



Forecasting of Indian and foreign tourist arrivals to Himachal Pradesh using Decomposition, Box–Jenkins, and Holt–Winters exponential smoothing methods

Keerti Manisha^{1,2} · Inderpal Singh¹

Received: 13 February 2023 / Accepted: 26 June 2024 / Published online: 23 July 2024
© The Japan Section of the Regional Science Association International 2024

Abstract

Himachal Pradesh offers many tourist experiences to Indian and foreign visitors. The mountainous state experiences a high degree of tourism seasonality, interrupting the efficient operation of the tourism infrastructure. Fundamental requirements for tourism planning are accurate projections of tourist demands. However, estimating future demands becomes challenging because of the high degree of seasonality. The compound and annual growth rate methods are used in state-level tourism research to forecast future growth. Since these models cannot manage seasonality and trends in data series, they inaccurately predict future demand. In this context, forecasting models such as the Decomposition, Box–Jenkins, and Holt–Winters exponential smoothing methods were used to forecast the seasonal tourism demands in the study area. The dataset utilized for the analysis was the monthly Indian and foreign tourist arrivals from 2008 to 2018. Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Theil's U1 coefficients validated that the forecast models produced accurate results. The Box–Jenkins model accurately forecasted the tourist arrivals (2019 to 2031) as reflected by the lowest error metrics (Indian: RMSE 36833.8, MAPE 3.0, Theil's U1 0.023; Foreign: RMSE 6907.59, MAPE 15.51, Theil's U1 0.10). This approach outperformed the traditional seasonal data series forecasting techniques and contributed to the literature on univariate tourist demand forecasting for hilly areas experiencing a high degree of seasonality. Estimating maximum tourist arrivals is crucial in long-term strategic planning for tourism expansion to withstand maximum loads and ensure efficient business flow, higher investment, enhanced economic growth, and environmental protection in the state. This study is a pioneer in examining tourism demands for stakeholders, tourism operators, and planners to plan the fluctuating tourism in Himachal Pradesh while adhering to sustainability principles. Furthermore, it provides inputs for effective planning and policy formulations specific to the tourism industry based on future demand assessments.

Keywords Demand · Forecasting · Seasonality · Tourism · Tourist arrivals

Extended author information available on the last page of the article

1 Introduction

Globally, tourism has evolved as the fastest-growing industry (Reivan-Ortiz et al. 2023), offered opportunities for sustainable development (Mustafa et al. 2021), generated employment, and alleviated poverty (Peiris 2016). In 2017, tourism contributed to around 10% GDP of the world (Duro and Turrión-Prats 2019), and international tourism receipts increased to US\$1340 billion in 2017 from US\$2 billion in 2008, as per the United Nations World Tourism Organization (UNWTO). “Travel and tourism percentage of global GDP for 2018, 2019, 2020, 2021 was 9.2%, 6.8%, 5.3%, 6.1%”, respectively (WTTC 2022). As WTTC reported, currently tourism is affected by seasonality which causes overcrowding, escalated prices, inadequacy in infrastructure, unsatisfactory hospitality services, and job loss during shoulder and lean seasons in a destination (Duro and Turrión-Prats 2019). Seasonality is studied concerning tourist arrival numbers with respect to tourist seasons, behaviour, preferences, and revenue generation (Choe et al. 2019). Seasonality relates to climatic conditions, institutional holidays, cultural events, and destination attractiveness (UNWTO 2004). Seasonality in tourism is a constraint in achieving sustainability, efficiency, and regional productivity (Wang et al. 2023). The specific indicators to measure the degree of seasonality include tourist arrival numbers, occupancy rate, and linked issues such as unemployment, seasonal opportunities, poor quality of services, and wear-tear of tourism infrastructure during the lean season.

Typically, tourism depends on seasonal patterns such as climatic conditions, socioeconomic activities, and consumption patterns (Ćorluka 2019). The presence of seasonality negatively impacts the equal distribution of tourists in a destination (Jangra and Kaushik 2018). Furthermore, tourism destinations that depend on natural and climate factors are more vulnerable to seasonal fluctuations than destinations depending on cultural factors, such as urban destinations. Although seasonality is a crucial tourism characteristic (Higham and Hinch 2002), it is primarily a disgrace because of migration, unemployment, low-income rate, and disinvestment (Martín et al. 2020). The growth in the tourism sector relies on the seasonality patterns which condition the market growth (Lee and Huang 2007). In numerous studies, seasonality has challenged planners and policymakers in building a sustainable tourism economy (Martín Martín et al. 2020).

India offers a wide range of tourist attractions, such as adventure, heritage, wildlife, and pilgrimage (Kumar and Singh 2019). The earnings from foreign tourism exchange in 2019 grew by 5.1% (US\$ 30.06 billion) from US\$ 28.59 billion in 2018. In 2019, foreign tourist arrivals increased to 10.93 million from 10.56 million in 2018 (Ministry of Tourism 2023). India has recovered about 60% of its pre-pandemic level, showing a growth of 321.54% from 2021 as reported by the Ministry of Tourism. Tourism in India will contribute 7.1% more to the GDP by 2033 (WTTC 2023). However, the tourism economy in the country gives birth to issues such as employment, transportation, the hotel industry, restaurants, resorts, and allied infrastructure (Kumar and Singh 2019). Also, tourism in the Indian Himalayan Region (IHR) is a predominantly religious type and later

changed to summer escapes during the nineteenth century (NITI Aayog 2018). The region experiences a significant degree of seasonality due to the weather pattern (NITI Aayog 2018) and faces the negative impacts of traditional architecture, poor infrastructure, pollution, resource depletion, and damage to biodiversity and ecosystem services (Chetty and Manisha 2022; Manisha et al. 2023).

Himachal Pradesh in the northern IHR embraces immense natural beauty through snow-covered mountains, meandering rivers, forests, flora, climatic conditions, and deep valleys. The GSDP contribution of Himachal Pradesh was at its lowest (0.73%) in 2022 among other northern Himalayan states (Statistics Times, 2022). Tourism in the state generated foreign exchange revenues and direct and indirect employment, and registered a growth of 8.16% between 2015–26 and 2023–24 (IBEF 2023). The fragile ecosystem of the state is adversely affected due to continuous growth, increased demand for exclusive tourism, and intensive infrastructure development (Badar and Bahadure 2020; Kumar et al. 2023). However, the scientific literature on analyzing the future tourism trend in the study area is scarce. In this context, there is a need to accurately forecast the tourist arrival numbers using scientific methods to enable effective tourism planning and mitigate the negative impacts of tourism practices on the environment. Accurately estimating tourist arrivals can guide planners and architects toward regional sustainability.

Himachal Pradesh, on the lap of the north-western Himalayas, is between $30^{\circ}22'$ and $33^{\circ}12'$ north latitude and $75^{\circ}47'$ and $79^{\circ}4'$ east longitude (Fig. 1). The state covers an area of 55,673 sq. km. and accommodates a total population of 6.08 million (Planning Commission 2011). The altitude variation of the state ranges from 190 to

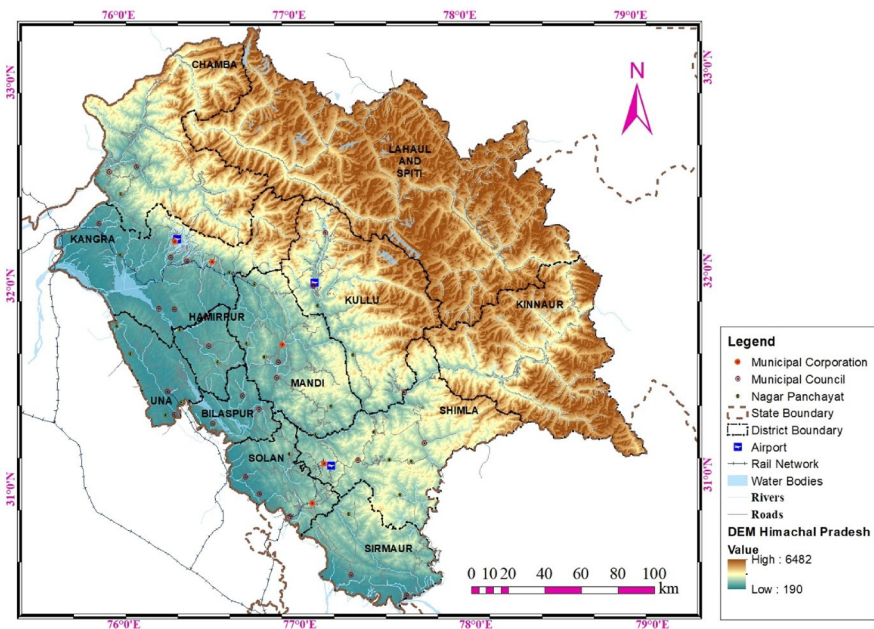


Fig. 1 Locational map of the study area

6482 m above mean sea level. The climate zone predominantly varies from semi-tropical to semi-arctic based on the altitude. The state witnesses three main seasons which directly impact its economic growth; rainy (July–September), Winter (October–March), and Summer (April–June). The state has five perennial water sources: Sutlej, Beas, Ravi, Chenab, and Yamuna. Due to undulating terrain and physiography in the state, the existing natural drainage system is challenging to exploit for irrigation. The productive period of the state depends on its unique geographic and climatic conditions (Planning Commission 2011).

The administrative divisions in the state comprise 12 districts (Fig. 1) and reached 16.09 million domestic tourists and 356,000 foreign tourists in 2018 (Pradesh 2020). However, hotels remained vacant during the peak months in 2019 due to the COVID-19 pandemic (April, May, and June) (Pradesh 2020). Tourism across the state got shut down because of the lockdown due to the coronavirus pandemic. As an implication of the COVID-19 pandemic, as per the Federation of Indian Chambers of Commerce & Industry, the hotel and allied service sectors remained closed since March 2020. Numerous local employees engaged in the hotel industry have returned to their place of origin, while many got stuck at their workplaces with almost no business. Research on state-level forecasting has primarily considered geometric models like the compound or annual growth rate method. A weighted average method gets applied for district-level forecasting to forecast Indian and foreign tourist arrivals. Typically, these models utilize the data from the past to estimate the future. However, these models cannot estimate the seasonality, trend, and level in a dataset. The state is a mountainous destination with varying climatic and socio-economic conditions that invite seasonal tourism. The presence of seasonality needs to be captured accurately for future estimation. Thus, applying the scientific method to forecast the growing tourism demand in the study area is necessary.

