

Hybrid SVR-SARIMA model for tourism forecasting using PROMETHEE II as a selection methodology: a Philippine scenario

Dharyll Prince Mariscal Abellana, Donna Marie Canizares Rivero, Ma. Elena Aparente and Aries Rivero

(Information about the authors can be found at the end of this article.)

Abstract

Purpose – This paper aims to propose a hybrid-forecasting model for long-term tourism demand forecasting. As such, it attempts to model the tourism demand in the Philippines, which is a relatively underrepresented area in the literature, despite its tourism sector's growing economic progress.

Design/methodology/approach – A hybrid support vector regression (SVR)– seasonal autoregressive integrated moving averages (SARIMA) model is proposed to model the seasonal, linear and nonlinear components of the tourism demand in a destination country. The paper further proposes the use of multiple criteria decision-making (MCDM) approaches in selecting the best forecasting model among a set of considered models. As such, a preference ranking organization method for enrichment of evaluations (PROMETHEE) II is used to rank the considered forecasting models.

Findings – The proposed hybrid SVR-SARIMA model is the best performing model among a set of considered models in this paper using performance criteria that evaluate the errors of magnitude, directionality and trend change, of a forecasting model. Moreover, the use of the MCDM approach is found to be a relevant and prospective approach in selecting the best forecasting model among a set of models.

Originality/value – The novelty of this paper lies in several aspects. First, this paper pioneers the demonstration of the SVR-SARIMA model's capability in forecasting long-term tourism demand. Second, this paper is the first to have proposed and demonstrated the use of an MCDM approach for performing model selection in forecasting. Finally, this paper is one of the very few papers to provide lenses on the current status of Philippine tourism demand.

Keywords Futures, Multiple criteria decision-making, Hybrid forecasting model, Philippine tourism, Tourism forecasting

Paper type Research paper

Received 12 July 2019
Revised 30 December 2019
9 March 2020
Accepted 31 March 2020

© Dharyll Prince Mariscal Abellana Donna, Marie Canizares Rivero, Ma. Elena Aparente and Aries Rivero. Published in *Journal of Tourism Futures*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

1. Introduction

The tourism industry is a growing economic sector in the globe. According to the World Travel and Tourism Council (WTTC), it has contributed to 10% of the global gross domestic product (GDP) (Ghalekhondabi *et al.*, 2019). As such, 1 out of 11 jobs was relevant to tourism. With such a contribution to the globe, tourism is considered an important economic sector. In the literature, tourism demand forecasting is one of the most active research areas. The perishable nature of tourism services primarily motivates the need for generating forecasts; hence, not selling it at the right time would lead to lost sales for all related businesses (Ghalekhondabi *et al.*, 2019). Scholars maintain that requirement of destination countries for substantial infrastructure investments and promotional activities, which aid in

safeguarding positive return on investments, further establish the need for accurate forecasting. Likewise, accurate forecasting supplements activities at the business level, particularly in devising strict budget requirements. The interest in developing accurate forecasting models has been active since the 1980s (Hassani *et al.*, 2017). Although Song and Li (2008) and Ghalekhondabi *et al.* (2019) report that much of the significant developments skyrocketed since the year 2000. Three significant primary modeling approaches formed from such developments:

1. time series;
2. causal modeling; and
3. emerging methodological approaches (e.g. artificial intelligence, fuzzy logic and gray-based methods).

Despite such progress, results in literature have yet to point at a single best approach in all situations. Scholars have attributed such conflicting results in the literature to the modeling capabilities of the different approaches. For instance, time-series methods are better at modeling the linear components of tourism data, whereas artificial intelligence methods are better at modeling the nonlinear components. As such, the hybridization of different models has become a contemporary research area in the literature primarily to combine the strengths of forecasting models. Several hybrid approaches have been proposed in the literature, such as Grey-Markov (GM) models with artificial neural networks (ANN) (Hu *et al.*, 2018), among others. However, most of the proposed approaches are challenged in their estimation approach. For instance, due to being a gradient-based approach, the optimization process of ANNs are susceptible to fall on local optima, which would adversely affect the predictive performance of the model. Such situations are crucial concerns, especially when voluminous data paired with non-convexity of the optimization problem governing a forecasting approach challenge the modeling process. These issues could be addressed by developing hybrid models that do not fall into those issues (e.g. adopting models that are governed by convex optimization formulation). In the relevant literature, scholars have found the capability of the hybrid seasonal autoregressive integrated moving averages (SARIMA) – support vector regression (SVR) in forecast modeling. Its applicability has achieved fruitful results in relevant domains, as demonstrated by Lee *et al.* (2017), Xu *et al.* (2019) and Ruiz-Aguilar *et al.* (2014), among others.

Mainly, the SVR is adopted to estimate the nonlinear components of the time series data, whereas the SARIMA is adopted for modeling the seasonal and linear components. Despite success in several application domains, its capability has not been explored in the tourism demand forecasting literature. The use of the SVR-SARIMA would require less specification and estimation procedures as compared to causal models (which is firmly based on theory). This paper proposes a hybrid SVR-SARIMA for tourism demand modeling. Features such as strong seasonality and nonlinear trend enable such an approach to be suitable in tourism demand forecasting. Furthermore, this paper proposes the use of multiple criteria decision-making (MCDM) approaches in the model selection phase. A study is conducted in the Philippines, a developing country that significantly depends on tourism as an economic sector, to provide lenses on the status of tourism demand in a developing country. Despite the significant dependence of the country on tourism as an economic sector, very few attempts have been made in the literature to explore its dynamics, with Rufino (2011) and Rufino (2016) as the only works tackling tourism forecasting. This paper has three significant contributions to the literature. First, it explores the capability of a hybrid model, particularly the SVR-SARIMA model, in developing an accurate tourism demand forecasting model. Second, it demonstrates the use of MCDM approaches in selecting the best forecasting model. Finally, it provides practical lenses on the dynamics of tourism demand in the Philippines, which is relatively underrepresented in the literature. Moreover, the results of this paper would be beneficial to various stakeholders such as tourism

industries, managers, government agencies and policymakers, among others, in that it provides them with a framework in developing a hybrid tourism forecasting method that accurately models tourism demand.

2. Literature review

2.1 Tourism demand modeling and forecasting

Tourism demand modeling and forecasting is the attempt to predict future values of tourism demand indicators such as tourist arrivals, search index and expenditure, among others (Song and Li, 2008). It is a contemporary research area in the tourism literature primarily due to the perishable nature of tourism services, which urges managers and decision makers to generate accurate forecasts that would reduce the risk of business decision-making (Gunter, 2018). As such, the tourism demand and modeling literature have placed much emphasis on surging methodological developments, which is evident with the spurt of published works in the tourism literature since 1960 (Song and Li, 2008). With this, several comprehensive and systematic reviews have been conducted by scholars in literature such as Song *et al.* (2019), Jiao and Chen (2018), Song and Li (2008), Li *et al.* (2005), Lim (1997a) and Lim (1997b).

With the upsurge of methodological developments in tourism demand modeling and forecasting, published works in literature cluster into two major categories:

1. quantitative; and
2. qualitative methods (Song *et al.*, 2019; Ghalekhondabi *et al.*, 2019).

For brevity, this paper focuses on providing reviews only in the first category, being more relevant to this study. While being quantitative entails that various indicators (e.g. index, count, proportions) could be adopted in developing models, the tourism literature has found demand to be most useful (Song *et al.*, 2019; Akin, 2015). As such, the tourist arrivals from several origin countries to a destination country has been the most widely adopted tourism demand indicator (Song *et al.*, 2019).

The quantitative methods in tourism demand modeling and forecasting are generally composed of:

- time series models;
- econometric (or causal); and
- emerging methodology (e.g. artificial intelligence, fuzzy logic, hybrid models) (Song and Li, 2008).

Despite reported success in the literature using such approaches, none has yet to be found to be the best model in all situations (Song and Li, 2008; Ghalekhondabi *et al.*, 2019). Several conflicting results have been found in the current literature. Above all, no single forecasting model is the best in all situations (Song and Li, 2008; Ghalekhondabi *et al.*, 2019).

2.1.1 Time series forecasting Time series forecasting approaches have been the most widely used in the tourism demand forecasting literature (Song and Li, 2008). Several approaches compose the time series forecasting literature such as exponential smoothing (ES) or error-trend-seasonal (ETS) model (Athanasopoulos and de Silva, 2012; Kourentzes and Athanasopoulos, 2019; Gunter and Onder, 2016), state-space modeling (Jorge-Gonzalez *et al.*, 2019), filtering (Bosupeng, 2019) and spatial-temporal modeling (Yang and Zhang, 2019), among others. Most of the published works in this domain use the use of the Box–Jenkins class of models (Ghalekhondabi *et al.*, 2019). This class of models includes the autoregressive moving averages (ARMA), autoregressive integrated moving averages (ARIMA) and SARIMA (Ghalekhondabi *et al.*, 2019; Song and Li, 2008).

