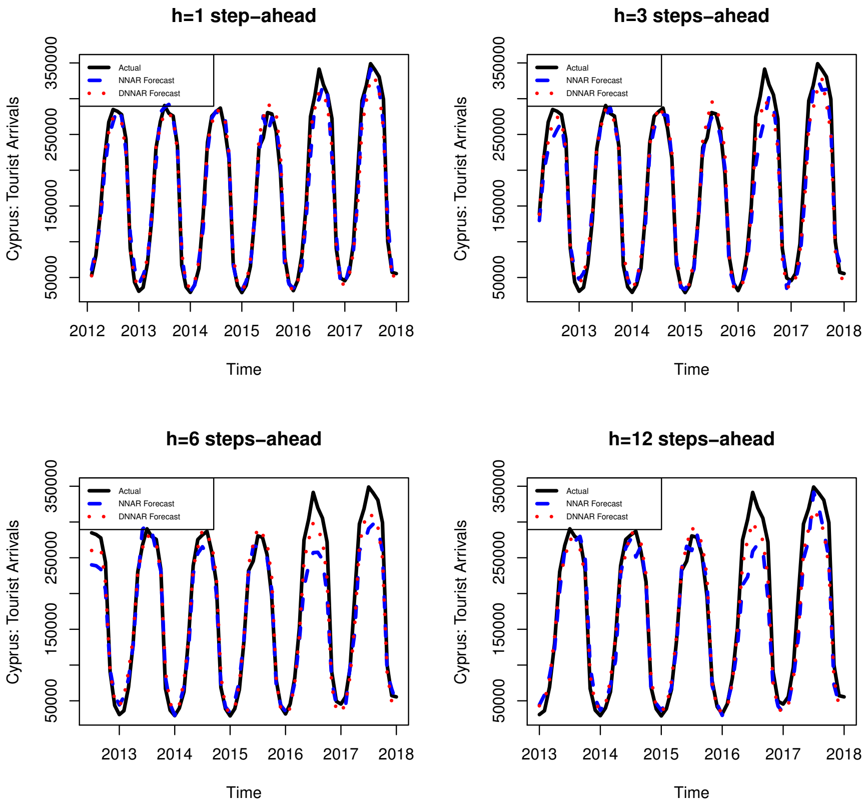

Figure 6: Out-of-sample forecasts for tourist arrivals in Cyprus.

In terms of forecasting tourism demand in Austria, we find that forecasts from the DNNAR model once again outperforms NNAR forecasts with statistically significant outcomes in both the short and long run horizons (as evident via the RMSE and MAPE criteria). Also, the DNNAR forecasts are found to be 31%, 33%, 28%, and 27% better than forecasts from NNAR models at $h = 1, 3, 6$ and 12 steps-ahead, respectively. For Austria, the RRMSE suggests that the largest forecast accuracy gain for the DNNAR model is at $h = 3$ steps-ahead, and this is plotted as part of Figure 8. In this case, we notice how the DNNAR model is once again superior at estimating the peaks in tourism demand comparatively better than the NNAR model right across the out-of-sample period with the exception of the smaller peak in 2013 where both models appear to struggle.

For Portuguese tourist arrivals, we find the DNNAR model outperforming forecasts from the NNAR model across all horizons based on the RMSE and MAPE criteria, but at $h = 1$ step-ahead the reported relative accuracy gain of 8% is not found to be statistically significant. Nevertheless, the DNNAR model is seen generating forecasts which are 31%, 48%, and 43% better than forecasts from the NNAR models at $h = 3, 6$ and 12 steps-ahead with statistically significant differences.

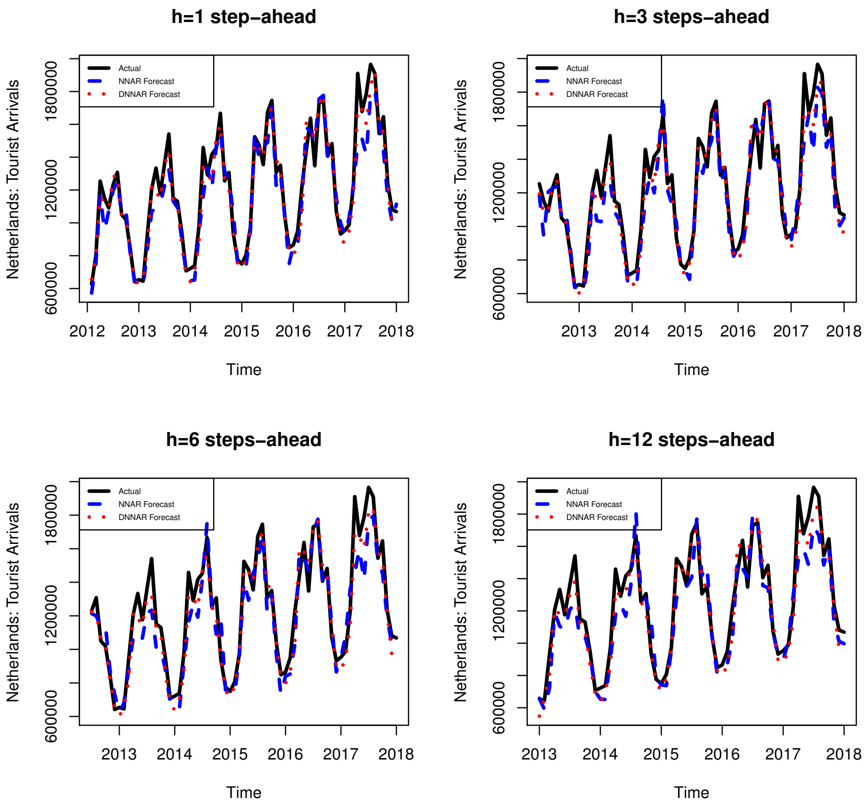

Figure 7: Out-of-sample forecasts for tourist arrivals in Netherlands.

The $h = 12$ steps-ahead plot in Figure 9 shows the DNNAR out-of-sample forecasts at this horizon are once again superior. However, we can also see that when the forecasting horizon increases to $h = 12$ steps-ahead, the NNAR model appears to find it extremely difficult to accurately predict the final peak in tourism demand for Portugal in 2017. In comparison, the DNNAR model is able to provide a better forecast though it still appears to consistently over or underestimate the peaks in tourism demand.