The rapidly increasing tourist arrivals is undoubtedly the main reason behind the environmental degradation in the study area. Estimating the maximum tourist arrivals is crucial for long-term strategic planning to expand tourism and withstand the foreseen maximum loads. Thus, long-term tourism forecasting is pivotal in providing insights for policymakers' decisions, such as expanding tourism infrastructure and amenities as per sustainable development principles. Basically, there are natural and institutional reasons for the high degree of seasonality. Although natural reasons for seasonality cannot be regulated, institutional-level measures can solve the problem of seasonality through policy formulation (Rizal and Asokan 2014). Typically, long-term, or medium-term forecasting is mandatory, allowing decisions to be made well in advance to fulfil future demand. Numerous studies have highlighted the importance of long-term forecasting while assessing future demands. A lack of adequate assessment and strategic planning would be disastrous for the tourism industry, environment, and society (Fahey et al. 1981; Koerner et al. 2023). Himachal Pradesh state authorities also invest in preparing a master plan specific to tourism and tourism infrastructure development for the coming 15 years. These scientific and statistical tourism projections play a significant role in dictating the quantum and type of infrastructure development, prioritizing interventions for future growth, optimum allocation of funds, identifying destinations/areas that are likely to exceed the carrying capacity, assessing tourism potential of less explored destinations, devising

mobility plan for compelling tourism experiences, and assisting in the development of overall development strategy for a region.

This study aims to accurately forecast tourist arrivals in Himachal Pradesh using the monthly Indian and foreign tourist arrival data from 2008 to 2018, i.e., pre-pandemic decadal data. This study compared the selected models, such as decomposition, the Box–Jenkins (B–J), and Holt–Winters (H–W) exponential smoothing methods, and identified the best-performing method in forecasting tourism demand from 2019 to 2031. The Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Theil's U1 coefficients validated the forecasting accuracy of these models. Here, the novel approach is the comprehensive methodology explicitly designed for the hilly tourist areas experiencing a high degree of seasonality. This research makes a unique contribution to the literature on univariate tourist demand forecasting of future tourism demand by using the number of tourist arrivals in the region. This approach on accurate forecasting will ensure efficient business flow, higher investment, and enhanced tourism growth in the hilly state of Himachal Pradesh. The rest of this study is as follows: a literature review to present a brief on the relevant empirical literature, a description of the forecasting method adopted, data and research methodology to describe the detailed approach adopted, results and discussions to demonstrate key findings from the analysis, policy recommendations to suggest mitigation measures; and conclusion to present the final remarks.

2 Literature review

The ability to predict future demand is the foundation of tourism planning (Li et al. 2020). The irregularity in tourism gets analyzed through seasonality (Mohd Lip et al. 2020). The economy, ecology, and society get strongly impacted because of seasonality in a tourism destination (Song et al. 2011; Vergori 2012; Chen et al. 2019; Duro and Turrión-Prats 2019; Manisha et al. 2023). The management of seasonality is used as a method to mitigate the adverse effects of peak season activities on the local economy and the environment. Seasonality in tourism affects how accurately a model can forecast travel demand (Chen et al. 2015; Wang et al. 2023). Generally, tourism forecasting involves both qualitative and quantitative techniques, including time series models (Yin 2020).

It has been noted from the literature review that numerous techniques adopted in tourism forecasting are the time-varying parameter methods (Song et al. 2011; Liu et al. 2024), multi-series structural time series (Chen et al. 2019), hybrid SARIMA-Long Short-Term Memory (LSTM) approach to forecast daily arrivals (Wu et al. 2021), visiting probability model to forecast tourist volume (Wu et al. 2019), hybrid SARIMA-ANN (Artificial Neural Network) model (Yollanda and Devianto 2020), machine learning and internet search index (Sun et al. 2019), Seasonal Index Method (Rizal and Asokan 2014), weighted fuzzy integrated time series (Suhartono and Javedani 2011), classical forecasting model (Intarapak et al. 2022), forecasting demand (Joyeux et al. 2012), SARIMA-MIDAS (moving average-mixed data sampling) (Wu et al. 2023), Markov switching approach (Botha and Saayman 2022), Holt's Weighted Exponential Moving Average (H-WEMA) (Hansun et al. 2019),

Multivariate Exponential Smoothing (Jiang 2023), Theta model and the new forecast Hybrid (Kolková and Rozehnal 2022), payment card data and Google Search indices (Crispino and Mariani 2024), and Facebook Prophet (Elseidi 2023). Seasonal Autoregressive Integrated Moving Average (SARIMA) and Autoregressive Integrated Moving Average (ARIMA) models are used to investigate the long-term forecasting of seasonal runoff (Valipour 2015). SARIMA model helps in modelling tourist forecasts in; the Chilean region (Brida and Garrido 2011), Malaysia (Mohd Lip et al. 2020), Sri Lanka (Peiris 2016; Thushara et al. 2019), and Spanish port (Grifoll et al. 2021).

In addition to the plethora of models, there is a grey-based model to forecast international tourism (Pirthee 2017), multi-horizon accommodation demand forecasting in New Zealand (Zhu et al. 2021), forecasting accuracy evaluation (Hassani et al. 2017). A comparative analysis between the models determines the most suitable forecasting model (Cho 2003). The forecasting models have developed over the past 50 years to enable research on efficiently and accurately forecasting tourism demand (Song et al. 2019). The most frequently considered models for comparative analysis are; SARIMA, Artificial Neural Network (ANN), and hybrid models (Aslanargun et al. 2007), WEMA and H-WEMA (Hansun et al. 2019), SARIMA and Holt-Winter method (Athanasopoulos and de Silva 2012; Hansun et al. 2019). The monthly annual tourist data helps estimate tourism seasonality and its impact on trend forecasting (Vergori 2012). Determining whether it is necessary to estimate seasonality within a stochastic model or eliminate the seasonal component from the time series is vital. SARIMA model can be used to analyze the time series, and the seasonal adjustment procedure (TRAMO-SEATS) can be used to remove the seasonal component from the original dataset (Brida and Garrido 2011). This data modification makes it fit for auto-regressive integrated moving average (ARIMA) models. SARIMA effectively analyzes the time series of visitor arrivals to model seasonality (Vergori 2012). Further to this, the forecasting techniques include the Box–Jenkins method (Goh and Law 2002; Shen et al. 2009; Kriti 2016), the Naïve method (Athanasopoulos et al. 2011; Makoni and Chikobvu 2018), ARIMA (Ampountolas 2021; Mustafa et al. 2021), Artificial Neural Network (ANN) (Panigrahi et al. 2018), exponential smoothing (ETS) model (Ke et al. 2016; Gunter 2021), Seasonal and Trend Decomposition (STL) model (Ke et al. 2016), and Holt–Winters method (Athanasopoulos and de Silva 2012).

The advantage of using these scientific models is their ability to predict future demand utilizing historical data and patterns. Typically, the trend, pattern, and seasonality observations might be absent in the data series, hindering accurate projections. These scientific models can individually identify trends, patterns, and seasonality for a data series and provide a framework for decision-making based on data-driven findings. As a result, different scenarios and assumptions can be drawn from these models to allow more informed decision-making. This study applied the decomposition, the SARIMA, and H–W exponential smoothing methods to develop different future tourism scenarios in the study area. The decomposition model applies the theory that a time series variable comprises four components: trend, seasonal, cyclical, and irregular (Konarasinghe 2017). These components act additively or multiplicatively to form the time series variable. It is one of the oldest models,

built on a simple and intuitive conceptual theory. It provides a clear and meaningful interpretation of the model parameters to identify the trend and patterns in the data. The H–W model can forecast time series data for short-term and long-term periods (Alonso Brito et al. 2021). This model helps forecast data that exhibit trends and seasonality, such as sales, tourist arrivals, weather, stock prices, and a wide range of industries, including finance, tourism, retail, and manufacturing. SARIMA is a powerful and popular method to forecast time series for short-term and long-term data exhibiting seasonality (Wu et al. 2021). However, each method possesses a set of assumptions and limitations. Therefore, it is paramount to note that the accuracy of models depends on the quality and relevance of the data used to build them.

3 Forecasting methods adopted in this study

3.1 Decomposition method

The decomposition method divides the timeseries (Y_t) into the trend (T_t) and season (S_t) components. The method generally follows two approaches, additive $Y_t = T_t + S_t + \varepsilon_t$ and multiplicative $Y_t = T_t * S_t * \varepsilon_t$. Both approaches have a component to express the random error (ε_t). In Decomposition method, the trend is determined using a simple linear regression model, and the seasonal component is determined by detrending the time series dataset (Intarapak et al. 2022). The prerequisite conditions to apply this method in the time series analysis are: the data display a seasonal trend with varying amplitude across the pattern, the magnitude of the seasonal variation depends on the magnitude of the trend, and the data exhibit an increasing and decreasing trend over time. In this study, the data analysis revealed that the multiplicative model outperformed the additive model by capturing the seasonality component in the time series to enable long-term projection. Basically, the multiplicative decomposition model was applied to capture the seasonal fluctuations in the time series data, as the magnitude of the seasonal variation is proportional to the magnitude of the trend.