The benchmark models used in testing the accuracy of proposed time series models include the ES, naïve models and seasonal naïve models. The ES is a conventional technique used for smoothing time series data using an exponential window function. Several works used the ES/ETS as a benchmark model such as [Rice et al. \(2019\)](#), [Li et al. \(2019\)](#) and [Gunter and Onder \(2016\)](#), among others. Likewise, the naïve and seasonal naïve models have also been used conventionally in literature for comparing the trend and seasonal change of proposed models, as can be seen in [Gunter and Onder \(2016\)](#). Moreover, as [Gunter and Onder \(2015\)](#) pointed out, the Box–Jenkins models have also been considered benchmark models in the literature due to numerous works using such techniques.

The time series forecasting approach has been successful in forecasting tourism demand, as pointed out by [Song et al. \(2019\)](#), [Rufino \(2011\)](#) and [Baldigara \(2015\)](#), among others. Several works in literature have successfully adopted time-series approaches in modeling and forecasting tourism demand. For instance, [Petrevska \(2017\)](#) tested several specifications of the ARIMA model to provide a short-run estimation of international tourism demand in Macedonia. Similarly, [Borhan and Arsad \(2014\)](#) adopted a SARIMA approach in forecasting international tourism demand to Malaysia from the USA, Japan and South Korea. Moreover, [Rufino \(2011\)](#) developed time series models using SARIMA on the top 12 tourist generating countries for Philippine tourism.

2.1.2 Causal modeling approaches Apart from time series forecasting, causal modeling approaches have also been extensively used to model time series data. Such approaches can also be seen as econometric models, as manifested by several scholars ([Gunter, 2018](#); [Assaf et al., 2018](#)). This type of approach makes use of exogenous variables to generate tourism demand forecasts ([Gunter, 2018](#)). Such techniques depend heavily on the assumption that all explanatory variables are exogenously determined ([Assaf et al., 2018](#)). Causal modeling techniques show potential in the literature, as they are easily interpretable due to being theory-grounded, which aid in further providing a thorough explanation of the meaning behind tourists' decision-making processes ([Assaf et al., 2018](#)).

The causal models used in literature include Vector Autoregressive Moving Average (VARMA), Bayesian global VAR (BGVAR), seasonal VARMA, autoregressive distributed lag (ARDL) and error correction models (ECM). As such, several works have shown the applicability of causal modeling approaches in forecasting tourist demand. For instance, [Assaf et al. \(2018\)](#) demonstrated the ability of the BGVAR model to capture the spillover effects of international tourism demand in the Southeast Asian region. Likewise, [Lopes et al. \(2017\)](#) modeled and forecasted the tourist demand in the Algarve region of Portugal using seasonal VARMA models.

Moreover, [Lin et al. \(2015\)](#) adopted the ARDL framework to model the main factors that affect Chinese outbound tourism and predict the total tourist outflows and departures from China between the periods 2012 to 2020. Likewise, ECM was adopted by [Pokharel et al. \(2018\)](#) to determine short-term and long-term relationships of the number of international tourists visiting Nepal in the years between 1962 and 2012. Another causal model called the time-varying parameter (TVP) had gained some attention from scholars in the tourism forecasting literature, as can be seen in [Jorge-Gonzalez et al. \(2019\)](#) and [Assadzadeh and Mamipour \(2013\)](#). Despite such attention, the TVP is not as prevalent as the other causal models in tourism demand forecasting ([Panahi et al., 2015](#)).

2.1.3 Artificial intelligence approaches Along with success in time series and econometric approaches to tourism demand forecasting, artificial intelligence (AI) approaches have emerged to be promising in the literature. AI is a derivative of human intelligence (HI) and is developed based on HI to solve problems by not only the routine and repetitive activities but also the behavior on intelligence level ([Fang et al., 2018](#)). The tourism demand forecasting literature consists of several AI approaches such as ANN ([Constantino et al., 2016](#)), deep learning ([Law et al., 2019](#)), SVR ([Liu et al., 2018](#)) and fuzzy time series

(Sakhuja *et al.*, 2016), among others. Such approaches have attained considerable success in the literature due to their independence of strong model assumptions and resilience in treating nonlinearities. As such, several works have adopted the use of AI approaches in tourism demand modeling and forecasting. For instance, Constantino *et al.* (2016) adopted the ANN for modeling and forecasting the tourism demand in Mozambique using monthly data of tourist overnight stays in hotels during the periods 2004 to 2013. Likewise, Law *et al.* (2019) used a deep learning approach to modeling monthly tourism demand in China. Similarly, Liu *et al.* (2018) developed a forecasting model of the monthly tourism demand in Taiwan in the period 2006 to 2016 using SVR. Also, Sakhuja *et al.* (2016) adopted a fuzzy time series to model the monthly tourist arrivals in Taiwan in the period 2000 to 2005.

2.1.4 Forecast combinations or hybrid approaches Driven by the absence of a single approach in forecasting tourism demand, several works in literature proposed the hybridization of different approaches to improve the accuracy of forecasting models. For instance, Pai *et al.* (2014) developed a combined fuzzy c-means (FCM) algorithm with logarithmic least-squares support vector regression (LLS-SVR) to model the monthly tourist arrivals in Taiwan and Hong Kong during the period from 1969 to 2010. Likewise, Hu *et al.* (2018) improved the predictive accuracy of the GM model with neural networks using two case studies. Similarly, Chen (2011) tested the combination of several linear (e.g. ARIMA, exponential smoothing) and nonlinear models (e.g. SVR, backpropagation neural networks [BPNN]) to increase the predictive accuracy of tourism forecasting models further.

Wen *et al.* (2019) demonstrated the capability of the hybridization of the autoregressive integrated moving averages with exogenous variable (ARIMAX) and non-linear autoregressive with exogenous variable (NARX) model in improving predictive accuracy by capturing the linear and nonlinear features of a data set. Moreover, Shen *et al.* (2019) combined feature selection (FS) and SVR with particle swarm optimization (PSO – SVR) to model the monthly tourist arrivals in Singapore from 1987 to 2017 and obtained a reasonably high predictive accuracy. General results in the literature point out improvement in predictive accuracy when combining several approaches. Hence, the hybridization of several approaches, particularly linear and nonlinear approaches, is a contemporary research area in the current literature.

2.2 Model selection

With the existence of several approaches in modeling and forecasting tourism demand, the decision to select the best model is a crucial process to be made in developing the most accurate forecasting model, which is essential in tourism managerial decision-making. As such, several forecasting criteria have been widely used in the literature to aid decision makers in selecting the best performing model. On top of it, three general measures have been adopted – measures of magnitude such as root mean square error (RMSE) (Hassani *et al.*, 2017); measures of turning point such as direction of change (DC) (Chen, 2011); and measure of trend change (Witt and Witt, 1991). Conventionally, a model is selected based on its performance of one or several of such criteria. For example, a decision maker may pertain to the best model as one having the lowest error, thus, would base on RMSE.

Conversely, a decision maker may also pertain to the best model as one which has precisely followed the direction of the actual series. Hence, the selection of the best forecasting model is mostly a preference of the decision maker. A critical argument in literature, however, is deciding the significant differences between two or more competing models, mainly if their differences are almost negligible. Scholars point out the role of statistical significance testing. In this respect, Song and Li (2008) argued that it would be difficult to assert whether one model is better than others without the support of statistical testing. Efforts have been made in the literature to establish means of selecting models based on formal statistical testing. Such methods include:

- the Morgan–Granger–Newbold (MGN) test which determines the equality of two forecast series values using a loss differential function;
- the Meese–Rogoff (MR) test which is a test of equal forecast accuracy when the forecast errors are serially and contemporaneously correlated;
- the Diebold–Mariano (DM) test which considers a model-free test of forecast accuracy; and
- West’s asymptotic inference on predictive ability which provides a formal asymptotic theory for inference about moments of smooth functions of out-of-sample predictions and prediction errors, among others (Mariano, 2004).

Although calling for the proposition of such an approach would aid decision makers in more confidently selecting the best model, such proposals were challenged by Armstrong (2007) and Kostenko and Hyndman (2008). Armstrong (2007) showed, in the context of forecasting, how significance testing would be harmful to the development of scientific knowledge. For one, they distract the researcher from using proper methods. As such, Armstrong (2007) asserted that scholars should instead report on effect sizes, confidence interval, and replications, among others. Likewise, Kostenko and Hyndman (2008) claimed that significance testing has little to do with forecasting. With this, they pointed out the existence and use of simple but powerful approaches that are not statistical tests but are useful in ranking and selecting forecasting models.