In terms of forecasting Swedish tourist arrivals, the findings from the RMSE and MAPE criteria show the DNNAR model generating statistically significantly better forecasts than the NNAR model across all horizons, with the DNNAR forecasts being 54%, 55%, 73%, and 61% better than forecasts from the NNAR model at $h = 1, 3, 6$ and 12 steps-ahead, respectively. Here, we interpret the out-of-sample forecasts at $h = 6$ steps-ahead (Figure 10). Forecasts from the DNNAR model are seen overestimating the peak in 2013, however, it is comparatively more accurate at forecasting the remaining peaks in relation to the NNAR model.

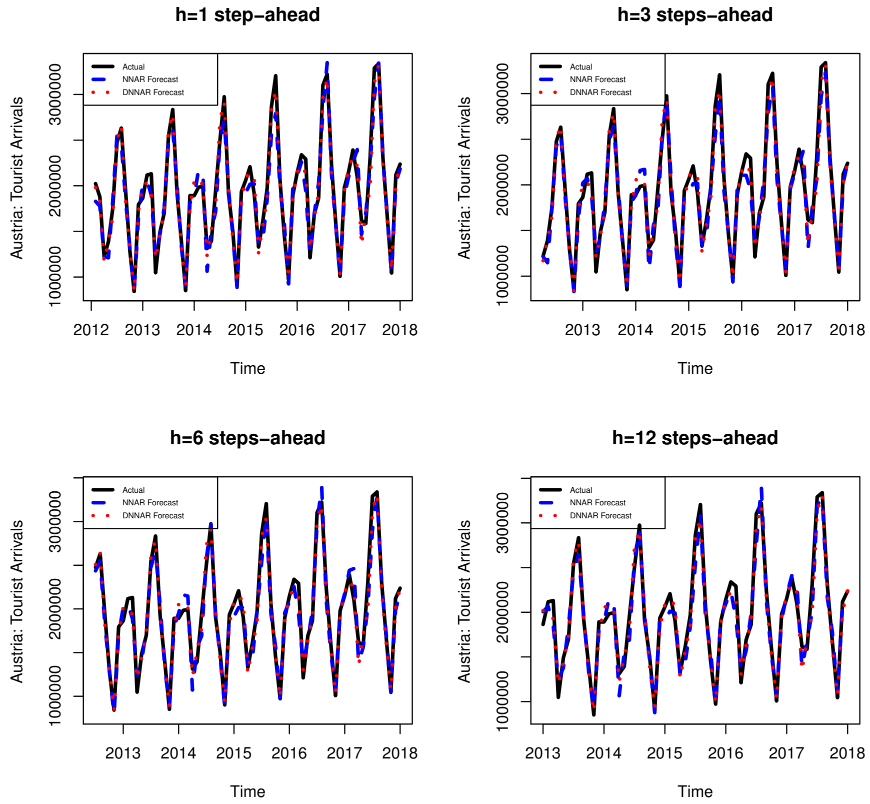

Figure 8: Out-of-sample forecasts for tourist arrivals in Austria.

Finally, in terms of forecasting UK tourist arrivals, the DNNAR model outperforms NN forecasts across all horizons based on the RMSE and MAPE criteria. However, none of the differences between forecasting results are found to be statistically significant. Nevertheless, based on the RRMSE criterion, the DNNAR forecasts are 45%, 44%, 39%, and 28% better than the NNAR forecasts at $h = 1, 2, 3$ and 12 steps-ahead respectively. As with the previous countries analysed here, we interpret the plot for the out-of-sample forecasts for UK tourism demand at $h = 6$ steps-ahead (Figure 11). In this case, its very easy to distinguish between the DNNAR and NNAR forecast accuracy. In fact, the NNAR model is seen predicting a major peak in tourism demand when the UK market is experiencing a trough. The DNNAR model too fails to predict the 2014 peak in UK tourist arrivals but is quick to adjust and pick up the trough in 2015.

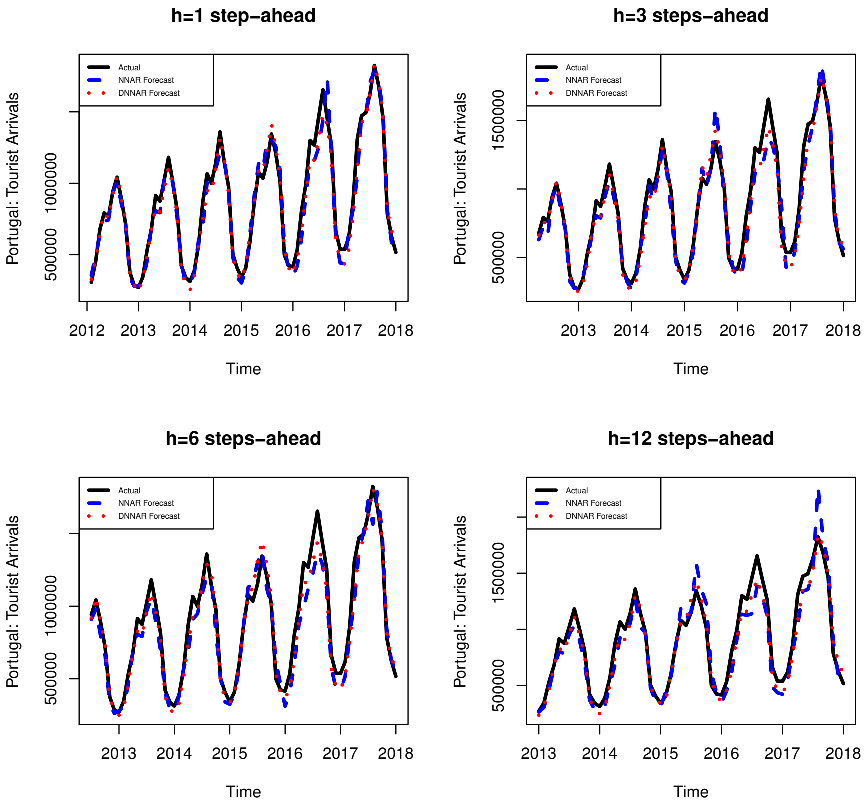

Figure 9: Out-of-sample forecasts for tourist arrivals in Portugal.