3.2 Box–Jenkins (B–J) method

The Box–Jenkins (B–J) method uses an iterative model-building strategy capable of handling stationary and non-stationary time series datasets with or without seasonal components (Intarapak et al. 2022). The data series goes through three stages; selection of the model (Identification), estimation of the parameter through coefficients (Estimation), checking the model by analyzing the residuals (Diagnostics), and forecasting. The selection of the model requires the determination of the order of ordinary differencing (d), seasonal differencing (D), non-seasonal and seasonal autoregressive component (p , P), and non-seasonal and seasonal moving average component (q , Q). Typically, the time series analysis is performed in statistical software like EViews, R- for Statistical Computing and Statistical Package for the Social Sciences (SPSS). The model checks for stationarity in the time series using

autocorrelation coefficient (ACF) and Partial autocorrelation coefficient (PACF) coefficients by counting the spikes in the plot. The next step is to check for the unit root in the time series by applying Augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests hypothesis. The original time series dataset gets modified in B–J method for modelling the time series. The dataset is non-stationary when it exhibits an increasing trend and a seasonal pattern. Then, the dataset is required to be transformed into stationarity using unit root tests like the ADF (Mohamed 2010) and KPSS (Ampountolas 2021). ADF and KPSS tests are the conjunction with the B–J method to identify and estimate ARIMA model. When the data is non-stationary, the difference of the time series removes the trend component.

“B–J and Seasonal Autoregressive Integrated Moving Average (SARIMA) are related but not replaceable” (Bolarinwa and Bolarinwa 2021). The B–J methodology identifies the fitted ARIMA model for a time series, including the order of differencing, autoregressive, and moving average components. This methodology can identify and estimate SARIMA and other time series models. SARIMA is a time series model that extends the ARIMA model to include the seasonal component. SARIMA effectively models the time series data and demonstrates seasonal patterns. In the SARIMA model $(p, d, q) (P, D, Q)_s$, s is the number of seasons per year, d is the order of ordinary differencing, and D is the seasonal order of differencing. The expression of the SARIMA model is; $\phi_p(B)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = \varnothing_0 + \omega_q(B)\Omega_Q(B^s)\varepsilon_t$ (Gujarati 2015).

There is a plethora of existing literature on the SARIMA for estimating tourist arrivals in Taiwan (Peiris 2016), Thailand (Intarapak et al. 2022), Chilean (Brida and Garrido 2011), Kenya (Lidiema 2017), and others. Previously, the model was applied to forecast the energy demand, changes in monthly rainfall patterns, rise in tuberculosis cases, and others. The SARIMA model applies the maximum likelihood estimation technique to determine the value of the parameters of the fitted model. The accuracy of the model got validated through minimum Akaike’s Information Criterion (AIC) and Bayesian Information Criterion (BIC) values. Thereafter, the fitted model is examined for white noise using the Ljung–Box test. The test confirms no significant autocorrelation in the residuals ($p > 0.05$) means that the residuals are independent and normally distributed in a fitted model.

3.3 Holt–Winters (H–W) method

The H–W method employs multiplicative or additive assumptions to forecast a long-term relationship with seasonal characteristics. The multiplicative assumption is applied when the size of seasonal variation increases with the time series. At the same time, an additive assumption occurs if the absolute size of the variation is unrelated. However, this study presented the analysis of the multiplicative method because of its capability to capture the seasonality in the time series and paucity of space. The forecasting equation of the multiplicative exponential smoothing method is expressed in Eq. 1. In this equation; there are components such as level component (L_t), trend component (T_t), seasonal component (S_t), number of tourists arrival

(Y_t), time period (t), and initial value of seasonal component (S_n). Here X_{t+p} is the forecast value for p -step ahead, $p=1, 2, 3 \dots n$. The α, β, γ are the level, trend, and seasonality components need calculation (Mohd Lip et al. 2020). The value of these components range between $0 \leq \alpha, \beta, \gamma \leq 1$ and p is the length of the forecasting period (Intarapak et al. 2022).

$$X_{t+p} = (L_t + pT_t)S_{t-s+p} \quad (1)$$

The Eq. 2 is the level component (L_t) of the H–W method,

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (2)$$

Equation 3 is the trend component (T_t) of the H–W method,

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (3)$$

Equation 4 is the seasonal component (S_t) of the H–W method,

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (4)$$

Equation 5 is the initial value of the level component (L_0),

$$L_0 = \frac{y_1 + y_2 + y_3 + \dots + y_n}{12} \quad (5)$$

Equation 6 is the initial value of trend component (bo) and s is the number of seasons.

$$bo = \frac{1}{s} \left(\frac{y_{s+1} - y_1}{s} \right) + \left(\frac{y_{s+2} - y_2}{s} \right) + \left(\frac{y_{s+3} - y_3}{s} \right) + \left(\frac{y_{s+4} - y_4}{s} \right) + \dots + \left(\frac{y_{s+n} - y_n}{s} \right). \quad (6)$$

Equation 7 is to calculate the initial value of seasonal component (S_n),

$$S_n = \frac{y_n}{L_0} \quad (7)$$

3.4 Validation of the models or forecasting accuracy

The selected models can forecast seasonal time series data with greater accuracy. The models are validated for fit using RMSE, MAPE, and Theil's U1 coefficients. RMSE is the square root of the mean error, which is the addition of the squared residuals per number of observations. The RMSE equation expresses the difference between the actual (A_t) and forecasted (F_t) values in period (t), and n is the count of observations (Eq. 8). The interpretation of MAPE states that there are four conditions to evaluate the accuracy of the forecast. Less than 10% in MAPE accuracy indicates precision, between 11 and 20% is a good forecast, 21–50% is a moderately

accurate forecast, and above 51% is an inaccurate forecast (Chen et al. 2003). The MAPE is calculated by dividing the absolute error for each period by the actual value observed during the period and taking the average of these absolute errors (Intarapak et al. 2022). MAPE can compare time series, methods, and different time intervals, as specified by Goh and Law (2002). Here, n represents the total number of periods considered, A_t is the actual values, and F_t is the forecasted values in period (t) (Eq. 9). Theil's U1 (Eq. 10) coefficient measures the forecast accuracy of the models (Goh and Law 2002). Theil's U1 statistics range from 0 to 1, where 0 indicates perfect equality between two data sets developed by the economist Henri Theil in 1967.

$$\text{RMSE} = \sqrt{\frac{\sum (A_t - F_t)^2}{n}} \quad (8)$$

$$\text{MAPE} = \frac{\sum |(A_t - F_t)/A_t|}{n} \quad (9)$$

$$\text{Theil's U1} = \frac{\sqrt{\sum (A_t - F_t)^2}}{\sqrt{\sum (A_t - A_{t-1})^2}}. \quad (10)$$

4 Data and research methodology

The tourist arrival data were collected from January 2008 to December 2021 from the Department of Tourism and Civil Aviation (DoTCA) in Shimla, Himachal Pradesh. The state witnessed a complete shutdown during the pandemic, which resulted in ingenuine tourist arrival data record between 2019 and 2021. Since the time series forecasting depends on the pattern and behaviour of the past records, the data pertaining the pandemic year (2019–2021) got eliminated. This elimination intended accurate long-term tourism forecasts outside the pandemic period, like the research of Makoni et al. (2021). Essentially, this study predicted future tourist arrivals based on pre-pandemic trends to show what tourist demand would have been without COVID-19. Nevertheless, the authors acknowledge that presenting the COVID-19 tourism scenario is vital and thus, recommend future research to employ advanced data mining techniques for assessment. The detailed research methodology adopted in this study is exhibited in Fig. 2.

The first step begins with collecting data from the DoTCA for the Indian and foreign tourist arrival numbers. The collected data is then converted into a time series object $Y(t)=X$ by assigning a unique identifier 2008M01, 2008M02...2018M12; where $Y(t)$ is the actual value of the time series at time t and represented in graph

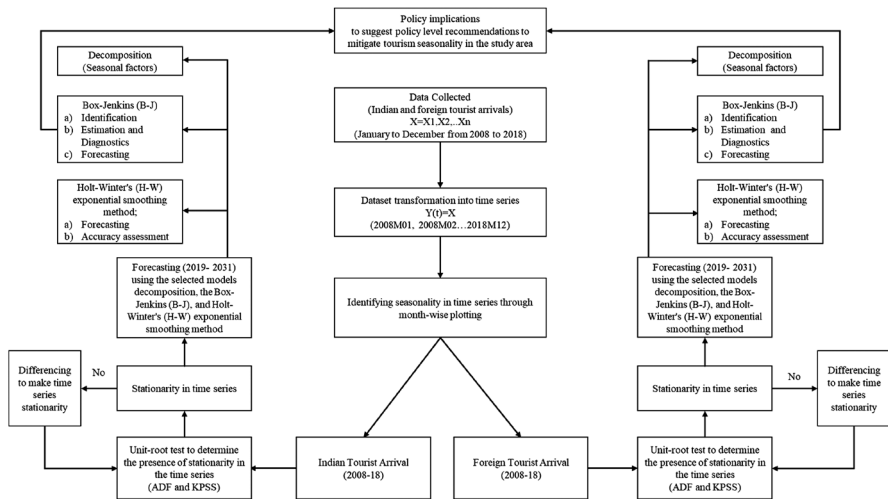


Fig. 2 Detailed research methodology adopted in this study

plots. Here, “ X ” (January to December from 2008 to 2018) represents the raw data entries that are not arranged by time into an orderly fashioned data series “ $Y(t)$.” In this step, the actual data entries get transformed into time series, wherein each entry gets labelled to a specific time. However, the actual data did not undergo a logarithmic transformation to prioritize the interpretability of results for stakeholders and decision-makers familiar with the original scale of the time series data. Several past studies, including Elamin and Fukushige (2018), Fatema et al. (2022), Intarapak et al. (2022), Kayral et al. (2023) have utilized actual data without logarithmic transformation. After that, the original time series (level) was utilized in standard techniques such as time series decomposition, autocorrelation analysis, and forecasting. Seasonal characteristics in the time series were detected using the unit root tests like ADF and KPSS. Afterward, the suitable lag order was identified in the level and differenced time series. If the test confirms the stationarity the model gets selected otherwise, the time series is differenced to make it stationary. In case of unsatisfactory results, the process of lag selection repeated to select a suitable model. In this study, methods selected for forecasting tourist arrivals are the decomposition, B–J, and H–W exponential smoothing models. The L-Jung test was applied to check the white noise in the dataset (Sun et al. 2019). These models can effectively analyze and forecast the tourist arrival time series in this study. The forecasting accuracy of the models chosen was achieved through RMSE, MAPE, and Theil’s $U1$ coefficients by comparing the actual and forecasted tourist arrivals. This comparative analysis determined the forecasting performance of the models both within the training sample and the estimation sample, as by Elamin and Fukushige (2018). Numerous studies have supported comparing the accuracy of forecasting techniques against the actual data, such as Fatema et al. (2022), Intarapak et al. (2022), Kayral et al. (2023). This precise estimation of visitor arrivals identified peak and lean months in the study area. The future projections from the systematic methodology would guide

stakeholders, tourism operators, and planners to plan for sustainable tourism development in Himachal Pradesh.