3. Methodology

3.1 Support vector regression

The support vector machine (SVM) is a machine-learning algorithm first proposed by Boser *et al.* (1992) to solve classification problems. The SVR is an extension of the SVM developed to deal with regression problems (Xu *et al.*, 2019). In the tourism demand literature, the SVR has been adopted in works such as Pai *et al.* (2006), Cai *et al.* (2009) and Akin (2015), among others, due to its ability to perform well in modeling nonlinear data. The SVR does this by mapping the input data into a higher dimensional space to formulate the non-linear relationship of the data. With the training data denoted as $\{(x_i, y_i) | x_i \in \mathbb{R}^n, y_i \in \mathbb{R}\}$, y_i is the actual output value, SVR function is formulated as follows:

$$f(x) = \omega^T \varphi(x) + b \quad (1)$$

where $\varphi(x)$ is the feature function of inputs; $f(x)$ is the output; ω and b are adjustable parameters. To simplify the estimation process, the SVR is converted to a quadratic programming (QP) problem in literature and generally formulated as follows:

$$\min(\omega, \xi, \xi^*) = \frac{1}{2} \omega^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (2)$$

$$\text{subject to } \begin{cases} y_i - \omega^T \varphi(x_i) - b & \leq \varepsilon + \xi_i^* \\ -y_i + \omega^T \varphi(x_i) + b & \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* & \geq 0 \end{cases} \quad (3)$$

where a hyperparameter, used to represent the precision to which the optimization problem is feasible, $C > 0$ determines the trade-off between the flatness of the function f , and ξ , ξ^* are slack variables introduced to cope with otherwise infeasible constraints of the QP. A thorough discussion of SVR is provided by Smola and Schölkopf (2004) and is not discussed in this paper for brevity.

3.2 Seasonal autoregressive integrated moving averages

The SARIMA model is an extension of the ARIMA model developed to capture the seasonal component of a time series (Baldigara, 2015). The SARIMA model is a widely used time series approach in the tourism demand forecasting literature and has been found by several scholars to generate forecasts with reasonable accuracy (Song *et al.*, 2019). The SARIMA model incorporates both the seasonal and non-seasonal components of a time series. As such, the SARIMA model is represented as follows:

$$ARIMA(p, d, q)(P, D, Q)_S \quad (3)$$

where p, d, q are non-seasonal AR, differencing and MA orders, P, D, Q are seasonal AR, differencing and MA orders, and S is the timespan of repeating seasonal patterns. The SARIMA model of Box and Jenkins (1970) is formulated as:

$$\phi(B) \Phi(B^S)(1-B)^d(1-B^S)^D Y_t = \theta_0 + \theta(B) \Theta(B^S) \varepsilon_t \quad (4)$$

where:

$\phi(B) = 1 - \sum_{i=1}^p \phi_i B^i$ is the p -order non-seasonal AR model;

$\theta(B) = 1 - \sum_{i=1}^q \theta_i B^i$ is the q -order non-seasonal MA model;

$\Phi(B^S) = 1 - \sum_{i=1}^P \Phi_i B^{iS}$ is the P -order seasonal AR model;

$\Theta(B^S) = 1 - \sum_{i=1}^Q \Theta_i B^{iS}$ is the Q -order seasonal MA model;

$(1-B)^d$ = denotes the non-seasonal differencing of order d ;

$(1-B^S)^D$ = denotes the seasonal differencing of order D

ε_t = the error term;

B = backshift operator; and

S = seasonal order.

For brevity, the SARIMA model is only discussed briefly in this paper. As such, a comprehensive discussion is provided by Baldigara (2015), which is contextualized for tourism demand forecasting.

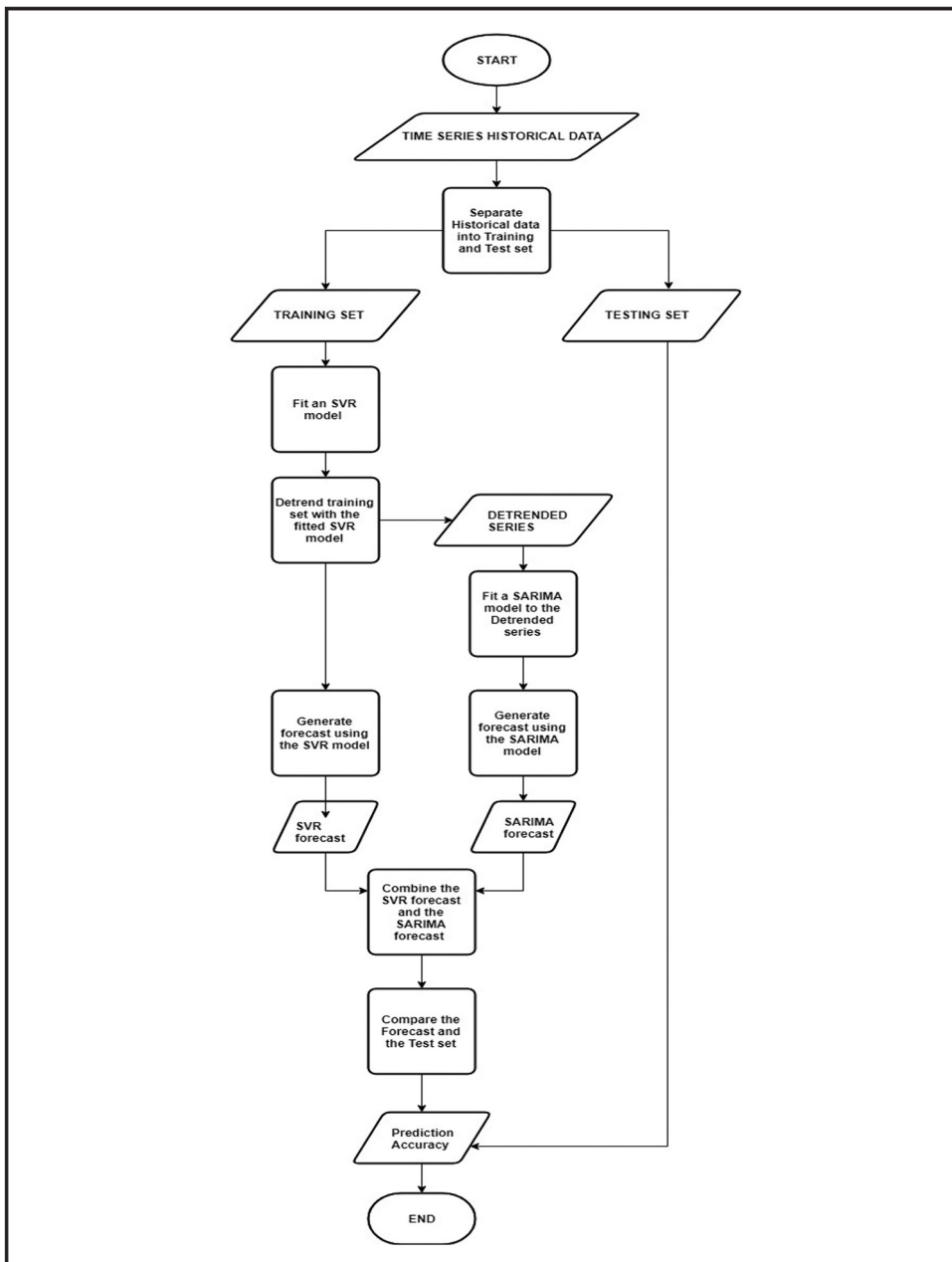
3.3 Proposed hybridization approach

The combination of the SARIMA and SVR has already been explored in the relevant literature for combining the strengths of modeling linear and nonlinear data, respectively. For instance, Xu *et al.* (2019) used a SARIMA-SVR to forecast statistical indicators in the aviation industry for capacity management and planning purposes. Likewise, Zou *et al.* (2019) used a combined SARIMA and SVR to develop a predictive model that could forecast the Hand-foot-mouth disease (HFMD) incidence in China. Moreover, Lee *et al.* (2017) used a SARIMA-SVR model to improve atmospheric pollution forecast accuracy based on the analysis of atmospheric pollution data in the Internet-of-Things (IoT) environment. Finally, Ruiz-Aguilar *et al.* (2014) adopted a SARIMA-SVR model to forecast the inspection volume at the European border, notably the Border Inspection Post of Port of Algeciras Bay. Such works provide support on the capability of the combined linear and nonlinear models, particularly, the SVR and SARIMA models, in improving forecast accuracy.