Overall, these results uncover several key points. First and foremost, that denoising seasonal tourism demand series with SSA and then fitting a NNAR model on the reconstructed series can enable the generation of superior NNAR forecasts than when a NNAR model is fitted on raw seasonal data for forecasting purposes. Secondly, we find evidence which supports the historical view of Gorr (1994) which suggests that NN models can detect both seasonality and nonlinear trend in time series, in contrast to the findings of Farway and Chatfield (1995), Nelson et al. (1999), Zhang and Qi (2005), and Curry (2007).

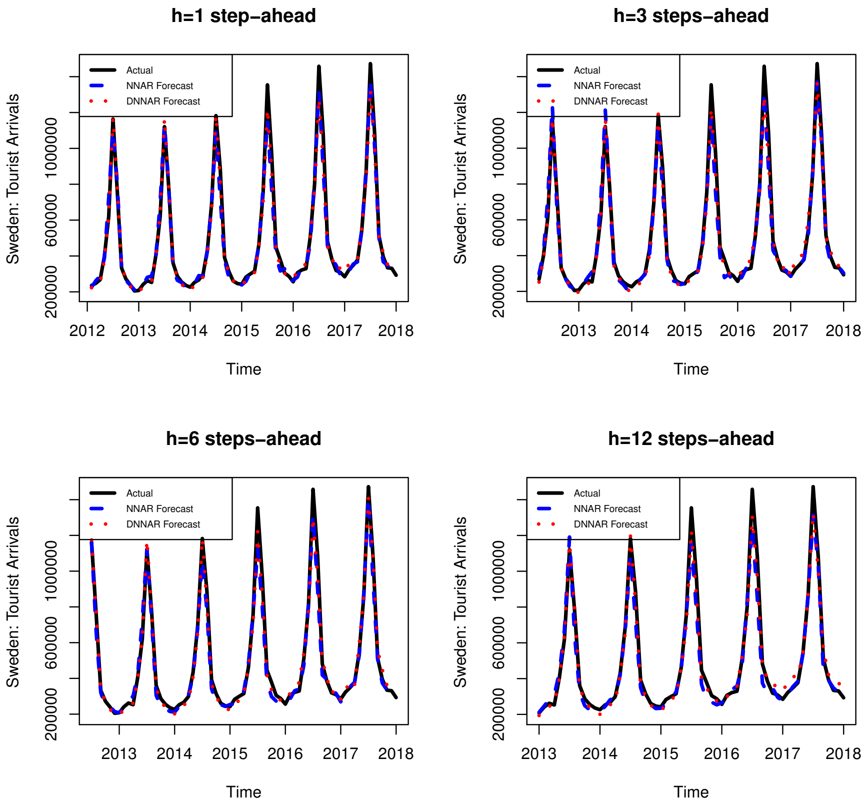

Figure 10: Out-of-sample forecasts for tourist arrivals in Sweden.

Whilst these findings are significant and very promising for those interested in NN models, forecasters and researchers alike would be equally interested in finding out how the newly proposed DNNAR model performs in comparison to other benchmarks. As such, we use the discussion which follows to address this issue by comparing the performance of the DNNAR model with popular time series analysis and forecasting benchmarks.

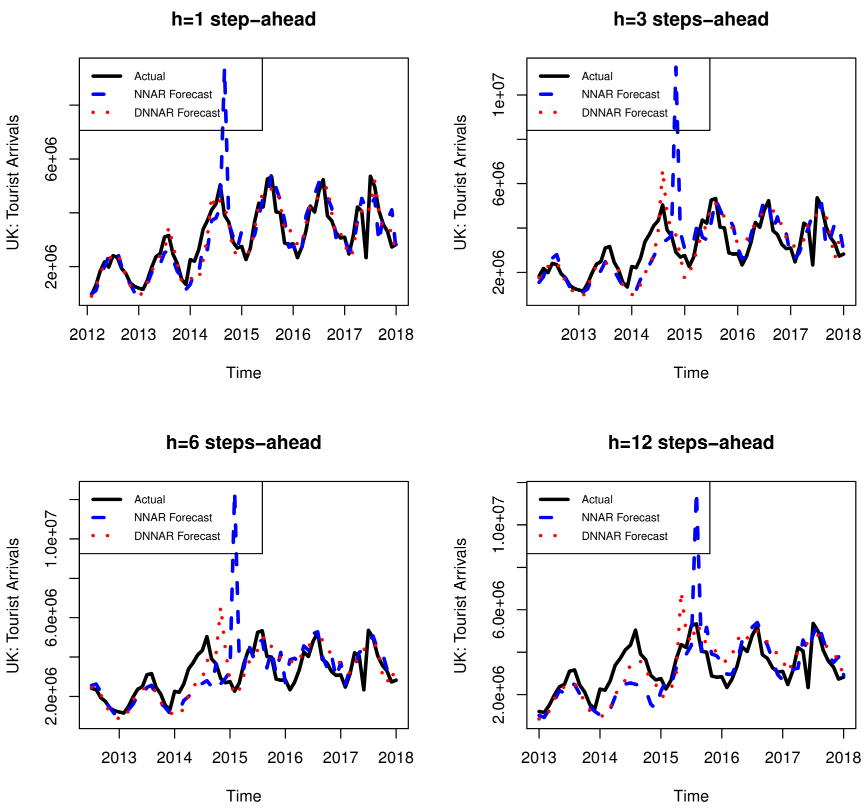

Figure 11: Out-of-sample forecasts for tourist arrivals in UK.

Table 4: Forecasting results for European tourist arrivals from NNAR and DNNAR models.