5 Results and discussions

5.1 Indian tourist arrival

The state experiences seasonality throughout the year, mainly due to climatic conditions, events, religious festivals, etc. The seasonality in the Indian tourist arrival was exhibited through the month-wise time series plot of variations in the dataset (Fig. 3). The time series demonstrated the peak season for Indian tourists from April, May, and June because of congregated holidays and institutional vacations. During summers the state becomes a destination of tourism pleasure. The time series data demonstrated irregular behaviour because of seasonality, trend, and cycle components. The fluctuations in the Indian tourist arrival data follow an upward trend. The graph shows the seasonality by identifying imbalanced spikes depicting the real-time tourist arrivals numbers. Besides summer, tourists visit the study area to relish the season of apple plantations/orchards, and cherry gardens during September and October. The state experiences the lean season during November, December, and January because of extreme winters and snowfall. The region observes a critical drop in tourist arrivals during July and August because of heavy rains and floods. The year 2013, witnessed a severe drop in tourist arrival due to extreme rainfall (Himdhara Environment Action and Research Collective 2016). It is evident from the plot that tourist arrivals started increasing in 2018. However, in 2018, the state faced an acute water crisis in June, impacting peak season tourist arrivals. This lowering trend in 2018 immediately followed the pandemic downfall from 2019 to 2021.

The results of ADF, and KPSS tests revealed that the data are non-stationary at level. The derived ADF value was 0.96 which is greater than 0.05, and KPSS value was 0.07 as shown in Table 1. However, the data became stationary after taking the first difference. The significance value of ADF test was 0.00 ($p < 0.05$) and KPSS 0.16 ($p > 0.05$) as shown in Table 2.

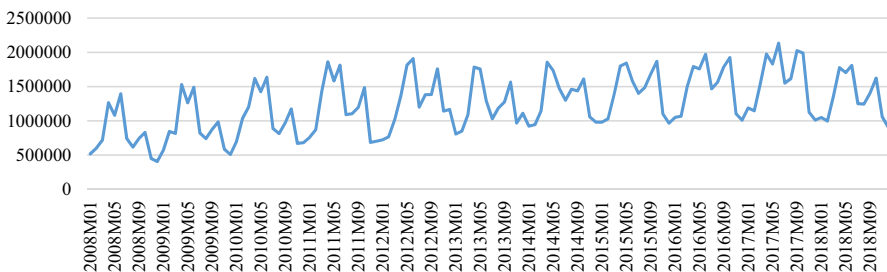


Fig. 3 Month wise time series plot of variations in tourist arrivals: peak season and lean season—Indian (2008–18)

5.1.1 Decomposition of Indian tourist arrivals

The multiplicative decomposition model is $Y_t = T_t S_t \epsilon_t$, the forecasting equation is $y_t = t_t s_t$, where t_t is the trend-cycle component which is $866,276 + 5412.5t$; $t = 1, 2, \dots$ wherein 1 represents January 2008 and s_t is the seasonal factor. Table 3 displays the seasonal factor in percentage that affects the Indian tourist arrivals in the state. The multiplicative decomposition model emphasized that the number of Indian tourist arrivals in the study area is higher during the summer months of April, May, and June, and then grows in September and October. The seasonal factor validates the forecast that the state experiences an increased tourist arrival by, 37.8%, 32.1%, 35.2%, (April to June), 4.9%, and 21.4% (September and October), respectively. Figure 9 exhibits the forecasting of Indian tourist arrival in the study area using Decomposition.

5.1.2 Exponential smoothing using Holt–Winters method for the Indian tourist arrivals

The analysis revealed that the multiplicative model outperformed the additive model by capturing the seasonality component in the time series to enable long-term projection. The multiplicative exponential smoothing model exhibits the level (α), trend (β), and seasonality (γ) to be 0.424, 0.001, and 0.482. The value of the level component indicates that more emphasis is on the recent observation of the data. The trend parameter is 0.001 exhibiting less emphasis on the recent observations, Lidiema (2017); Intarapak et al. (2022) presented similar findings. Hyndman and Athanasopoulos (2018) noted that forecasting models with lower weight on recent observations are generally more stable and less prone to overreacting to sudden changes in the data. Remarkably, Bermúdez et al. (2007) analyzed that including a

Table 1 ADF and KPSS tests at level

Test	Value	Lag	P value
ADF	− 2.37	12	0.96
KPSS	14.40	12	0.07

Table 2 ADF and KPSS tests at first difference

Test	Value	Lag	P value
ADF	− 5.12	12	0.00
KPSS	0.64	12	0.16

Table 3 Seasonal factor in percentage of Indian tourist arrivals in Himachal Pradesh

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
69.3	75.6	97.9	137.8	132.1	135.2	93.5	95.6	104.9	121.4	70.2	66.7

trend component did not improve the forecasting accuracy. As the trend parameter has considerably less weight, seasonality determines the patterns and variations in the dataset. The seasonality parameter is 0.482, indicating a significant weight on recent observations to analyze the seasonal patterns and variation while generating forecasts. Figure 9 exhibits the H–W forecasting of Indian tourist arrival in the study area.

5.1.3 Box–Jenkins method for the Indian tourist arrivals

The increasing trend and seasonal component in the Indian tourist arrival data exhibited non-stationarity (Fig. 3). The time series was transformed by taking the first season difference [D(Forecast)] (Fig. 4). The suitable model was selected using the ACF coefficient which reduces to 0 (Fig. 5), whereas, the PACF drops to 0 (Fig. 6). For the non-seasonal component, the AR (p) is significant at lag 3, 5, 9, 10, 11, 12 and the partial correlation coefficients MA (q) is significant at lag 3, 6, 9, 12. The seasonal difference in this model is 1, therefore, the value of $d=1$. The Indian tourist arrival data has a seasonal component which requires undergoing Seasonal Autoregressive Integrated Moving Average (SARIMA). The MA (Q) seasonal term was examined at lag value 12. The difference value (D) of the model is 1 and the seasonal lag (s) is 12. From the ACF and PACF plots, the suitable SARIMA models were selected. The most fitted model identified based on minimum AIC and BIC values was SARIMA (2,1,2) (1,1,1)₁₂. Thereafter, the correlogram of residuals squared (L-Jung test) found no autocorrelation in the SARIMA (2,1,2) (1,1,1)₁₂ model. The probability values are greater than 0.05 (95%), indicating that the model is well constructed and the residuals are fitted (Fig. 7). Additionally, the residual plot exhibited equal scattering of the error terms (Fig. 8). Lastly, the arrival of Indian tourists was forecasted from 2019 to 2031 using the fitted model (Fig. 9).

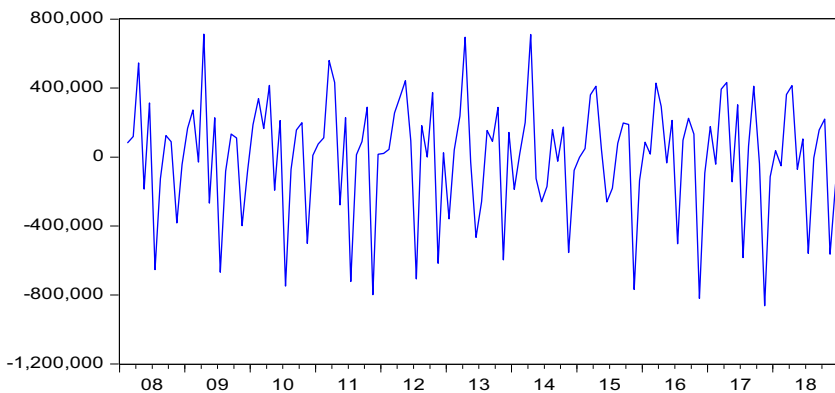


Fig. 4 Time series transformation by seasonal difference [D(Forecast)]

Date: 03/17/23 Time: 14:56
 Sample: 2008M01 2018M12
 Included observations: 132

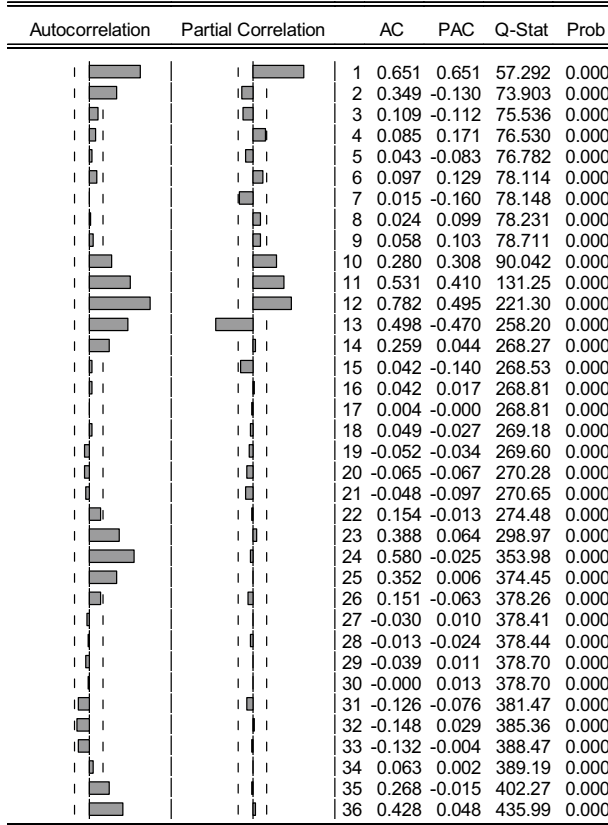


Fig. 5 ACF and PACF plots for the Indian tourist arrival (Tourist arrivals I)