In tourism demand forecasting, no work has explored on the hybridization of the SVR and SARIMA in developing a forecasting model. As such, this paper explores the applicability of the combined SVR and SARIMA models in generating tourist demand forecasts. In the relevant literature, the hybrid SVR-SARIMA model uses the SARIMA model first to fit the seasonal

components of the original series and fits an SVR model afterward to estimate the nonlinear residuals. However, in this paper, the SVR model will be fitted first, followed by the fitting of the SARIMA model afterward. This paper argues that such strategy is more fit in the tourism demand forecasting literature due to the nonlinear trends exhibited by tourism demands as can be verified in the time series data used by [Burger et al. \(2001\)](#), [Goh & Law \(2002\)](#), and [Rufino \(2011\)](#), among others. On the contrary, the SARIMA-SVR developed in literature dealt with time series that have stationary or linear trends. The proposed methodology in this paper is similar to the SVR-ARIMA algorithm used by [Alwee et al. \(2013\)](#). Hence, in this paper, the SVR will be used to fit the trend of the tourism demand time series; then afterward, the SARIMA will be used to fit the seasonal and non-seasonal residuals of the SVR-detrended tourism demand time series. The procedure proposed in this paper, as presented in [Figure 1](#), is described as follows:

Figure 1 Proposed Hybrid SVR-SARIMA algorithm



- *Step 1. Visualize the series.* Plotting the series helps in identifying the behavior of the time-series data. As such, it allows the decision-maker to decide on the potential models that would be appropriate to the time series data being studied.
- *Step 2. Separate training set and testing set.* After having a visualization of the time series data. The ratio of the training and test set must be determined. No standard paradigm exists in the literature in guiding scholars regarding the selection of the proportion of training and test sets. As such, the researchers decide this.
- *Step 3. Fit an SVR model to the training set.* The SVR model will be used to fit the trend of the original time series data because of the nonlinear trend behavior of tourism time series data, as can be observed in several works in the literature.
- *Step 4. Detrend the original training set by the SVR function obtained in Step 3.* After obtaining an SVR model for the original series, the series can be made stationary by subtracting the SVR trend from the original series. This process decomposes the original time series into its trend and non-trend components. Whereas the SVR estimated the trend component, the non-trend components (i.e. seasonality, cycles, noise) will be fitted using the SARIMA model.
- *Step 5. Fit a SARIMA model to the residuals obtained in Step 4.* As mentioned previously, the residuals generated from Step 4 will be fitted using a SARIMA model. The SARIMA model will fit the seasonal and cycles of the time series data.
- *Step 6. Generate separate ex-post forecasts using the models obtained in Step 3 and Step 5.* The SVR and SARIMA model will generate separate ex-post forecasts.
- *Step 7. Combine the forecasts obtained in Step 6.* Consistent with the hybridization approaches in the literature. The nonlinear component (i.e. estimated by the SVR) and linear component (i.e. estimated by the SARIMA) will be combined to reconstruct the original time series. As such, their combined forecast will be compared to the test data to evaluate the performance of the hybrid model.

3.4 Preference ranking organization method for enrichment of evaluations

The selection of the best forecasting model can be considered as a decision-making problem. In a decision-making problem, a decision maker has a set of alternatives to select from and a set of criteria to evaluate such alternatives. In selecting the best forecasting model, a decision maker is confronted with several competing models and several performance criteria. As such, the set of forecasting models can be considered as the decision alternatives, whereas the performance criteria (e.g. MAPE, RMSE) can be considered as the decision criteria. With such attributes, the selection of the best forecasting model can be approached in an MCDM environment. Several MCDM approaches exist in the literature, such as the analytic hierarchy process (AHP), analytic network process (ANP) and preference ranking organization method for enrichment of evaluations (PROMETHEE) I and II. Among such methods, the PROMETHEE II method will be selected to rank the forecasting models according to their performance.

Unlike AHP and ANP, PROMETHEE is a preference function based outranking method, which is a particular type of MCDM approach that can provide a ranking ordering of the decision options ([Athawale and Chakraborty, 2010a, 2010b](#)). The PROMETHEE II will be selected over PROMETHEE I, as the objective of the paper is only to provide a rank ordering of the alternatives. While the PROMETHEE I is capable of providing such rank-ordering, it goes further into deriving the partial ordering, which adds more computational steps into the procedure. For brevity, the procedure of the PROMETHEE II will not be discussed in this paper. A comprehensive discussion on PROMETHEE II is provided by [Athawale and Chakraborty \(2010a, 2010b\)](#), [Brans et al. \(1986\)](#) and [Sen et al. \(2015\)](#).

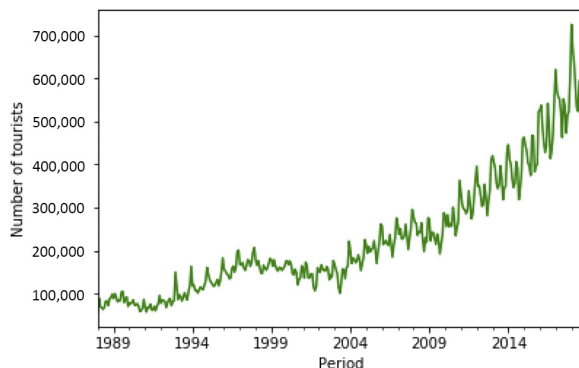
4. The Philippine scenario and results

The Philippines is an archipelagic country in Southeast Asia. It consists of about 7,641 islands that are categorized into three main geographical divisions from north to south: Luzon, Visayas and Mindanao. Being a tropical country makes the Philippines one of the active nations when it comes to tourism. Some of the most popular tourist destinations of the country are the Boracay White Beach, Bohol Chocolate Hills, Puerto Princesa Subterranean River National Park, Intramuros, Mactan, Oslob and Vigan City, among others. Indeed, tourism plays a huge role in the country's economic progress. Republic Act No. 9593, otherwise known as the Tourism Act of 2009, recognizes tourism as an indispensable element of the Philippine economy that generates investment, employment, growth and national development. The tourism industry's contribution to the economy grew in 2018, according to data released by the Philippine Statistics Authority (PSA) in June of 2019. Moreover, the tourism sector contributes 12.2% share to the country's total gross domestic product (GDP) in 2017 and has been increasing since 2012 (PSA, 2018). The tourist demand of the Philippines from January 1988 to December 2018, as presented in Figure 2, can be seen to have increased over time.

Although such growth entails economic progress for the Philippines, it is an indicator for tourism managers, government agencies, policymakers and other stakeholders to be more proactive in dealing with the capacity decisions of the Philippines' tourist destinations. With the growing tourism demand of the country, capacity-planning activities become more and more relevant. One of the most crucial activities in capacity planning is forecasting future demand. In the Philippines, it is crucial to develop annual forecasts. Such forecasts are used in the development of long-term investment plans (e.g. airports, hotels and destination developments), policymaking and annual budget allocation, among others.

Moreover, such initiatives can also be linked to the political and administrative agenda of the government. Primarily, that in the Philippines, economic activities are highly driven by the executive branch of the government (which changes administration every six years through a national election). With such initiatives being linked to political and government activities, developing annual forecasts would aid the government and other stakeholders in implementing proactive plans, especially with long-term investments. Likewise, with the current administration's goals of sustainable development through the 'Build, build, build'

Figure 2 Philippines inbound tourist arrivals from January 1988 to December 2018



Source: The dataset is obtained from www.ceicdata.com/en/indicator/philippines/visitor-arrivals. The image is generated using the pandas library in Python through Google Colaboratory

advocacy, long-term forecasting would be highly relevant to the country in facilitating proactive development strategies.

As such, this paper attempts to model tourism demand using a hybrid SVR-SARIMA due to the nonlinear trend and seasonal pattern exhibited by the tourism demand in the Philippines. With the proposed model, estimation would be relatively more straightforward as compared to causal models, which would require careful statistical testing and estimation procedures. With the absence of a consensus in the current literature regarding which approach would yield more accurate forecasts (i.e. disintegrating source markets versus forecasting the tourism demand as a whole), a direct approach of forecasting the aggregate tourism demand is useful in simplifying the modeling process. The study is conducted in the Philippines to provide insights on its tourism dynamics as well as other developing nations. As such, very few efforts are visible in the current literature despite the growing importance of the tourism sector in such countries, particularly the Philippines.

Using Step 1 in the SVR-SARIMA methodology, the historical data is plotted in [Figure 2](#). With Step 2, the series is divided into training and testing sets. In this paper, the training set is the series from January 1988 to December 2013, while the testing set is the series from January 2014 to December 2018. This splitting of the dataset follows from a focus group discussion of a panel of experts regarding which periods to consider given the disturbances that may affect the tourism demand. It can be observed that the series is nonlinearly trending.

In the Philippine scenario, disruptions (e.g. natural disasters, acts of terrorism and political unrest) have occurred, which may cause structural breaks in the forecasting model. For instance, with the change in the constitution of the Philippines in the year 1986 following from the people power revolution, tourism demand may have dropped, as can be seen in [Figure 2](#) ([Totanes, 1998](#)). In the year 1992, it can also be seen that a structural change occurred. Such structural change can be attributed to the 1992 Presidential election being the first general election held under the new regime. In the year 2001, the second EDSA people power revolution took place, which may have caused a structural change in the tourism demand. In February of the year 2004, a terrorist attack that is considered as the worst in the Philippines, bombed the *SuperFerry XIV* killing 116 people. Such an event may have caused a decrease in tourism demand. In the same year, from November to December, four typhoons devastated the country, with 1,060 people dead, more than 560 missing, and 850,000 displaced ([World Almanac Books, 2007](#)).