		RMSE(MAPE)	RMSE(MAPE)	RRMSE
Country	h	**NNAR**	**DNNAR**	$\frac{DNNAR}{NNAR}$
Germany	1	114225(2.64%)	62542(1.72%)	$0.55^{**,\ddagger}$
	3	108092(2.64%)	64343(1.74%)	0.60^{*}
	6	108944(2.59%)	61693(1.63%)	$0.57^{*,\ddagger}$
	12	112084(2.61%)	64081(1.74%)	0.57^{**}
Greece	1	128576(11.74%)	101406(11.12%)	0.79^{*}
	3	187311(12.74%)	106148(11.17%)	$0.57^{**,\ddagger}$
	6	163306(10.99%)	111385(10.58%)	0.68^{*}
	12	194482(11.15%)	93876(9.62%)	$0.48^{***,\ddagger}$
Spain	1	245826(4.68%)	203676(4.20%)	0.83^{***}
	3	281799(5.40%)	230396(4.50%)	0.82^{***}
	6	288385(5.70%)	224654(4.53%)	$0.78^{***,\flat}$
	12	301279(5.77%)	239078(5.06%)	0.79^{***}
Italy	1	330012(5.09%)	211462(3.53%)	0.64^{***}
	3	340517(5.25%)	220984(3.81%)	0.65^{***}
	6	324428(4.92%)	223237(3.80%)	0.69^{***}
	12	341156(5.31%)	227447(3.96%)	0.67^{***}
Cyprus	1	17093(9.57%)	14776(7.76%)	0.86
	3	21201(11.55%)	18293(10.33%)	0.86
	6	29330(12.76%)	19437(11.59%)	0.67^{*}
	12	29143(15.45%)	21340(12.57%)	0.73
Netherlands	1	119099(6.71%)	75385(4.57%)	$0.63^{***,\ddagger}$
	3	125174(7.23%)	80267(4.91%)	$0.64^{***,\ddagger}$
	6	138548(8.24%)	85540(5.21%)	$0.62^{***,\ddagger}$
	12	140701(7.97%)	85505(5.43%)	$0.61^{***,\flat}$
Austria	1	146758(5.70%)	101585(4.13%)	$0.69^{***,\flat}$
	3	154106(6.27%)	102824(4.21%)	$0.67^{***,\ddagger}$
	6	145152(5.83%)	104908(4.26%)	0.72^{***}
	12	149190(6.42%)	108879(4.52%)	0.73^{**}
Portugal	1	89075(8.15%)	68563(6.73%)	0.77^{**}
	3	101262(9.06%)	81368(7.88%)	0.80^{**}
	6	116003(10.26%)	84509(7.99%)	0.73^{**}
	12	135755(12.04%)	90762(9.38%)	$0.67^{**,\ddagger}$
Sweden	1	56347(5.95%)	43555(4.88%)	0.77^{*}
	3	59288(6.46%)	43863(5.91%)	0.74^{*}
	6	59809(6.77%)	40325(5.60%)	0.67^{***}
	12	72098(8.12%)	47703(7.16%)	0.66^{***}
UK	1	839403(15.31%)	460113(11.30%)	0.55
	3	1243092(22.99%)	693622(17.63%)	0.56
	6	1459881(24.97%)	889587(22.48%)	0.61
	12	1182159(23.33%)	855771(22.92%)	0.72

Note: $^{\flat}$ - a statistically significant difference between the distribution of forecasts based on the two-sided HS test at $p = 0.10$, ‡ $p = 0.05$, or † $p = 0.01$. * - a statistically significant difference between the DNNAR forecast and the NNAR forecast based on the modified DM test at $p = 0.10$, ** $p = 0.05$, or *** $p = 0.01$

5 Discussion

As in Cho (2003), we consider ARIMA and ETS models as benchmarks to further compare and evaluate the performance of the DNNAR model. However, it is noteworthy that unlike in Cho (2003), we use automated and optimized versions of ARIMA and ETS models as introduced to the reader in Section 2.2.

Table 5 below reports the out-of-sample forecasting results. We also compare our results with the findings in Hassani et al. (2017b) where the authors considered the same data set (up until December 2013) and generated forecasts from NNAR, ARIMA and ETS (amongst other models). Unlike in the previous section, where we found the DNNAR model outperforming the NNAR model for all countries across all horizons, here we find varying outcomes. However, the DNNAR model continues to outperform the benchmark models for majority of the countries and horizons considered in this study (see Table 6). At the outset, it is noteworthy that in Hassani et al. (2017b) there was not a single instance when forecasts from the NNAR model outperformed ARIMA or ETS forecasts.

In the case of German tourist arrivals, we find the DNNAR forecasts outperform both ARIMA and ETS forecasts across all horizons with statistically significant results, in contrast to the findings in Hassani et al. (2017b). Moreover, the DNNAR forecasts are found to be 19% and 12% better than ARIMA and ETS forecasts respectively at $h = 1$ step-ahead, 32% and 25% better at $h = 3$ steps-ahead, 44% and 36% better at $h = 6$ steps-ahead, and finally 37% and 51% better than ARIMA and ETS forecasts at $h = 12$ steps-ahead. It is interesting that the ARIMA model outperformed 8 competing models in Hassani et al. (2017b) as it reported the best out-of-sample forecast for German tourist arrivals at $h = 1$ step-ahead, but the improved DNNAR model has succeeded in outperforming ARIMA significantly in this latest study.

However, the DNNAR model is only found to be superior above both benchmark models at a particular horizon at $h = 3, 6$ and 12 steps-ahead when forecasting tourist arrivals in Greece. All differences between forecasts, except between ARIMA and DNNAR models at $h = 1$ step-ahead, are found to be statistically significant. Here, the ARIMA forecast is found to be 18% better than the DNNAR forecast at $h = 1$ step-ahead whilst the DNNAR forecast outperform the ETS forecast by 24% at this horizon. At $h = 3$ steps-ahead, the DNNAR forecasts are 26% and 65% better than forecasts from ARIMA and ETS models, 29% and 70% better at $h = 6$ steps-ahead, and 41% and 55% better at $h = 12$ steps-ahead. In comparison to the findings in Hassani et al. (2017b), the DNNAR model forecasts are now far better than the ARIMA and ETS forecasts across horizons, with the exception of ARIMA which continues to remain superior above ETS and NNAR forecasts for Greece tourist arrivals at $h = 1$ step-ahead.

For Spain, the findings are similar to that of Greece in terms of the DNNAR models superiority above both benchmark models at a particular horizon. In contrast to the findings in Hassani et al. (2017b), the ARIMA model is seen providing better forecasts than ETS at $h = 1$ step-ahead (in addition to outperforming the accuracy of DNNAR forecasts by 10% at this horizon). The DNNAR forecast is 4% better than the ETS forecast at $h = 1$ step-ahead, 5% better than ARIMA and 16% better than ETS at $h = 3$ steps-ahead (though these results are not statistically significant). At $h = 6$ steps-ahead, the DNNAR forecasts are 31% and 40% better than ARIMA and ETS respectively, with statistically significant results. Moreover, in the long run, the DNNAR forecast is 22% better than ARIMA (not statistically significant) and significantly better than the ETS forecast for tourist arrivals in Spain by 29%.