5.1.4 Comparison between the three selected model of forecasting Indian tourist arrivals

Table 4 exhibits the comparison between the selected model to forecast Indian Tourist arrivals. The comparative graph between the actual values and forecasting values of the decomposition, SARIMA, and H–W models are exhibited in Fig. 9. The accuracy measurements are based on the predicted tourist arrival values from 2019 and onwards instead of the actual values during that time. Estimating the accuracy of forecast beyond the training period and assessing the degree of fit between the forecast curve and the actual data from 2008 to 2018 are the main objectives of these measurements. The more the actual and forecast curve overlap the historical trend from 2008M01 to 2018M12, the more accurate predictions are. The forecasting graph demonstrates an upward-moving diagram anticipating an increase in tourist arrivals. The forecast accuracy of H–W is reasonably accurate at 10.18%,

Date: 03/17/23 Time: 14:22
 Sample: 2008M01 2018M12
 Included observations: 131

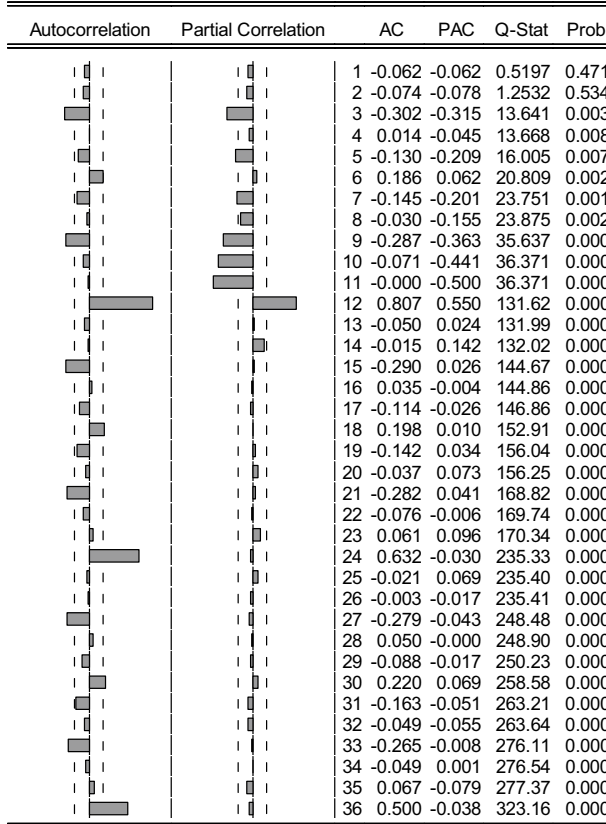


Fig. 6 ACF and PACF plots for the seasonal differences in Indian tourist arrival [D(Tourist arrivals I)]

Decomposition is good accuracy at 9.67%, and SARIMA is highly accurate at 3% MAPE coefficient value. The RMSE values of the H–W model (170,399.2) through values is more than SARIMA model and decomposition model. The RMSE values of Decomposition model is 148,977.2 which is higher than the SARIMA model. Theil’s U1 coefficient results of all the three models indicate a higher accuracy and equal distribution in the dataset. The SARIMA (2,1,2) (1,1,1)₁₂ model demonstrates the best forecasting results according to Table 4. Therefore, the findings discovered that the coefficients evaluated the forecast accuracy beyond the training period, making it useful for out-of-sample performance.

The H–W overestimated the forecast value than the actual value, whereas, SARIMA accurately forecasted the tourist arrivals with the lowest risk of overestimation or underestimation as reflected by the RMSE, MAPE, and Theil’s U1 coefficients. Here, the overestimation means that the forecasted value is higher than the actual tourist arrivals, and lower in case of underestimation. Figure 9

Date: 03/18/23 Time: 10:30
 Sample: 2008M01 2018M12
 Included observations: 131
 Q-statistic probabilities adjusted for 6 ARMA terms

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.057	-0.057	0.4311	
		2	0.001	-0.003	0.4311	
		3	0.082	0.082	1.3413	
		4	0.087	0.097	2.3837	
		5	0.053	0.065	2.7706	
		6	-0.051	-0.052	3.1364	
		7	-0.023	-0.048	3.2138	0.073
		8	-0.098	-0.125	4.5636	0.102
		9	-0.046	-0.066	4.8606	0.182
		10	-0.038	-0.034	5.0638	0.281
		11	0.028	0.059	5.1758	0.395
		12	0.035	0.082	5.3521	0.500
		13	-0.095	-0.060	6.6986	0.461
		14	0.144	0.133	9.7911	0.280
		15	0.012	0.003	9.8118	0.366
		16	0.039	0.019	10.038	0.437
		17	-0.017	-0.046	10.084	0.523
		18	0.017	-0.014	10.129	0.605
		19	0.106	0.087	11.877	0.538
		20	-0.009	0.024	11.891	0.615
		21	-0.041	-0.036	12.157	0.667
		22	-0.095	-0.098	13.594	0.629
		23	0.143	0.126	16.878	0.463
		24	-0.088	-0.064	18.128	0.447
		25	-0.014	0.008	18.160	0.512
		26	0.034	0.023	18.355	0.564
		27	-0.103	-0.073	20.121	0.514
		28	-0.045	-0.080	20.456	0.555
		29	0.008	0.007	20.468	0.614
		30	0.096	0.083	22.063	0.576
		31	-0.050	-0.009	22.504	0.606
		32	-0.035	-0.002	22.722	0.649
		33	0.026	-0.008	22.844	0.693
		34	0.030	-0.002	23.006	0.733
		35	0.010	-0.028	23.022	0.776
		36	-0.130	-0.100	26.127	0.669

Fig. 7 Correlogram of residuals squared (L-Jung test) for SARIMA (2,1,2) (1,1,1)₁₂

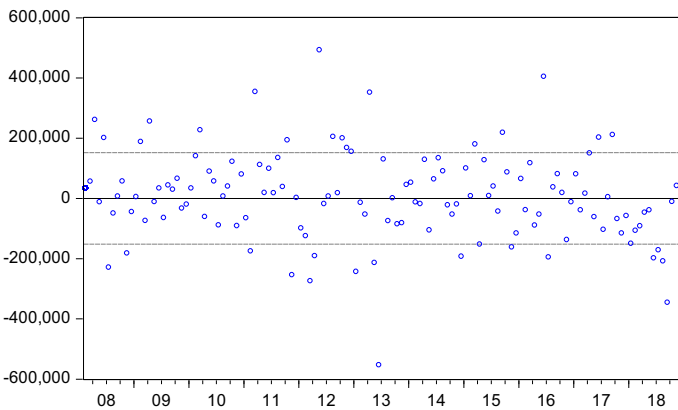
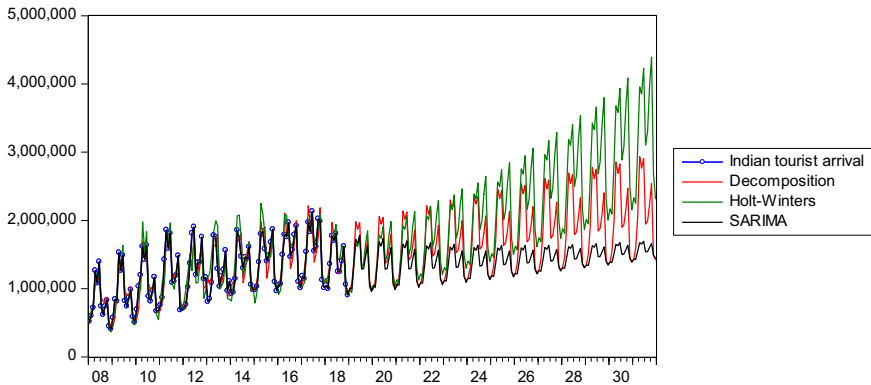


Fig. 8 Dot plot of residuals vs forecast values of Indian tourist arrivals to the study area

Table 4 Comparison between the selected model to forecast Indian Tourist Arrivals

Forecast	RMSE (Arrival numbers)	MAPE (%)	Theil's U1
Decomposition	148,977.2	9.67	0.05
H–W	170,399.2	10.18	0.06
SARIMA (2,1,2) (1,1,1) ₁₂	36,833.8	3.0	0.02



Note: Training sample 2008M01 2018M12 Estimation sample 2019M01 2031M12

Fig. 9 Comparative forecasting of Indian Tourist Arrivals from selected models

demonstrates the outcome variation in the forecasting model and confirms that SARIMA best fits the historical data. The state will experience overcrowding by 2031, especially during peak seasons, causing adverse resource scarcity and a dearth of infrastructure. However, these forecasts are merely scientific estimations and are bound to change with the progress in time due to unexpected events and external factors. This study solely assessed the future tourist arrivals, which follows the pre-pandemic trend. COVID-19 was a significant event severely impacting the environment and socio-economic richness in the state. However, this study acknowledges that presenting pandemic tourism scenario is vital and recommends that future research could use advanced data mining techniques to improve forecasting accuracy. M. Fátima and Rocha (2023) analyzed the “pre-pandemic tourism forecasts and post-pandemic signs of recovery assessment for Portugal.” Thus, this study unlocks opportunities for future researchers to focus on the post-pandemic Indian tourist forecast in Himachal Pradesh and outline a sustainable recovery plan.

5.2 Foreign tourist arrival

The fluctuation in foreign tourist arrivals plot as a month-wise time series is shown in Fig. 10. Evidently, the fluctuations in the tourist arrivals follow a subtle

downward trend. The time series data demonstrated irregular behaviour because of seasonality, trend, and cycle components. Foreign visitors travel from May to September which compliments domestic arrivals from April to June. The highest foreign arrivals are recorded in the tribal districts of Kullu, Kinnaur, and Lahaul and Spiti particularly from May to July. August observed a reduced foreign tourist arrival because of heavy rains, raising the risk of landslides and floods. Tourist arrivals dropped in 2013 and 2018 due to heavy snowfall and the water crisis. The temporal imbalance is exhibited by the seasonality of foreign arrivals, particularly during the peak season. Local business owners in Lahaul and Spiti district rely on foreign tourists for income rather than Indians. Foreign visitors are more sensitive to the native diversity and traditional values of the region. Thus, the state envisions rejuvenating foreign tourist arrivals and receipts.