In the year 2008, several disruptions, including the 2008 Global Financial Crisis, terrorist attacks, and devastation by Typhoon Frank, have caused a structural change to the tourism demand in the Philippines. Followingly, in the year 2013, three significant disturbances affected Philippine tourism. Such disturbances are:

1. terrorist attacks at Cagayan de Oro and Cotabato City;
2. the devastation of a 7.2 magnitude earthquake killing 144 people ([Whaley, 2013](#)); and
3. the catastrophic destruction by Typhoon Haiyan (known as Typhoon Yolanda in the Philippines) killing about 6,300 people ([Mullen, 2013](#); [Williams, 2013](#)).

It can be observed that after the year 2013, structural changes have not been exhibited by tourism demand in the Philippines.

The strength of the combined SVR-SARIMA lies in its ability to estimate the nonlinear, linear, and seasonal components of the tourism demand with less model specification in contrast to strictly causal and time series models. As such, the SVR becomes an appropriate regression technique due to its capability to fit nonlinear relationships. In Step 3, an SVR is fitted to the training set in Python using the SVR class in *scikit-learn*. The radial basis function (RBF) is used as the kernel for the SVR due to the non-polynomial behavior of the

training data. The fitted SVR of the training set is presented in Figure 3. The trend fitted by the SVR will be used to detrend the training set to transform the series into a stationary process, as presented in Figure 4. By conducting an augmented Dickey–Fuller (ADF) test, it was verified that the detrended series is stationary. The nonstationarity of a series is a null hypothesis of the ADF test. Thus, the rejection of the null hypothesis implies that the series is already stationary. As such, the p -value obtained in the test is $p = 0.015045$, which makes the result significant at the 0.05 level of significance. Hence, the detrending of the training data with the SVR trend transforms it into a stationary process.

A SARIMA model is then fitted to the detrended series, as presented in Figure 4. The ‘*auto-arma*’ functionality of Python’s ‘*pmdarima*’ library allows the automatic estimation of the SARIMA parameters. The criteria used by the function to select the best model is the Akaike information criterion (AIC). As such, the best model (i.e. having the lowest AIC) is the $SARIMA(2,0,0)(2,0,0)_{12}$, as presented in Table 1. Using the fitted SVR and SARIMA models, separate forecasts are generated by each model. As such, separate *ex-post* forecasts are made from January 2014 to December 2018. The SVR and SARIMA forecasts are presented in Figures 5 and 6, respectively. To reconstruct the time series, the SVR and SARIMA forecasts are combined, as presented in Figure 7. The combined forecast of the SVR and SARIMA models compose the proposed hybrid SVR-SARIMA model. The forecasts of the proposed model are then compared to the actual tourist arrivals to evaluate its performance.

Figure 3 Fitted SVR to training data

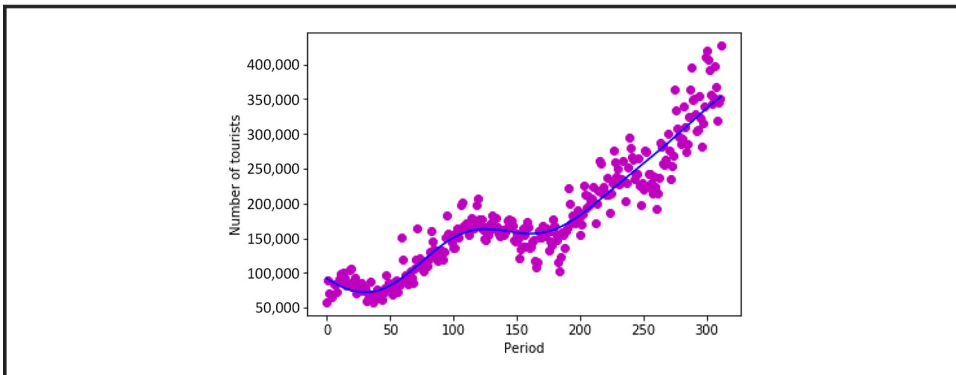


Figure 4 Detrended series

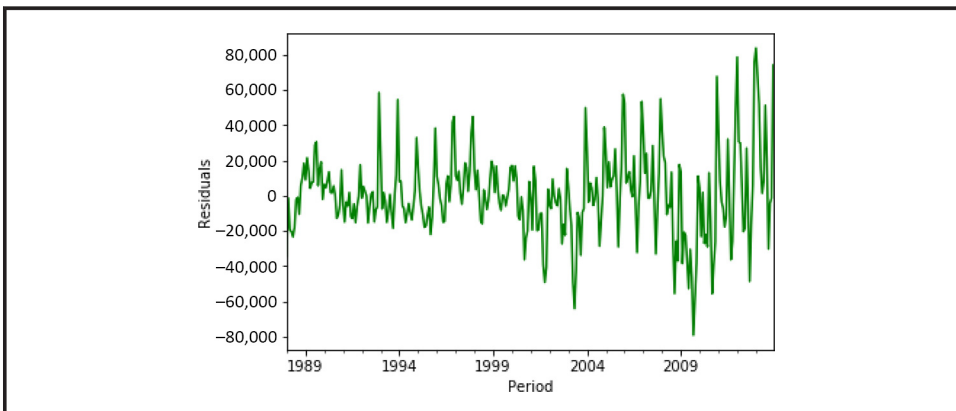
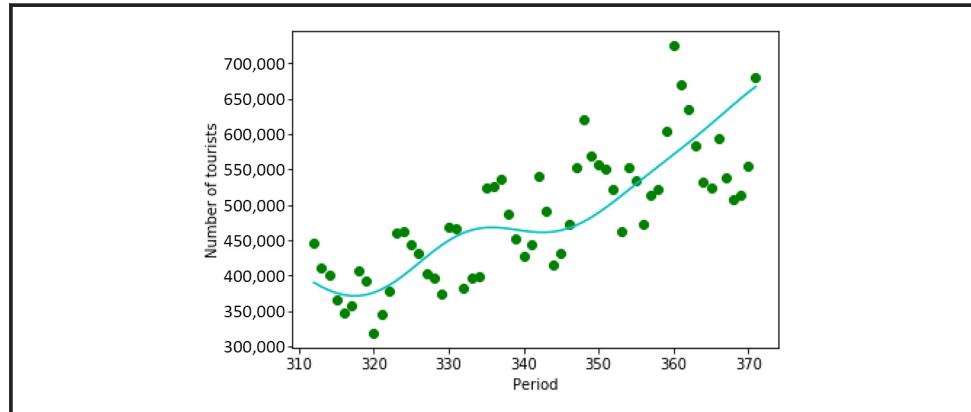
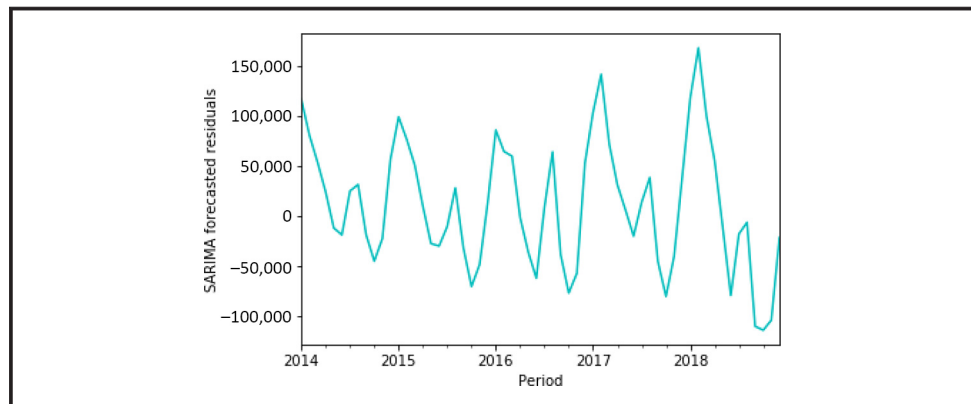