In terms of forecasting tourist arrivals in Italy, forecasts from the DNNAR model outperforms both ARIMA and ETS across all horizons, and reports statistically significant differences in forecasts (except when it outperforms the ARIMA forecast at $h = 12$ steps-ahead). In terms of the accuracy gains, forecasts from the DNNAR model are 25% and 32% better than forecasts from ARIMA and ETS at $h = 1$ step-ahead, 34% and 36% better at $h = 3$ steps-ahead, 33% and 35% better at $h = 6$ steps-ahead, and 33% and 39% better at $h = 12$ steps-ahead. Interestingly, in Hassani et al. (2017b), the ARIMA forecast for Italian tourist arrivals was the best at $h = 6$ steps-ahead in comparison to 8 competing models based on the lowest RMSE. In contrast, we find the DNNAR model is significantly better than ARIMA at $h = 6$ steps-ahead.

The forecast evaluation for tourist arrivals in Cyprus indicates that the DNNAR forecasts outperform ARIMA and ETS forecasts across all horizons (based on the RMSE) with statistically significant results in majority of the horizons considered here. The RRMSE criterion shows that the DNNAR forecasts are 7% better than ARIMA (not statistically significant) and 28% better than ETS at $h = 1$ step-ahead. At $h = 3$ steps-ahead, the DNNAR forecasts are 23% and 48% better than ARIMA and ETS respectively, whilst at $h = 6$ steps-ahead, the DNNAR forecasts are 25% and 54% better than ARIMA and ETS forecasts. At $h = 12$ steps-ahead, the DNNAR forecasts continue to remain 19% and 4% more accurate than ARIMA and ETS forecasts, but these differences are not statistically significant. In comparison to the findings in Hassani et al. (2017b) for tourism demand in Cyprus, unlike the NNAR model, the new DNNAR model is a serious contender for the generation of forecasts for tourist arrivals in Cyprus.

When forecasting tourism demand in the Netherlands, forecasts from the DNNAR model is seen outperforming those from ARIMA and ETS models with statistically significant results across all horizons (except for the differences in forecasts between ARIMA and DNNAR models at $h = 12$ steps-ahead). In fact, at $h = 1$ step-ahead, the DNNAR model is 18% and 15% more accurate than forecasts from ARIMA and ETS respectively, 31% and 18% more accurate at $h = 3$ steps-ahead, 34% and 24% more accurate at $h = 6$ steps-ahead, and 34% and 40% more accurate at $h = 12$ steps-ahead. Once again, these findings from the new DNNAR model represent a significant change in terms of the findings in Hassani et al. (2017b) where the NNAR model failed to outperform ARIMA or ETS forecasts across any of the horizons considered there.

In the case of Austria, we find the new DDNAR forecasts are significantly better than the ARIMA and ETS forecasts across all horizons. The RRMSE criterion enables us to conclude that the DNNAR forecasts are 20% and 19% better than forecasts from ARIMA and ETS at $h = 1$ step-ahead, 19% and 18% better at $h = 3$ steps-ahead, 18% and 14% better at $h = 6$ steps-ahead, 19% and 21% better at $h = 12$ steps-ahead. The NNAR model in Hassani et al. (2017b) provided very poor forecasts for Austria in relation to ARIMA and ETS. However, we find the improved DNNAR model overturns those findings.

In the case of Portugal, we find forecasts from the ARIMA model are significantly better than those from the DNNAR model by 19% at $h = 1$ step-ahead whilst the ETS model forecasts are 17% better than DNNAR forecasts at this same horizon. At $h = 3$ steps-ahead, the DNNAR forecast is 10% better than the ARIMA forecast (not statistically significant) whilst the ETS forecast is 5% better than the DNNAR forecast. However, at $h = 3$ and $h = 6$ steps-ahead, we find the DNNAR model forecasts outperforming both ARIMA and ETS forecasts with statistically significant results. In fact, the DNNAR forecasts are 22% and 21% better than ARIMA and ETS forecasts at $h = 6$ steps-ahead, and 13% and 38% better at $h = 12$ steps-ahead. It is noteworthy that the Hassani et al. (2017b) study found ARIMA to provide the best forecast for Portugal tourist arrivals at $h = 1$ step-ahead in comparison to 8 other models, and our study finds similar evidence for ARIMA at $h = 1$ step-ahead, as it provides significantly better forecasts than the DNNAR model and outperforms the ETS forecast as well. However, unlike forecasts from the NNAR model in Hassani et al. (2017b) at $h = 6$ and $h = 12$ steps-ahead, we find our DNNAR model provides significantly better forecasts than ARIMA and ETS at these horizons.

For Sweden, the ARIMA forecast is 32% better than the DNNAR forecast at $h = 1$ step-ahead (but this result is not statistically significant), whilst the DNNAR forecast is 37% significantly better than the ETS forecast at $h = 1$ step-ahead. At $h = 3$ steps-ahead the ARIMA forecast is 4% better than the DNNAR forecast (not statistically significant) whilst the DNNAR forecast is significantly better than 77% more accurate than the ETS forecast. At $h = 6$ steps-ahead, the DNNAR forecast is 9% better than the ARIMA forecast (not statistically significant) and 50% more accurate and significantly better than the ETS forecast. However, at $h = 12$ steps-ahead, the DNNAR forecast is only 1% better than the ARIMA forecast whilst the ETS forecast for tourist arrivals in Sweden is 21% more accurate and significantly better than the DNNAR forecast. In comparison to the findings in Hassani et al. (2017b), where the ARIMA forecast was the best forecast at $h = 1$ and $h = 6$ steps-ahead, and more accurate than ETS at $h = 3$ steps-ahead, our conclusions are similar at $h = 1$ and $h = 3$ steps-ahead where ARIMA outperforms both DNNAR and ETS forecasts based on the RMSE. However, at $h = 6$ steps-ahead the DNNAR forecast performs better than ARIMA. Our findings at $h = 12$ steps-ahead are consistent with that of Hassani et al. (2017b) where the ETS forecasts were better than ARIMA and NNAR forecasts.

Finally, in the case of UK tourist arrivals, we do not find any instances where the DNNAR forecasts can outperform either ARIMA or ETS forecasts. In fact, at $h = 1$ step-ahead, the ARIMA forecast is 7% better than the DNNAR forecast whilst the ETS forecast is 12% better than the DNNAR forecast (even though the results are not statistically significant). At $h = 3$ steps-ahead, the ARIMA forecast is 31% better than the DNNAR forecast and the ETS forecast is 33% more accurate and significantly better than the DNNAR forecast. At $h = 6$ steps-ahead, both findings are statistically significant and both ARIMA and ETS forecasts are found to be 28% better than the DNNAR forecast.