The results of ADF and KPSS tests for foreign tourists revealed that the data was non-stationary at level. The estimated ADF value was 0.65 ($p > 0.05$), and the KPSS value (0.05) shown in Table 5. However, the significance value of ADF was 0.00 ($p < 0.05$), and KPSS was 0.13 ($p > 0.05$) after taking the first difference confirmed stationarity in the dataset in Table 6.

5.2.1 Decomposition of foreign tourist arrivals

In forecasting equation $y_t = t_i s_p$, $t_i = 36,780 + 7.6628t$ and s_i is the seasonal factor. Table 7 displays the seasonal factor in percentage that affects the foreign tourist arrivals

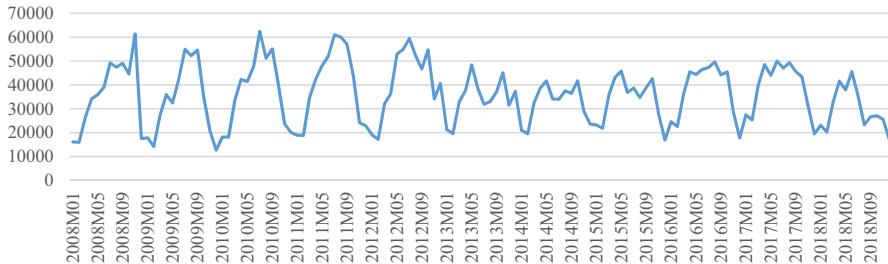


Fig. 10 Month wise time series plot of variation of tourist arrivals: Peak season and lean season- Foreign (2008–18)

Table 5 ADF and KPSS tests at level

Test	Value	Lag	P value
ADF	-1.89	12	0.65
KPSS	17.35	12	0.05

Table 6 ADF and KPSS tests at first difference

Test	Value	Lag	P value
ADF	-5.21	12	0.00
KPSS	0.00	12	0.13

in the state. The multiplicative decomposition model emphasized that the number of foreign tourist arrivals increase from April to July and descend from August to October. The seasonal factor validates the forecast that the state experiences an increased tourist arrival by 13.3%, 22.2%, 23.4%, 30.7%, 26.5%, 23.9%, and 18.6% from April to October. Furthermore, the lean period spans from November to March.

5.2.2 Exponential smoothing using Holt–Winters method for foreign tourist arrivals

For the long-term projection, the multiplicative exponential smoothing model exhibited the level (α), trend (β), and seasonality (γ) to be 0.054, 0.001, and 0.940. Generally, smoothing parameter do not vary considerably and they almost achieve consistency in the training sample. α emphasized on the previous observations suggesting a slow and stable change over time. β indicates a greater importance on the past observations while generating the forecast, which is suitable for long-term projections, as presented by Lidiema (2017); Alonso Brito et al. (2021); Intarapak et al. (2022), among others. Although weak presently, the trend would gain prominence in the future because of revised tourism policies by the State Tourism Authority. Likewise, γ assigns heavy weight on recent observations presenting a higher sensitivity to specific parts. Figure 16 exhibits the H–W forecasting of foreign tourist arrival in the study area.

5.2.3 Box–Jenkins method for foreign tourist arrivals

The decreasing trend and seasonal component in foreign tourist arrivals exhibited non-stationarity (Fig. 10). The time series was transformed by taking the first season difference [D(Forecast)] (Fig. 11). The suitable model was selected using the ACF coefficient which gradually reduces to 0 (Fig. 12), whereas, the PACF drops to 0 (Fig. 13). For the non-seasonal component, the AR (p) is significant at 4, 8, 9, 12 and MA (q) is significant at lag 4, 12, 16, 24. The seasonal difference in this model is 1, therefore, the value of $d=1$. The foreign tourist arrival data has a seasonal component which requires undergoing Seasonal Autoregressive Integrated Moving Average (SARIMA). The MA(Q) seasonal term was examined at lag value 12. From the ACF and PACF plots, the suitable SARIMA models are selected. The most fitted model identified based on minimum AIC and BIC values was SARIMA (3,1,3) (1,1,1)₁₂. After that, the correlogram of residuals squared (L–Jung test) found no autocorrelation in the SARIMA (3,1,3) (1,1,1)₁₂ model. The probability values are greater than 0.05 (95%), indicating that the model is well constructed and the residuals are fitted (Fig. 14). Additionally, the residual plot exhibited equal scattering of the error

Table 7 Seasonal factor in percentage of foreign tourist arrivals in Himachal Pradesh

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
58	54.1	93	113.3	122.2	123.4	130.7	126.5	123.9	118.6	77.3	59

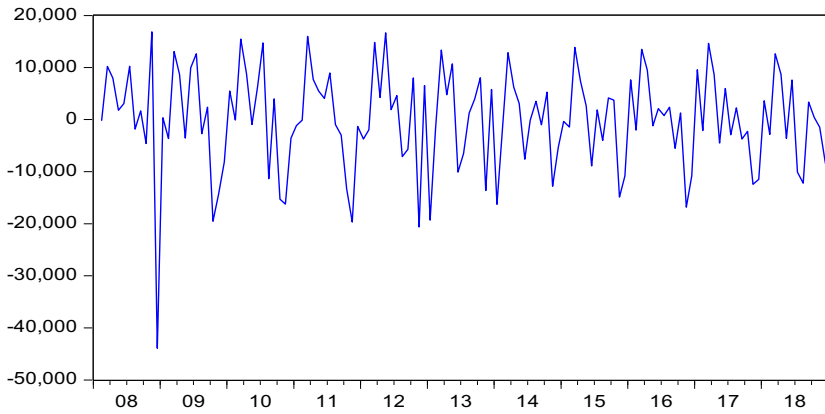


Fig. 11 Time series transformation by seasonal difference [D(Forecast)]

Date: 03/21/23 Time: 12:34
 Sample: 2008M01 2018M12
 Included observations: 132

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.683	0.683	62.909	0.000
		2 0.354	-0.209	79.969	0.000
		3 -0.037	-0.366	80.160	0.000
		4 -0.355	-0.250	97.618	0.000
		5 -0.501	-0.071	132.60	0.000
		6 -0.571	-0.241	178.33	0.000
		7 -0.495	-0.153	213.03	0.000
		8 -0.326	-0.086	228.16	0.000
		9 -0.011	0.187	228.18	0.000
		10 0.351	0.277	245.99	0.000
		11 0.598	0.162	298.21	0.000
		12 0.720	0.240	374.65	0.000
		13 0.475	-0.290	408.14	0.000
		14 0.190	-0.034	413.57	0.000
		15 -0.103	0.079	415.16	0.000
		16 -0.367	-0.068	435.67	0.000
		17 -0.462	0.050	468.45	0.000
		18 -0.510	-0.020	508.87	0.000
		19 -0.441	-0.082	539.28	0.000
		20 -0.241	0.030	548.44	0.000
		21 0.024	-0.080	548.54	0.000
		22 0.310	-0.075	563.96	0.000
		23 0.462	-0.069	598.52	0.000
		24 0.516	0.067	642.11	0.000
		25 0.339	0.006	661.11	0.000
		26 0.083	-0.138	662.27	0.000
		27 -0.128	0.017	665.04	0.000
		28 -0.336	-0.008	684.25	0.000
		29 -0.402	-0.018	712.02	0.000
		30 -0.413	0.006	741.63	0.000
		31 -0.339	-0.040	761.72	0.000
		32 -0.154	-0.040	765.91	0.000
		33 0.064	0.038	766.65	0.000
		34 0.272	-0.053	780.02	0.000
		35 0.370	-0.055	804.97	0.000
		36 0.375	-0.006	830.93	0.000

Fig. 12 ACF and PACF plots for the foreign tourist arrival (Tourist arrivals F)

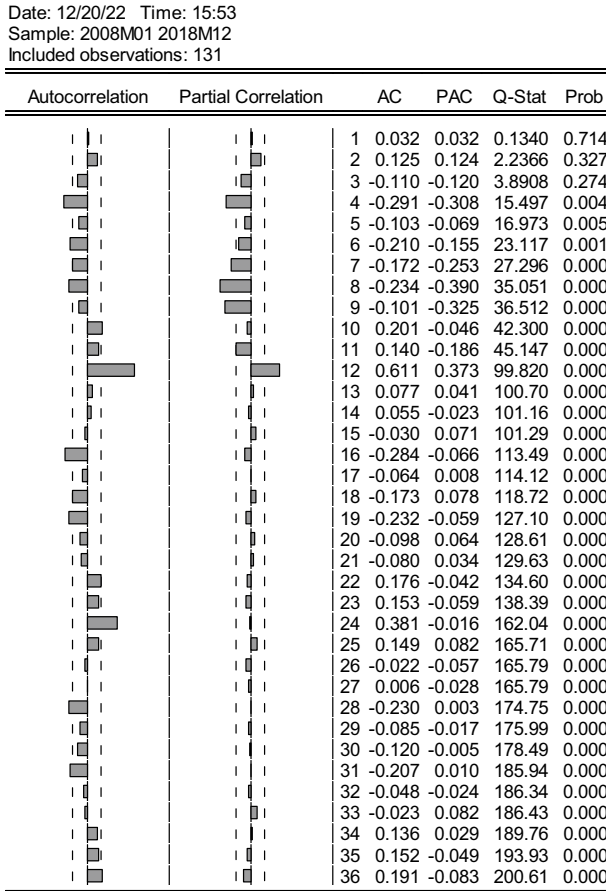


Fig. 13 ACF and PACF plots for the seasonal differences in foreign tourist arrival [D(Tourist arrivals F)]

terms (Fig. 15). Lastly, the arrival of foreign tourists was forecasted from 2019 to 2031 using the fitted SARIMA model.