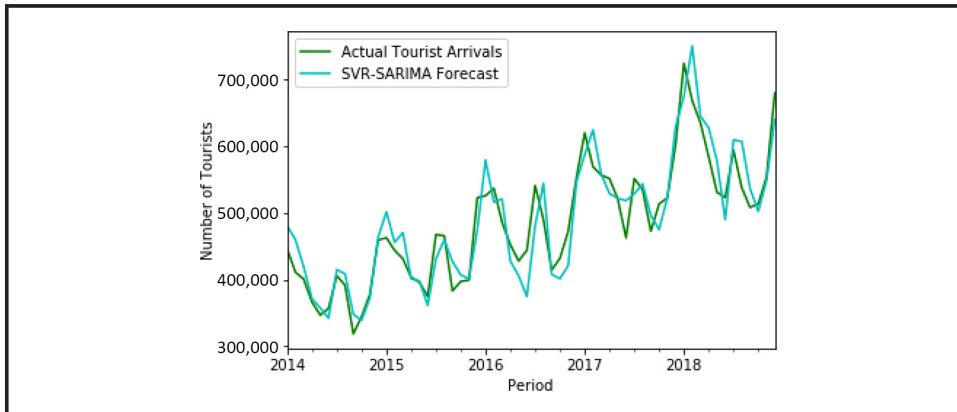
Table 1 Seasonal autoregressive integrated moving averages summary statistics

Model parameters	Coefficients	Standard error	z-score	p > z
AR lag_1	0.6554	0.153	4.270	0.000
AR lag_2	0.1065	0.135	0.788	0.431
Seasonal AR lag_12	0.5030	0.129	3.907	0.000
Seasonal AR lag_24	0.3237	0.131	2.468	0.014
Variance	3.60e+08	8.98e-11	4.02e+18	0.000

Figure 5 SVR forecast**Figure 6** SARIMA residuals forecast

Moreover, its performance is compared to benchmark models in literature such as the NAÏVE I, seasonal NAÏVE I and ES model (Athanasopoulos and de Silva, 2012; Rice *et al.*, 2019; Gunter and Onder, 2016). The Holt-Winters (HW) method is used as the ES benchmark model due to its capability to capture the trend and seasonality of a time series (Koehler *et al.*, 2001; Jiang *et al.*, 2020). Expert decision-making was also used to judge if an additive or multiplicative trend (and seasonality) will be used in the HW method. As a result, a multiplicative trend and an additive seasonality were used.

Figure 7 SVR-SARIMA forecast vs actual tourist arrivals



Moreover, the smoothing parameters of the HW method were selected using a hyper-optimization technique implemented in the python programming language. A comprehensive discussion on the HW method is provided by [Koehler et al. \(2001\)](#) and [Jiang et al. \(2020\)](#), among others. To determine the improvement provided by the hybridization approach, the SVR-SARIMA is also compared to the individual SVR and SARIMA models. The performance criteria used are the MAPE, RMSPE, MSE, RMSE, MAE, directionality change error (DCE), and trend change error (TCE). Such measures were selected to obtain a holistic evaluation of the proposed model being judged according to magnitude, directionality and trend change. The performance of each model is tabulated in [Table 2](#). With the presence of several criteria to evaluate the models, the decision to select the best model among the alternatives becomes a complex decision-making problem. To objectively select the best model relative to the determined performance criteria, PROMETHEE II is adopted. The weights of the performance criteria are decided to be uniform by the expert decision-makers due to the equal priority given to such weights. The final ranking of the forecasting models using the PROMETHEE II is presented in [Table 3](#). As such, the SVR-SARIMA model outranks all other models considered in this paper. Hence, the SVR-SARIMA model is considered the best model among the alternatives.

Table 2 Performance of each model

	NB	NB	NB	NB	NB	B	B
Weights	0.1428	0.1428	0.1428	0.1428	0.1428	0.1428	0.1428
Model	MAGNITUDE				DIRECTIONALITY		
	MAPE (%)	RMSPE (%)	MSE	RMSE	MAE	DCE (%)	TCE (%)
SVR-SARIMA	5.45	6.76	1139044456.87	33749.73	26873.71	73.33	25
SARIMA	27.88	28.14	20190652287.61	142093.82	136877.68	93.33	86.84
SVR	10.15	12.02	3772424888.83	61420.07	50045.23	53.33	46.15
NAÏVE I	8.40	10.41	2772371956.25	52653.32	41509.42	50	22.03
Seasonal Naïve I	8.50	9.97	2693419436.32	51898.16	43019.38	50	47.5
Holt–Winters method	16.30	18.67	11067576807.15	105202.55	86591.71553	78.33	51.72
$(\alpha = 0.035,$ $\beta = 0.011,$ $\gamma = 0.237)$							
Note: NB denotes non-beneficial criterion while B pertain to beneficial criterion							

Table 3 Preference ranking organization method for enrichment of evaluation II outranking results

	<i>Leaving flow</i>	<i>Entering flow</i>	<i>Difference</i>	<i>Rank</i>
SVR-SARIMA	0.337	0.075	0.262	1
SARIMA	0.228	0.565	-0.337	6
SVR	0.186	0.128	0.057	3
NAÏVE I	0.193	0.150	0.043	4
Seasonal Naïve I	0.216	0.104	0.112	2
Holt-Winters method	0.150	0.287	-0.136	5

5. Discussion and managerial implications

The development of hybrid forecasting models is a contemporary research area in the tourism demand forecasting literature, as pointed out by [Song and Li \(2008\)](#) and [Ghalekhondabi et al. \(2019\)](#). Collectively, results in the literature point toward the success of hybrid models in increasing forecast accuracy. In this paper, the SVR is hybridized with the SARIMA model to handle both linear and nonlinear components of the tourism demand time-series data. Consistent with the results found in the literature, the hybridization of linear and nonlinear models (i.e. in this case, the SVR-SARIMA model) generates more accurate forecasts than linear (i.e. SARIMA) and nonlinear (i.e. SVR) models treated separately. As can be observed in [Table 3](#), the proposed hybrid model outranks all other models tested in this paper.

Moreover, several interesting points can be drawn from the results. First, although the SVR outranked the NAÏVE I and the HW method, the difference is not apparent without performing the outranking. Second, the SVR and SARIMA models are most of the time outperformed by the benchmark models. Third, the proposed hybrid SVR-SARIMA model significantly improved the individual forecasts of the SVR and SARIMA models. Fourth, although the SARIMA model underperformed under the criteria of magnitude, it outperforms all other models in terms of directionality and trend change.

The small difference in MAPE of the proposed model with the NAÏVE models is driven by the lack of hyperparameter tuning of the SVR models. As found in several works in literature such as [Hasanipanah et al. \(2017\)](#) and [Laref et al. \(2019\)](#), the performance of SVR models heavily depended on the choice of hyperparameters, hence making the optimization process relevant. An important point to be acknowledged, however, is that the proposed hybridization approach outperformed the benchmark models even without performing a hyperparameter optimization, which demonstrates the inherent improvement of the hybridization approach. Moreover, a significant decrease in errors has been attained by the hybridization approach relative to its component models (i.e. SVR and SARIMA). The significant improvement attained by the proposed hybridization approach is likely due to significant individual contributions of the SVR and SARIMA models. For one, the ability of the hybrid model to decrease the magnitude errors could be inherited from the SVR model, which is very suitable in fitting nonlinear data such as the trend exhibited by tourism demand. The proposed model's sensitivity to directionality and trend change is inherited from SARIMA, which was found to be very responsive to directionality and trend changes.

Moreover, the results obtained in this paper demonstrates the significance of developing hybrid models. Such inference is manifested by the underperformance of the separate SVR and SARIMA models relative to the benchmark models and the outranking of the hybrid SVR-SARIMA model relative to the benchmark models. Hence, consistent with some results in the literature, hybridization significantly improves forecasting accuracy. Another significant finding that arrived in this paper is the relevance of the MCDM approach in selecting the best forecasting model. With the absence of statistical testing, such finding provides significant insights to the literature with regards to model selection in forecasting.

The approach adopted in this paper in selecting the best model is novel, as this has not been attempted in the forecasting literature and in tourism demand forecasting. The approaches taken in this paper significantly demonstrate the essence and importance of tourism demand forecasting. Tourism demand forecasting is at the core of tourism management. One of the widely known reasons why forecasting is especially useful in the tourism sector is that tourism products and services are highly perishable. Transportation is a critical component of tourism, which may enhance or undermine the tourist experience. Inadequate transport resulting from poor demand forecasting limits tourists to access attractions, which could be detrimental to stakeholders' businesses. Such illustrates the unfulfilled demand along with opportunity loss associated with them. On the other hand, unsold seats, vacant rooms in a hotel, unsold admissions in theme parks, along with the associated lost revenue also illustrates how products and services in tourism, are perishable since these are then considered unsold "inventory."

Another reason for the need for forecasting tourist arrivals is that people are inseparable from the production-consumption process. In most cases, the provision of the products and services in tourism occurs at the same time as consumption or interaction with the people providing it. Thus, determining the number of service crews in a fast-food chain, hotel staff, flight attendants, tourist guides and attendants in resorts ensures that there is enough available personnel whenever tourists or visitors need them. Furthermore, forecasts can help stakeholders identify and ensure the availability of other complementary services, which could augment tourist satisfaction. For example, a specific resort may be able to recognize that tourist arrivals in their facility depend on several factors such as the quality of recreational activities offered, the type of food served, or the availability of transport services. With this information, the resort would be able to foresee which additional services would be made available for the tourist's future needs. The ability to forecast events may be human-made, or natural disasters can also help minimize the disastrous effect of loss in sales, income, employment, and damages on any tourism facility. Lastly, decisions in investing infrastructures, equipment, airports, roads and bridges much depend on the accuracy of the forecasts. Considering that these are long-term investments, future demand must be correctly anticipated.