At $h = 12$ steps-ahead, the gains are comparatively lower, but the ARIMA forecast continues to outperform the DNNAR forecast by 1% whilst the ETS forecast outperforms the DNNAR forecast significantly by 11%. In contrast to the findings in Hassani et al. (2017b), our results indicate that ETS can provide a more accurate forecast than ARIMA across all horizons for UK.

Table 5: Forecasting results for European tourist arrivals from ARIMA, ETS and DNNAR models.

		RMSE(MAPE)	RMSE(MAPE)	RMSE(MAPE)	RRMSE	RRMSE
Country	h	**ARIMA**	**ETS**	**DNNAR**	$\frac{DNNAR}{ARIMA}$	$\frac{DNNAR}{ETS}$
Germany	1	77141(2.27%)	70475(1.97%)	62542(1.72%)	0.81^{***}	0.88
	3	94787(2.93%)	85661(2.39%)	64343(1.74%)	$0.68^{***,\dagger}$	0.75^{***}
	6	109998(3.30%)	96208(2.61%)	61693(1.63%)	$0.56^{***,\dagger}$	$0.64^{**,\ddagger}$
	12	101903(2.53%)	131853(3.73%)	64081(1.74%)	0.63	$0.49^{**,\dagger}$
Greece	1	85641(10.81%)	132901(7.92%)	101406(11.12%)	1.18	0.76^{*}
	3	142993(23.45%)	300093(14.41%)	106148(11.17%)	$0.74^{*,ddag}$	$0.35^{*,\ddagger}$
	6	157201(27.34%)	369571(17.34%)	111385(10.58%)	$0.71^{***,\dagger}$	0.30^{*}
	12	159580(16.37%)	170626(10.57%)	93876(9.62%)	$0.59^{***,\flat}$	$0.55^{**,\dagger}$
Spain	1	184420(3.38%)	211231(3.45%)	203676(4.20%)	1.10	0.96
	3	243188(4.99%)	275151(4.28%)	230396(4.50%)	0.95	0.84
	6	323815(6.31%)	374300(5.83%)	224654(4.53%)	$0.69^{***,\dagger}$	$0.60^{***,\ddagger}$
	12	305390(5.08%)	337033(5.82%)	239078(5.06%)	0.78	$0.71^{*}, \flat$
Italy	1	282950(5.05%)	311960(4.64%)	211462(3.53%)	0.75^{***}	0.68^{***}
	3	334954(5.73%)	344108(4.79%)	220984(3.81%)	0.66^{***}	0.64^{**}
	6	332345(5.15%)	345950(5.26%)	223237(3.80%)	0.67^{***}	0.65^{***}
	12	340124(5.10%)	373507(5.67%)	227447(3.96%)	0.67	0.61^{***}
Cyprus	1	15944(10.88%)	20411(10.75%)	14776(7.76%)	0.93	$0.72^{***,\ddagger}$
	3	23692(16.90%)	35178(26.02%)	18293(10.33%)	0.77^{*}	0.52^{***}
	6	26065(16.60%)	42444(38.45%)	19437(11.59%)	$0.75^{\ddagger}$	$0.46^{***,\dagger}$
	12	26502(18.76%)	22264(11.52%)	21340(12.57%)	0.81	0.96
Netherlands	1	91104(5.24%)	88614(5.39%)	75385(4.57%)	0.82^{*}	0.85^{*}
	3	116460(6.49%)	97892(6.49%)	80267(4.91%)	0.69^{**}	$0.82^{**,\flat}$
	6	129567(7.45%)	112914(7.21%)	85540(5.21%)	$0.66^{*,\ddagger}$	$0.76^{***,\ddagger}$
	12	129583(7.56%)	143001(9.19%)	85505(5.43%)	0.66	$0.60^{***,\dagger}$
Austria	1	127446(5.15%)	125017(5.50%)	101585(4.13%)	$0.80^{**,\ddagger}$	$0.81^{**,\ddagger}$
	3	127452(5.24%)	125214(5.71%)	102824(4.21%)	$0.81^{**,\flat}$	$0.82^{**,\ddagger}$
	6	128478(5.25%)	122117(5.61%)	104908(4.26%)	0.82^{**}	0.86^{***}
	12	134891(5.33%)	137317(5.99%)	108879(4.52%)	$0.81^{***,\ddagger}$	$0.79^{**\ddagger}$
Portugal	1	57429(6.16%)	58700(5.50%)	68563(6.73%)	1.19^{**}	1.17
	3	90778(10.32%)	77508(7.31%)	81368(7.88%)	0.90	1.05
	6	107912(11.04%)	106617(9.49%)	84509(7.99%)	$0.78^{*,\ddagger}$	$0.79^{\ddagger}$
	12	103939(9.24%)	146690(12.72%)	90762(9.38%)	$0.87^{\dagger}$	$0.62^{**,\dagger}$
Sweden	1	32922(4.86%)	69316(6.31%)	43555(4.88%)	1.32	0.63^{**}
	3	42180(4.90%)	131078(11.22%)	43863(5.91%)	1.04	$0.33^{**,\flat}$
	6	44240(5.36%)	80853(10.61%)	40325(5.60%)	0.91	$0.50^{**,\dagger}$
	12	48178(5.79%)	39497(4.43%)	47703(7.17%)	0.99	$1.21^{\ddagger}$
UK	1	431788(9.71%)	409854(9.56%)	460113(11.30%)	1.07	1.12
	3	529319(13.97%)	521734(12.88%)	693622(17.63%)	1.31	$1.33^{**,\ddagger}$
	6	693950(17.98%)	693877(16.30%)	889587(22.48%)	1.28^{*}	1.28^{*}
	12	848969(21.25%)	769017(17.63%)	855771(22.92%)	1.01	$1.11^{\ddagger}$

Note: $^{\flat}$ - a statistically significant difference between the distribution of forecasts based on the two-sided HS test at $p = 0.10$, ‡ $p = 0.05$, or † $p = 0.01$. * - a statistically significant difference between the DNNAR forecast and the competing univariate forecast based on the modified DM test at $p = 0.10$, ** $p = 0.05$, or *** $p = 0.01$

Table 6: Best forecasting model based on lowest RMSE following comparison with benchmark models.