5.2.4 Comparison between the three selected model of forecasting foreign tourist arrivals

Table 8 exhibits the comparison between the selected model to forecast foreign Tourist arrivals. The forecasting graph shows a linear downward curve in the foreign arrivals. The forecast accuracy for Decomposition is reasonably accurate at 23.72%, H–W is good accuracy at 17.22%, and SARIMA is highly accurate at 15.51% MAPE coefficient value. The RMSE value of the Decomposition model is 11,126.87, which is higher than the SARIMA model and Decomposition model. The RMSE value of the H–W model is 6907.59 which is equivalent to the SARIMA model. Their’s U1

Date: 01/03/23 Time: 20:32
 Sample: 2008M01 2018M12
 Included observations: 131

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.113	0.113	1.7254	0.189
		2	-0.088	-0.102	2.7751	0.250
		3	0.046	0.071	3.0680	0.381
		4	0.000	-0.025	3.0680	0.547
		5	-0.032	-0.018	3.2132	0.667
		6	-0.094	-0.096	4.4518	0.616
		7	0.037	0.060	4.6433	0.703
		8	0.055	0.027	5.0708	0.750
		9	-0.057	-0.050	5.5417	0.785
		10	0.048	0.065	5.8782	0.825
		11	-0.008	-0.044	5.8880	0.881
		12	0.042	0.064	6.1432	0.909
		13	-0.026	-0.047	6.2408	0.937
		14	-0.016	0.015	6.2787	0.959
		15	0.042	0.016	6.5435	0.969
		16	0.015	0.025	6.5769	0.981
		17	0.014	0.013	6.6058	0.988
		18	-0.045	-0.052	6.9139	0.991
		19	-0.004	0.013	6.9168	0.995
		20	0.058	0.038	7.4520	0.995
		21	-0.048	-0.037	7.8222	0.996
		22	0.037	0.051	8.0413	0.997
		23	0.198	0.185	14.355	0.916
		24	0.201	0.168	20.929	0.643
		25	-0.013	-0.026	20.958	0.695
		26	-0.011	0.026	20.979	0.743
		27	0.025	-0.017	21.080	0.783
		28	-0.043	-0.017	21.399	0.808
		29	-0.031	0.019	21.566	0.838
		30	-0.051	-0.056	22.013	0.854
		31	-0.002	-0.008	22.014	0.883
		32	-0.018	-0.040	22.069	0.906
		33	-0.074	-0.057	23.039	0.902
		34	-0.002	-0.028	23.040	0.923
		35	0.022	0.013	23.128	0.938
		36	0.006	-0.003	23.135	0.952

Fig. 14 Correlogram of Residuals Squared (L-Jung test) for SARIMA (3,1,3) (1,1,1)₁₂

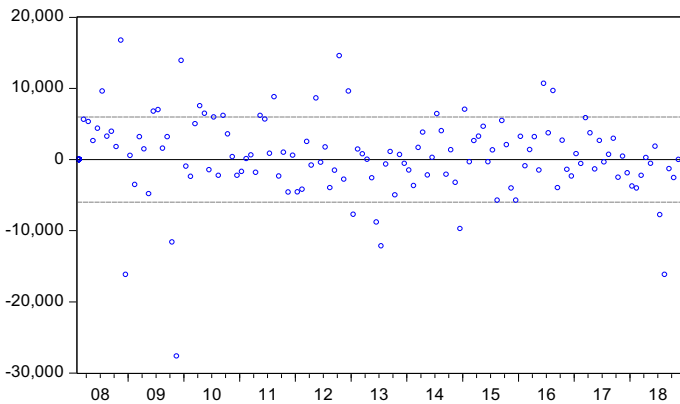


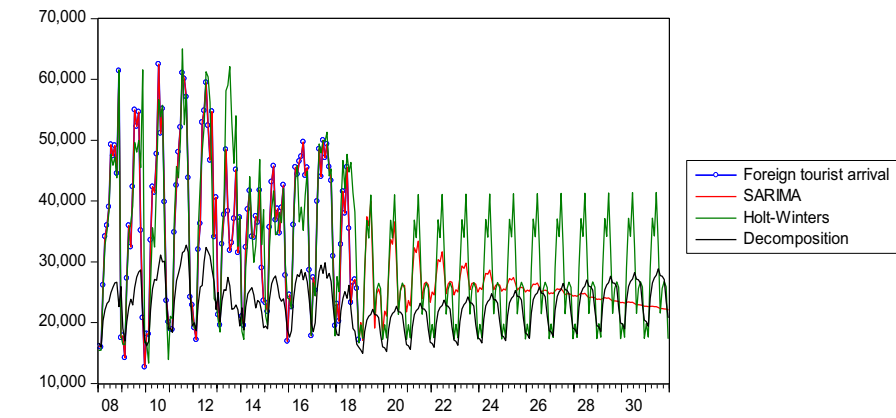
Fig. 15 Dot plot of residuals vs forecast values of foreign tourist arrivals to the study area

Table 8 Comparison between the selected model to forecast foreign Tourist Arrivals

Forecast	RMSE (Arrival numbers)	MAPE (%)	Theil's U1
Decomposition	11,126.87	23.72	0.20
H–W	6907.59	17.22	0.10
SARIMA (3,1,3) (1,1,1) ₁₂	6907.59	15.51	0.10

coefficient results of all three models indicate a higher accuracy and equal distribution in the dataset. The SARIMA (3,1,3) (1,1,1)₁₂ model accurately forecasts foreign tourist arrivals. The comparative graph between the actual and forecasted values of the decomposition, B–J, and H–W exponential smoothing methods is shown in Fig. 16. Thus, the findings revealed that the coefficients evaluated the forecast accuracy beyond the training period from 2019M01 to 2031M12.

The graph shows a critical drop in tourist arrivals in 2013 (natural disaster) and 2018 (water crisis) affected the time series trend. The foreign tourist arrivals have inconsistencies shown by the trend because of the high degree of seasonality due to natural and climatic adversities. Furthermore, destinations get tremendous tourist footfall during peak seasons, impacting regional stability and safety. Notably, H–W model exhibit that the trend component has no visible influence after the year 2018, which experienced an acute water shortage impacting the foreign arrivals. The H–W model overestimated the future demand more than the actual value during peak seasons, whereas the decomposition curve underestimated peaks; Fosgerau et al. (2013) recorded similar findings. In contrast, the SARIMA model accurately estimated the tourist arrivals with the lowest risk of overestimation or underestimation, as reflected by accuracy coefficients. The SARIMA model exhibits a declining trend and reduced seasonal fluctuations



Note: Training sample 2008M01 2018M12 Estimation sample 2019M01 2031M12

Fig. 16 Comparative forecasting of foreign Tourist Arrivals from selected models

because of the strong influence of recent data values (refer Fig. 16); notably the year 2018, since there was an acute water shortage severely impacted the peak season for foreign tourists. The present study realized that the time series considered is relatively short, therefore, the H–W model might not have enough information to evidently detect the trend. In contrast to the H–W model, B–J can capture trend even in shorter series by analyzing the relationship between past and present observations. Thus, the authors suggest that future researchers collate long-term time series data to assess the smoothing parameters efficiently and present a reliable forecast of foreign tourist arrivals. Although the forecasts present the pre-pandemic scenario, they explicitly highlight the need to assess the pandemic scenario through a comparative assessment. The region witnesses significant foreign tourist arrivals from July to August. During the monsoon period, the state gets flooded with a high risk of landslides, hindering traffic movement. Since the state is vulnerable to natural calamities resulting in declined foreign demand, implementing bioengineering is essential for long-term tourism planning, as Kohler et al. (2012) and Cantasano et al. (2023) suggested. Bioengineering would promote environmentally responsible tourism with minimum environmental impacts to attract foreign tourists. Finally, this study sets a base for future studies to accurately estimate future demand by employing efficient models capturing seasonality in time series.

6 Policy recommendations

The results established a high degree of tourism seasonality in the mountainous state which interrupts the efficient operation of tourism infrastructure. Seasonality is a critical factor in developing tourism policy for sustainable development (Su et al. 2022). The natural factors of tourism seasonality cannot get eradicated, but the institutional factors can improve the tourism seasonality in a region (Rizal and Asokan 2014). An accurate forecast of future demand is necessary to strengthen the current regulatory framework for sustainable tourism growth. To sum up, a precise tourist projection is a fundamental requirement for tourism planning to mitigate the adverse effects of tourism, diversify destinations and develop communities. Keeping this in mind; policy level recommendations to mitigate tourism seasonality in the study area are as follows;

-
- Strengthening the existing Himachal Pradesh Tourism Policy’19:** The policy aims at protecting destinations through sustainable policies and initiatives. The existing policy needs alignment with the national strategies of the Ministry of Tourism India for environmental promotion, economic and socio-cultural sustainability, certification schemes, capacity building, and destination/product development.
- The main objective of the policy is building infrastructure and working towards a new marketing plan for tourist destinations. Therefore, a policy on tourism marketing adhering to the carrying capacity-based model would ensure natural preservation for future development.
-

Tourism diversification and destination management: Every season, visitors witness festivals and events like Dussehra, Shivratri, Holi, and Yatras. These festivals resulted in unfathomable problems concerning solid waste, pollution, and disasters which required an immediate response from the authorities.

- The biggest competitors to Himachal Pradesh are Uttarakhand and Jammu & Kashmir, offering diverse tourism. Thus, the state should develop tourism specializations supported by eco-tourism and forest policy.
- Preparation of destination development plans to regulate the impacts of increased tourist footfall. These plans permit regular monitoring of the saturated or to-be-saturated destinations by introducing tourism laws and strategies.
- Tourism policies on offering incentives for travel and tour packages during shoulder or lean seasons to diversify tourism, promote local marketing, and reduce and mitigate the impacts of seasonality.
- The Government must frame policies for travel insurance to ensure safe travel by partnering with authorized insurance providers and supportive infrastructure. The policies must deliberate on the safety of foreign tourists visiting the state during monsoon season, which is highly prone to disasters. This intervention can be extended further for severe winter seasons.