6. Conclusion and future works

The hybridization of several forecasting models has achieved success in literature in enhancing the accuracy of forecasting models. This paper proposes the use of a hybrid SVR-SARIMA model in forecasting tourism demand due to its capability to handle the linear, nonlinear and seasonal components of the tourism demand data. This paper finds that in the case when the benchmark models outperform such non-hybrid models (i.e. individual SVR and SARIMA models), their hybridization significantly enhances the forecasting accuracy and outperforms the accuracy of the benchmark models. Such results are consistent with those found in literature, which suggest that hybridization approaches of forecasting models, particularly, linear and non-linear models outperform their non-hybrid (i.e. individual models) counterparts.

Furthermore, this paper also proposes the use of MCDM approaches in selecting the best forecasting model. The PROMETHEE II method is used to rank the forecasting models using a preference function. With such an approach, the proposed hybrid SVR-SARIMA model outranked all other considered forecasting models. This paper provides significant contributions to the current literature. First, it investigates the capability of the SVR-SARIMA in generating highly accurate forecasts. Second, it shows how MCDM methods are used to select the best forecast model. Third, it offers insights into the dynamics of tourism demand from the Philippine viewpoint, which in the literature is comparatively underrepresented. Also, the findings of this paper benefit several stakeholders, including tourism sectors,

managers, government entities and policymakers, in that it offers them with a basis for creating a hybrid tourism forecasting technique that accurately models tourism demand.

This paper, however, only used an MCDM method without comparing it to other MCDM approaches for brevity. Moreover, a hyperparameter optimization approach has not been conducted to tune the proposed hybrid model. As such, for future works, several MCDM approaches could be compared to determine their applicability in the forecasting domain further. Likewise, a hyperparameter optimization can be conducted to determine the further improve the accuracy of the proposed hybrid model.

References

- Akin, M. (2015), "A novel approach to model selection in tourism demand modeling", *Tourism Management*, Vol. 48 No. 1, pp. 64-72.
- Alwee, R., Siti Mariyam, H. and Sallehuddin, R. (2013), "Hybrid support vector regression and autoregressive integrated moving average models improved by particle swarm optimization for property crime rates forecasting with economic indicators", *The Scientific World Journal*, Vol. 2013, p. 11.
- Armstrong, J. (2007), "Significance tests harm progress in forecasting", *International Journal of Forecasting*, Vol. 23 No. 2, pp. 321-327.
- Assadzadeh, A. and Mamipour, S. (2013), "Estimation of tourism demand function for Australian destinations: a time-varying parameter approach", *Anatolia*, Vol. 25 No. 1, pp. 108-116.
- Assaf, A., Li, G., Song, H. and Tsionas, M. (2018), "Modeling and forecasting regional tourism demand using Bayesian global vector autoregressive (BGVAR) model", *Journal of Travel Research*, Vol. 58 No. 3, pp. 383-397.
- Athanasopoulos, G. and de Silva, A., (2012), "Multivariate exponential smoothing for forecasting tourist arrivals", *Journal of Travel Research*, Vol. 51 No. 5, pp. 640-652.
- Athawale, V. and Chakraborty, S. (2010a), "Facility location selection using PROMETHEE II method", Proceedings of the 2010 international conference on industrial engineering and operations management, January, pp. 9-10.
- Athawale, V. and Chakraborty, S., (2010b), "Facility layout using PROMETHEE II method", *IUP Journal of Operations Management*, Vol. 9 Nos 1/2, pp. 81-98.
- Baldigara, T., (2015), "Modelling international tourism demand using seasonal ARIMA models", *Tourism and Hospitality Management*, Vol. 21 No. 1, pp. 19-31.
- Borhan, N. and Arsad, Z. (2014), *Forecasting International Tourism Demand from the US, Japan, and South Korea to Malaysia: A SARIMA Approach*, AIP Publishing LLC, Penang.
- Boser, B., Guyon, I. and Vapnik, V. (1992), "A training algorithm for optimal margin classifier", *Proceedings of the fifth annual workshop on Computational learning theory, 27-29 July*, pp. 144-152.
- Bosupeng, M. (2019), "Forecasting tourism demand: the hamiltonian filter", *Annals of Tourism Research*, Vol. 79, p. 102823.
- Box, G. and Jenkins, G. (1970), *Time Series Analysis Forecasting and Control*, WISCONSIN UNIV MADISON DEPT OF STATISTICS, p. I.
- Brans, J., Vincke, P. and Mareschal, B., (1986), "How to select and how to rank projects: the PROMETHEE method", *European Journal of Operational Research*, Vol. 24 No. 2, pp. 228-238.
- Burger, C., Dohnal, M., Kathrada, M. and Law, R., (2001), "A practitioner's guide to time series methods for tourism demand forecasting – a case study of durban", *South Africa. Tourism Management*, Vol. 22 No. 4, pp. 403-409.
- Cai, Z., Lu, S., and Zhang, X. (2009), *Tourism Demand Forecasting by Support Vector Regression and Genetic Algorithm*, IEEE, Beijing, pp. 144-146.
- Chen, K., (2011), "Combining linear and nonlinear model in forecasting tourism demand", *Expert Systems with Application*, Vol. 38 No. 8, pp. 10368-10376.
- Constantino, H.A., Fernandes, P.O. and Teixeira, J., (2016), "Tourism demand modelling and forecasting with artificial neural network models: the Mozambique case study", *Tekhne*, Vol. 14 No. 2, pp. 113-124.

- Fang, J. Su, H. and Xiao, Y. (2018), "Will artificial intelligence surpass human intelligence?", SSRN Electronic Journal.
- Ghalekhondabi, I., Ardjmand, E., Young, W.A. and Weckman, G.R., (2019), "A review of demand forecasting models and methodological developments within tourism and passenger transportation industry", *Journal of Tourism Futures*, Vol. 5 No. 1, pp. 75-93.
- Goh, C. and Law, R., (2002), "Modeling and forecasting tourism demand arrivals with stochastic nonstationary seasonality and intervention", *Tourism Management*, Vol. 23 No. 5, pp. 499-510.
- Gunter, U. and Onder, I., (2015), "Forecasting international city tourism demand for Paris: accuracy of uni-and multivariate models employing monthly data", *Tourism Management*, Vol. 46, pp. 123-135.
- Gunter, U. and Onder, I., (2016), "Forecasting city arrivals with google analytics", *Annals of Tourism Research*, Vol. 61, pp. 199-212.
- Gunter, U., (2018), "Conditional forecasts of tourism exports and tourism export prices of the EU-15 within a global vector autoregression framework", *Journal of Tourism Futures*, Vol. 4 No. 2, pp. 121-138.
- Hasanipanah, M., Shahnazari, A., Amnieh, H. and Armaghani, D., (2017), "Prediction of air-overpressure caused by mine blasting using a new hybrid PSO-SVR model", *Engineering with Computers*, Vol. 33 No. 1, pp. 23-31.
- Hassani, H., SirimalSilva, E., Antonakakis, N., Filis, G. and Gupta, R., (2017), "Forecasting accuracy evaluation of tourist arrivals", *Annals of Tourism Research*, Vol. 63 No. 1, pp. 112-117.
- Hu, Y., Jiang, P. and Lee, P., (2018), "Forecasting tourism demand by incorporating neural networks into grey-Markov models", *Journal of the Operational Research Society*, Vol. 70 No. 1.
- Jiang, W., Wu, X., Gong, Y., Yu, W. and Zhong, X., (2020), "Holt-winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption", *Energy*, Vol. 193, p. 116779.
- Jiao, E. and Chen, J., (2018), "Tourism forecasting: a review of methodological developments over the last decade", *Tourism Economics*, Vol. 25 No. 3, pp. 1-14.
- Jorge-Gonzalez, E., Gonzalez-Davila, E., Martin-Rivero, R. and Lorenzo-Diaz, D., (2019), "Univariate and multivariate forecasting of tourism demand using state-space models", *Tourism Economics*, doi: [10.1177/1354816619857641](https://doi.org/10.1177/1354816619857641).
- Koehler, A.B., Snyder, R.D. and Ord, J.K., (2001), "Forecasting models and prediction intervals for the multiplicative holt-winters method", *International Journal of Forecasting*, Vol. 17 No. 2, pp. 269-286.
- Kostenko, A.V. and Hyndman, R. (2008), "Forecasting without significance test?", manuscript, Monash University.
- Kourentzes, N. and Athanasopoulos, G., (2019), "Cross-temporal coherent forecasts for Australian tourism", *Annals of Tourism Research*, Vol. 75, pp. 393-409.
- Laref, R., Losson, E., Sava, A. and Siadat, M., (2019), "On the optimization of the support vector machine regression hyperparameters setting for gas sensors array applications", *Chemometrics and Intelligent Laboratory Systems*, Vol. 184 No. 1, pp. 22-27.
- Law, R., Li, G., Fong, D. and Han, X., (2019), "Tourism demand forecasting: a deep learning approach", *Annals of Tourism Research*, Vol. 75 No. 1, pp. 410-423.
- Lee, N., Shim, J., Ju, Y. and Seok-Cheon, P., (2017), "Design and implementation of the SARIMA-SVR time series analysis algorithm for the improvement of atmospheric environment forecast accuracy", *Soft Computing*, Vol. 22 No. 13, pp. 4275-4281.
- Li, G., Song, H. and Witt, S., (2005), "Recent developments in econometric modeling and forecasting", *Journal of Travel Research*, Vol. 44 No. 1, pp. 82-99.
- Li, G., Wu, D., Zhou, M. and Liu, A., (2019), "The combination of interval forecasts in tourism", *Annals of Tourism Research*, Vol. 75, pp. 363-378.
- Lim, C., (1997a), "Review of international tourism demand models", *Annals of Tourism Research*, Vol. 24 No. 4, pp. 835-849.
- Lim, C., (1997b), "An econometric classification and review of international tourism demand models", *Tourism Economics*, Vol. 3 No. 1, pp. 69-81.
- Lin, V., Liu, A. and Song, H., (2015), "Modeling and forecasting Chinese outbound tourism: an econometric approach", *Journal of Travel and Tourism Marketing*, Vol. 32 Nos 1/2, pp. 34-49.