Country	$h = 1$	$h = 3$	$h = 6$	$h = 12$
Germany	**DNNAR**	**DNNAR**	**DNNAR**	**DNNAR**
	(ARIMA)	(ARIMA)	(ARIMA)	(ARIMA)
Greece	**ARIMA**	**DNNAR**	**DNNAR**	**DNNAR**
	(ARIMA)	(ARIMA)	(ARIMA)	(ETS)
Spain	**ARIMA**	**DNNAR**	**DNNAR**	**DNNAR**
	(ETS)	(ARIMA)	(ARIMA)	(ETS)
Italy	**DNNAR**	**DNNAR**	**DNNAR**	**DNNAR**
	(ARIMA)	(ARIMA)	(ARIMA)	(ARIMA)
Cyprus	**DNNAR**	**DNNAR**	**DNNAR**	**DNNAR**
	(ARIMA)	(ARIMA)	(ARIMA)	(ETS)
Netherlands	**DNNAR**	**DNNAR**	**DNNAR**	**DNNAR**
	(ETS)	(ETS)	(ETS)	(ETS)
Austria	**DNNAR**	**DNNAR**	**DNNAR**	**DNNAR**
	(ETS)	(ETS)	(ETS)	(ETS)
Portugal	**ARIMA**	**ETS**	**DNNAR**	**DNNAR**
	(ARIMA)	(ARIMA)	(ARIMA)	(ETS)
Sweden	**ARIMA**	**ARIMA**	**DNNAR**	**ETS**
	(ARIMA)	(ARIMA)	(ARIMA)	(ETS)
UK	**ETS**	**ETS**	**ETS**	**ETS**
	(ARIMA)	(ARIMA)	(ARIMA)	(ARIMA)

Note: Shown within brackets is the best model for forecasting tourist arrivals in a particular country at a given horizon, as per the findings in Hassani et al. (2017b) if one compared the ARIMA, ETS and NNAR forecasting RMSE values in that study. Shown in bold is the best forecasting model as per the forecast evaluation within the current study.

6 Conclusion

This paper begins with the aim of improving NNAR forecasts for tourism demand. In particular, our interest lies in the NNAR forecasts attainable via the *nnetar* algorithm which was considered in Hassani et al., (2015;2017b) as a method for forecasting tourist arrivals. In the aforementioned papers, the results indicated the algorithm performed very poorly when faced with forecasting the highly seasonal tourism demand series. In hope of improving the performance of this NNAR algorithm, we propose the use of SSA for denoising European tourism demand for 10 selected countries and forecasting the reconstructed series with the NNAR model. The results are compared with NNAR forecasts on the raw European tourism demand series and with two other parametric and nonparametric benchmark models.

Our findings present several contributions. Firstly, we find statistically significant evidence which points towards the effectiveness of denoising with SSA prior to fitting a NNAR model when seeking to forecast tourism demand series. In fact the results show considerable gains in accuracy levels when considering a denoised series. Accordingly, the proposed approach has the potential to alter the findings in Hassani et al. (2015; 2017b) and improve upon the accuracy of the NNAR forecasts reported there. Therefore, future forecasting evaluations of tourism demand are strongly advised to consider the DNNAR model introduced in this paper in order to obtain a better forecast from a NNAR model. Secondly, our findings show that the accuracy of forecasts for seasonal time series attainable via Hyndman and Athanasopoulos' (2018) *nnetar* algorithm can be significantly improved by denoising with SSA. Thirdly, our findings show different outcomes to the conclusions in Farway and Chatfield (1995), Nelson et al., (1999), Cho (2003), and Zhang and Qi (2005) where the authors concluded that NN models struggle to generate accurate forecasts when faced with seasonal time series. More importantly, recent evidence in Claveria et al. (2017) also points towards the importance of deseasonalizing when seeking to forecast with NN. In contrast, we clearly demonstrate that denoising is sufficient to significantly improve the accuracy of NNAR forecasts and that NN can handle less noisy seasonal data. Thirdly, the comparison between benchmark forecasting models further evidence the power of the DNNAR model at generating superior forecasts. In fact, when we compare the conclusions from the Hassani et al. (2017b) paper with those from this current study in Table 6, we find strong evidence which shows that the DNNAR model alters previous European tourist arrivals forecasting conclusions.

Nevertheless, this study is not without its limitations. Firstly, as we opt for a recursive forecasting approach, we are not able to comment on the impact of denoising on the NNAR architectures. Secondly, our findings relate only to the NNAR model from the Forecast Package in *R* which is effectively a feed-forward NN model with one hidden layer. It is possible that more complex NN models might perform differently.

Future research should consider a more varied comparison between DNNAR forecasts and those generated via other univariate time series analysis and forecasting approaches in order to determine to what extent these forecasting gains are useful in relation to alternative benchmarks. Moreover, a simulation study which considers the sensitivity of NNAR parameters when forecasting with denoised data could uncover more useful insights into the performance of NNAR models.

References

[1] Alon, I., Qi, M., and Sadowski, R. J. (2001). Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, **8**(3), 147–156.

[2] Altavilla, C. and De Grauwe, P. (2010). Forecasting and combining competing models of exchange rate determination, *Applied Economics, 42*, 3455–3480.

[3] Artola, C., Pinto, F., and Pedraza, P. (2015). Can internet searches forecast tourism inflow? *International Journal of Manpower*, **36**(1), 103–116.

Bagwayo-Skeete, P. F. and Skeete, R. W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, **46**, 4454–4646.

[4] Beneki, C., Eeckels, B., and Leon, C. (2012). Signal extraction and forecasting of the uk tourism income time series: a singular spectrum analysis approach. *Journal of Forecasting*, **31**, 391–400.

[5] Cang, S. (2014). A comparative analysis of three types of tourism demand forecasting models: individual, linear combination and non-linear combination. *International Journal of Tourism Research*, **16**(6), 596–607.

[6] Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. *Tourism Management*, **24**(3), 323–330.

[7] Claveria, O., Monte, E., Torra, S. (2015). Tourism demand forecasting with neural network models: different ways of treating information. *International Journal of Tourism Research*, **17**(5), 492–500.

[8] Claveria, O., Monte, E., and Torra, S. (2017). Data pre-processing for neural network-based forecasting: does it really matter? *Technological and Economic Development of Economy*, **23**(5), 709–725.