- Liu, H., Chang, L., Li, C. and Yang, C., (2018), "Particle swarm optimization-based support vector regression for tourist arrivals forecasting", *Computational Intelligence and Neuroscience*, Vol. 2018.
- Lopes, I., Soares, F., and Silva, E. (2017), *Tourism Demand in the Algarve Region: Evolution and Forecast Using SVARMA Models*, AIP Publishing, Rome.
- Mariano, R. (2004), "Chapter thirteen: testing forecast accuracy", *A Companion to Economic Forecasting*, Blackwell Publishing, p. 1.
- Mullen, J. (2013), "CNN", available at: <http://edition.cnn.com/2013/11/07/world/asia/philippines-typhoon-haiyan/index.html> (accessed 30 December 2019).
- Pai, P., Huo, K. and Lin, K., (2014), "Tourism demand forecasting using novel hybrid system", *Expert Systems with Applications*, Vol. 41 No. 8, pp. 3691-3702.
- Pai, P., Wei-Chiang, H., Ping-Teng, C. and Chen-Tung, C., (2006), "The application of support vector machines to forecast tourist arrivals in Barbados: an empirical study", *International Journal of Management*, Vol. 23 No. 2, pp. 375-385.
- Panahi, H., Mamipour, S. and Nazari, K., (2015), "Tourism and economic growth: a time-varying parameter approach", *Anatolia*, Vol. 26 No. 2, pp. 173-185.
- Petrevska, B., (2017), "Predicting tourism demand by A.R.I.M.A. models", *Economic Research-Ekonomska Istraživanja*, Vol. 30 No. 1, pp. 939-950.
- Pokharel, R., Poudel, J. and Grala, R., (2018), "Impact of economic factors and a political conflict on international tourism demand in Nepal: a vector error correction model", *International Journal of Tourism Policy*, Vol. 8 No. 1, pp. 18-41.
- PSA (2018), *Contribution of Tourism to the Economy is 12.2 Percent in 2017*, Philippine Statistics Authority, Manila.
- Rice, W., Park, S., Pan, B. and Newman, P., (2019), "Forecasting campground demand in US national parks", *Annals of Tourism Research*, Vol. 75, pp. 424-438.
- Rufino, C., (2011), "Forecasting international demand for Philippine tourism", *DLSU Business & Economics Review*, Vol. 20 No. 1, pp. 61-76.
- Rufino, C. (2016), *Forecasting Monthly Tourist from ASEAN+3 Countries to the Philippines for 2015-2016 Using SARIMA Noise Modeling*, DLSU, Manila.
- Ruiz-Aguilar, J., Turias, I., Jiménez-Come, M., and Cerbán, M. (2014), "Hybrid approaches of support vector regression and SARIMA models to forecast the inspections volume", *Hybrid Artificial Intelligence Systems*, Springer, Cham, pp. 502-514.
- Sakhuja, S., Jain, V., Kumar, S., Chandra, C. and Ghildayal, S.K., (2016), "Genetic algorithm based fuzzy time series tourism demand forecast model", *Industrial Management & Data Systems*, Vol. 116 No. 3, pp. 483-507.
- Sen, D., Datta, S., Patel, S. and Mahapatra, S., (2015), "Multi-criteria decision making towards selection of industrial robot: exploration of PROMETHEE II method", *Benchmarking: An International Journal*, Vol. 22 No. 3, pp. 465-487.
- Shen, M., Liu, H.H., Lien, Y.H., Lee, C.F., and Yang, C.H. (2019), *Hybrid Approach for Forecasting Tourist Arrivals*, ACM, Penang.
- Smola, A. and Schölkopf, B., (2004), "A tutorial on support vector regression", *Statistics and Computing*, Vol. 14 No. 3, pp. 199-222.
- Song, H. and Li, G., (2008), "Tourism demand modelling and forecasting – a recent review of recent research", *Tourism Management*, Vol. 29 No. 2, pp. 203-220.
- Song, H., Qiu, R. and Park, J., (2019), "A review of research on tourism demand forecasting: launching the annals of tourism research curated collection on tourism demand forecasting", *Annals of Tourism Research*, Vol. 75, pp. 338-362.
- Song, H., Wen, L. and Liu, C., (2019), "Density tourism demand forecasting revisited", *Annals of Tourism Research*, Vol. 75, pp. 379-392.
- Totanes, H. (1998), *A Timeline of Philippine History*, 10th ed., Vol. 1, Asia Publishing Company.
- Wen, L., Liu, C. and Song, H., (2019), "Forecasting tourism demand using search query data: a hybrid modelling approach", *Tourism Economics*, Vol. 25 No. 3, pp. 309-329.

Whaley, F. (2013), "Death toll in the Philippines earthquake rises to 144", *The New York Times*, 16 October.

Williams, R. (2013), "Typhoon haiyan: most powerful storm to ever hit land batters Philippines with 200 mph winds", *The Independent*, 8 November.

Witt, S. and Witt, C., (1991), "Tourism forecasting: error magnitude, direction of change error, and trend change error", *Journal of Travel Research*, Vol. 30 No. 2, pp. 26-33.

World Almanac Books (2007), *The World Almanac and Book of Facts 2007*, World Almanac Books, New York, NY.

Xu, S., Chan, H. and Zhang, T., (2019), "Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 122 No. 1, pp. 169-180.

Yang, Y. and Zhang, H., (2019), "Spatial-temporal forecasting of tourism demand", *Annals of Tourism Research*, Vol. 75, pp. 106-119.

Zou, J., Jiang, G.F., Xie, X.X., Huang, J. and Yang, X.B., (2019), "Application of a combined model with seasonal autoregressive integrated moving average and support vector regression in forecasting handfoot-mouth disease incidence in Wuhan", *China. Medicine*, Vol. 98 No. 6.

Further reading

Croce, V., (2016), "Can tourism confidence index improve tourism demand forecasts?", *Journal of Tourism Futures*, Vol. 2 No. 1, pp. 6-21.

Goodwin, P., (2010), "The holt-winters approach to exponential smoothing: 50 years old and going strong", *Foresight*, Vol. 19, pp. 30-33.

Kim, J. and Moosa, I., (2005), "Forecasting international tourist flows to Australia: a comparison between the direct and indirect methods", *Tourism Management*, Vol. 26 No. 1, pp. 69-78.

Author affiliations

Dharyll Prince Mariscal Abellana is based at the Department of Computer Science, University of the Philippines Cebu, Cebu City, Philippines and the Department of Industrial Engineering, Cebu Technological University, Cebu City, Philippines.

Donna Marie Canizares Rivero is based at the College of Technology, Cebu Technological University, Cebu City, Philippines.

Ma. Elena Aparente is based at the Department of Industrial Engineering, Cebu Technological University, Cebu City, Philippines.

Aries Rivero is based at Department of Industrial Engineering, Cebu Institute of Technology University, Cebu City, Philippines.

Corresponding author

Dharyll Prince Mariscal Abellana can be contacted at: dmabellana@up.edu.ph

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com