[9] Claveria, O., and Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models. *Economic Modelling*, **36**, 220–228.

[10] Curry, B. (2007). Neural networks and seasonality: Some technical considerations. *European Journal of Operational Research*, **179**(1), 267–274.

[11] Farway, J., and Chatfield, C. (1995). Time series forecasting withneural networks: A comparative study using the airline data. *Applied Statistics*, **47**, 231–250.

[12] Franses, P. H., and Draisma, G. (1997). Recognizing changing seasonal patterns using artificial neural networks. *Journal of Econometrics*, **81**, 273–280.

[13] Ghodsi, Z., Silva, E. S., and Hassani, H. (2015). Bicoid Signal Extraction with a Selection of Parametric and Nonparametric Signal Processing Techniques. *Genomics Proteomics and Bioinformatics*, **13**, 183–191.

[14] Golyandina, N., Nekrutkin, V., Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques.* New York-London: Chapman and Hall/CRC.

[15] Gorr, W. L., (1994). Research prospective on neural networkforecasting. *International Journal of Forecasting*, **10**, 1–4.

[16] Hamzacebi, C. (2008). Improving artificial neural networks performance in seasonal time series forecasting. *Information Sciences*, **178**(23), 4550–4559.

[17] Harvey, D. I., Leybourne, S. J., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, **13**, 281–291.

[18] Hassani, H., Ghodsi, Z., Silva, E. S., and Heravi, S. (2016). From nature to maths: Improving forecasting performance in subspace-based methods using genetics Colonial Theory. *Digital Signal Processing*, **51**, 101–109.

[19] Hassani, H., Mahmoudvand, R., Omer, H. N., and Silva, E. S. (2014). A Preliminary Investigation into the Effect of Outlier(s) on Singular Spectrum Analysis. *Fluctuation and Noise Letters*, **13**(4), DOI: 10.1142/S0219477514500291.

[20] Hassani, H., and Silva, E. S. (2015). A KolmogorovSmirnov based test for comparing the predictive accuracy of two sets of forecasts. *Econometrics*, **3**, 590–609.

[21] Hassani, H., Silva, E. S., and Ghodsi, Z. (2017a). Optimizing bicoid signal extraction. *Mathematical Biosciences*, **294**, 46–56.

[22] Hassani, H., Silva, E. S., Antonakakis, N., Filis, G., and Gupta, R. (2017b). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research*, **63**, 112–127.

[23] Hassani, H., Silva, E. S., Gupta, R., and Das, S. (2018). Predicting global temperature anomaly: A definitive investigation using an ensemble of twelve competing forecasting models. *Physica A: Statistical Mechanics and its Applications*, **509**, 121–139.

[24] Hassani, H., Webster, A., Silva, E. S., and Heravi, S. (2015). Forecasting U.S. Tourist arrivals using optimal Singular Spectrum Analysis. *Tourism Management*, **46**, 322–335.

[25] Hirashima, A., Jones, J., Bonham, C. S., and Fuleky, P. (2016). Forecasting in a mixed up world: Now casting Hawaii tourism. UHERO Working Paper, No. 2015-3R.

[26] Hyndman, R. J., and Athanasopoulos, G. (2018). *Forecasting: Principles and practice.* 2nd ed., OTexts. Available via: https://otexts.org/fpp2/. [Accessed: 25.06.2018].

[27] Hydman, R. J., and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, **27**(3), 1–22.

[28] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, **54**(13), 159–178.

[29] Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management*, **21**(12), 331–340.

[30] Law, R., and Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, **20**(1), 89–97.

[31] Li, Y., and Cao, H. (2018). Prediction for Tourism Flow based on LSTM Neural Network. *Procedia Computer Science*, **129**, 277–283.

[32] Li, X., Pan, B., Law, R., and Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management*, **59**, 57–76.

[33] Li, J., Xu, L., Tang, L., Wang, S., and Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, **68**, 301–323.

[34] Liu, Y., Tseng, F., and Tseng, Y. (2018). Big data analytics for forecasting tourism destination arrivals with the applied Vector Autoregressive model. *Technological Forecasting & Social Change*, **130**, 123–134.

[35] Makridakis, S. G., Wheelwright, S. C. and Hyndman, R. J. (1998). *Forecasting: Methods and applications.* Wiley.

[36] Nelson, M., Hill, T., Remus, T., and OĆonnor, M. (1999). Timeseries forecasting using NNs: Should the data be deseason-alized first? *Journal of Forecasting*, **18**, 359–367.

[37] Olmedo, E. (2016). Comparison of Near Neighbour and Neural Network in Travel Forecasting. *Journal of Forecasting*, **35**(3), 217–223.

Pan, B., Wu, D. C., and Song, H. (2012). Forecasting hotel room demand using search engine data. *Journal of Hospitality and Tourism Technology*, **3**(3), 196–210.

[38] Sanei, S., and H. Hassani. (2015). *Singular Spectrum Analysis of Biomedical Signals.* United States: CRC Press.

[39] Silva, E. S., Ghodsi, Z., Ghodsi, M., Heravi, S., and Hassani, H. (2017). Cross country relations in European tourist arrivals. *Annals of Tourism Research*, **63**, 151–168.

[40] Silva, E. S., Hassani, H., and Heravi, S. (2018). Modeling European industrial production with multivariate singular spectrum analysis: A cross-industry analysis. *Journal of Forecasting*, **37**(3), 371–384.

[41] Tseng, F-M., Yu, H-C., and Tzeng, G-H. (2002). Combining neural network model with seasonal time series ARIMA model. *Technological Forecasting and Social Change*, **69**(1), 71–87.

[42] Uysal, M., and El Roubi, M. S. (1999). Artificial Neural Networks versus Multiple Regression in Tourism Demand Analysis. *Journal of Travel Research*, **38**(2), 111–118.

[43] Wu, D. C., Song, H., and Shen, S. (2017). New developments in tourism and hotel demand modelling and forecasting. *International Journal of Contemporary Hospitality Management*, **29**(1), 507–529.

[44] Yang, X., Pan, B., Evans, J. A., and Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, **46**, 286–397.

[45] Zhang, G. P., and Kline, D. M. (2007). Quarterly Time-Series Forecasting With Neural Networks. *IEEE Transactions on Neural Networks*, **18**(6), 1800–1814.

[46] Zhang, G. P., and Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, **160**(2), 501–514.