Long-term tourism forecasting: Long-term estimates will assist policymakers in preparing destination-level crisis management plans, strategically elevating regional tourism, and regulating the massive tourist inflow and pilgrim traffic for sustainable transformation.

- Regular long-term tourism forecasting using reliable and robust methods for assessing the policy alternatives in the region. Plans for long-term tourism development aligned with a robust policy framework would ensure "tourism for all," expand the sector, protect tourist destinations, promote host communities, develop human resources, and stimulate private investment opportunities.
- Effective public-private participation (PPP) tourist packages through a partnership between the government and the private sector. The PPP interventions would significantly transform tourism, particularly during shoulder or lean seasons.
- In the long run, policy formulation to promote environmentally responsible tourism with minimum environmental impacts is essential in the state.
- The state experiences soil erosion, which needs controlling by adopting a stringent policy on sustainable construction techniques such as Bioengineering and plantation.

Infrastructure recommendations: The region has insufficient water resources to meet the existing demand. Thus, an adequate water conservation policy is necessary to meet the escalating water demand.

- The hills in the region are under massive impact from solid waste dumping. It indicates an immediate ban on single-use plastic through a regulatory framework in protected areas, religious sites, and rural areas.
- The primary mode of transportation, roads, need protection from harsh climatic conditions and disasters. Therefore, road alignment should strictly avoid fault zones and landslide-prone areas to minimize the risks.
- Strengthening existing airports and helipads is crucial for sharing the future tourist demand while improving regional air connectivity through UDAN (Ude Desh ka Aam Naagrik) scheme.

7 Conclusion

Estimating future demand is challenging because of the high seasonality in the hilly state of Himachal Pradesh. In this context, this study framed a comprehensive approach toward forecasting tourist demand using a univariate time series analysis. The methods applied are; the Decomposition, Box–Jenkins, and Holt–Winters exponential smoothing methods, based on their ability to assess the historical data and trends, determine the degree of seasonality, and ability for long-term forecasting. The empirical findings reveal a robust and reliable approach to forecasting tourist

demand in destinations facing seasonality for a longer span; from 2019 to 2031. The forecasting accuracy of the SARIMA model has the lowest MAPE values, 3% and 15.51%, for Indian and foreign visitors, respectively. This study is a pioneer in forecasting tourist demand in Himachal Pradesh and has significant implications for tourism planning and management, crisis management, and policy formulation. The suggested approach will also benefit sales, weather, stock prices, finance, retail, and manufacturing industries. The research findings would assist policymakers and tourism managers in making informed decisions about resource allocation, infrastructure development, and marketing strategies. Although comprehensive, this study has a few limitations; it solely assessed the future tourist arrivals, which follows the pre-pandemic trend. COVID-19 was a significant event that severely impacted the environment, social, and economic richness in the state. Furthermore, the scope of this study does not cover the treatment of the actual data between 2019 and 2021. The limitations in this study set avenues for future research to employ cutting-edge techniques such as artificial neural systems, artificial neural network, fuzzy neural systems, theta models, genetic algorithms, and other advanced data mining techniques for comparative forecasting.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

- NITI Aayog (2018) Sustainable tourism in the Indian Himalayan region contributing to sustainable development in Indian Himalayan region NITI Aayog 2 sustainable tourism in the Indian Himalayan Region. 1–100
- Alonso Brito GR, RiveroVillaverde A, Lau Quan A, Ruíz Pérez ME (2021) Comparison between SARIMA and Holt-Winters models for forecasting monthly streamflow in the western region of Cuba. *SN Appl Sci*. <https://doi.org/10.1007/s42452-021-04667-5>
- Ampointolas A (2021) Modeling and forecasting daily hotel demand: a comparison based on SARI-MAX, neural networks, and GARCH models. *Forecasting* 3:580–595. <https://doi.org/10.3390/forecast3030037>
- Aslanargun A, Mammadov M, Yazici B, Yolacan S (2007) Comparison of ARIMA, neural networks and hybrid models in time series: tourist arrival forecasting. *J Stat Comput Simul* 77:29–53. <https://doi.org/10.1080/10629360600564874>
- Athanasopoulos G, de Silva A (2012) Multivariate exponential smoothing for forecasting tourist arrivals. *J Travel Res* 51:640–652. <https://doi.org/10.1177/0047287511434115>
- Athanasopoulos G, Hyndman RJ, Song H, Wu DC (2011) The tourism forecasting competition. *Int J Forecast* 27:822–844. <https://doi.org/10.1016/j.ijforecast.2010.04.009>
- Badar RN, Bahadure SP (2020) Assessing tourism sustainability in hill towns: case study of Shimla India. *J Mt Sci* 17:2241–2261
- Bermúdez JD, Segura JV, Vercher E (2007) Holt-Winters forecasting: an alternative formulation applied to UK air passenger data. *J Appl Stat* 34:1075–1090. <https://doi.org/10.1080/02664760701592125>
- Bolarinwa BT, Bolarinwa IA (2021) Single versus combined forecasting: The case of wind run data. *Euro Sci J ESJ* 17:189. <https://doi.org/10.19044/esj.2021.v17n21p189>

- Botha BI, Saayman A (2022) Forecasting tourism demand cycles: a Markov switching approach. *Int J Tour Res*. <https://doi.org/10.1002/jtr.2543>
- Brida JG, Garrido N (2011) Tourism forecasting using SARIMA models in Chilean regions. *Int J Leisure Tour Market* 2:176. <https://doi.org/10.1504/ijltm.2011.038888>
- Cantasano N, Boccalaro F, Letto F (2023) Assessing of detached breakwaters and beach nourishment environmental impacts in Italy: a review. *Environ Monit Assess*. <https://doi.org/10.1007/s10661-022-10666-9>
- Chen RJC, Bloomfield P, Fu JS (2003) An evaluation of alternative forecasting methods to recreation visitation. *J Leis Res* 35:441–454. <https://doi.org/10.1080/00222216.2003.11950005>
- Chen R, Liang CY, Hong WC, Gu DX (2015) Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm. *Appl Soft Comput* 26:435–443. <https://doi.org/10.1016/j.asoc.2014.10.022>
- Chen JL, Li G, Wu DC, Shen S (2019) Forecasting seasonal tourism demand using a multiseriess structural time series method. *J Travel Res* 58:92–103. <https://doi.org/10.1177/0047287517737191>
- Chetty V, Manisha K (2022) Assessing and predicting urban growth patterns using ANN-MLP and CA Model in Jammu Urban agglomeration, India BT - modeling, simulation and optimization. In: Patgiri R, Bandyopadhyay S, Balas VE (eds) Das B. Springer Nature Singapore, Singapore, pp 387–397
- Cho V (2003) A comparison of three different approaches to tourist arrival forecasting. *Tour Manag* 24:323–330. [https://doi.org/10.1016/S0261-5177\(02\)00068-7](https://doi.org/10.1016/S0261-5177(02)00068-7)
- Choe Y, Kim H, Joun HJ (2019) Differences in tourist behaviors across the seasons: THE case of Northern Indiana. Sustainability (Switzerland). <https://doi.org/10.3390/su11164351>
- Planning Commission (2011) Himachal pradesh development report. planning commission, government of India 39–58
- Ćorluka G (2019) Tourism seasonality - an overview
- Crispino M, Mariani V (2024) A tool to nowcast tourist overnight stays with payment data and complementary indicators. *Ital Econ J*. <https://doi.org/10.1007/s40797-024-00266-6>
- de Fátima BM, Rocha ML (2023) COVID-19 pre-pandemic tourism forecasts and post-pandemic signs of recovery assessment for Portugal. *Res Glob*. <https://doi.org/10.1016/j.resglo.2023.100167>
- Duro JA, Turrión-Prats J (2019) Tourism seasonality worldwide. *Tour Manag Perspect* 31:38–53. <https://doi.org/10.1016/j.tmp.2019.03.010>
- Elamin N, Fukushima M (2018) Forecasting extreme seasonal tourism demand: the case of Rishiri Island in Japan. *Asia-Pac J Regional Sci* 2:279–296. <https://doi.org/10.1007/s41685-017-0048-y>
- Elseidi M (2023) Forecasting temperature data with complex seasonality using time series methods. *Model Earth Syst Environ* 9:2553–2567. <https://doi.org/10.1007/s40808-022-01632-y>
- Fahey L, King WR, Narayanan VK (1981) Environmental scanning and forecasting in strategic planning—the state of the art. *Long Range Plan* 14:32–39. [https://doi.org/10.1016/0024-6301\(81\)90148-5](https://doi.org/10.1016/0024-6301(81)90148-5)
- Fatema N, Malik H, Abd Halim MS (2022) Hybrid approach combining EMD, ARIMA and Monte Carlo for multi-step ahead medical tourism forecasting. *J Intell Fuzzy Syst* 42:1235–1251. <https://doi.org/10.3233/JIFS-189785>
- Fosgerau M, Frejinger E, Karlstrom A (2013) A link based network route choice model with unrestricted choice set. Munich Personal RePEc Archive
- Goh C, Law R (2002) Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tour Manag* 23:499–510. [https://doi.org/10.1016/S0261-5177\(02\)00009-2](https://doi.org/10.1016/S0261-5177(02)00009-2)
- Grifoll M, Sanchez-Espigares J, Feng H (2021) Forecasting performance of cruise passengers: the Spanish ports case. *Int J Tour Res* 23:667–676. <https://doi.org/10.1002/jtr.2433>
- Gujarati D (2015) *Econometrics by example second edition*. 466
- Gunter U (2021) Improving hotel room demand forecasts for Vienna across hotel classes and forecast horizons: single models and combination techniques based on encompassing tests. *Forecasting* 3:884–919. <https://doi.org/10.3390/forecast3040054>
- Hansun S, Kristanda MB, Subanar, et al (2019) Forecasting domestic tourist arrivals to Bali: H-WEMA approach. *Proceedings of 2019 5th international conference on new media studies, CONMEDIA 2019* 121–124. <https://doi.org/10.1109/CONMEDIA46929.2019.8981825>
- Hassani H, Silva ES, Antonakakis N et al (2017) Forecasting accuracy evaluation of tourist arrivals. *Ann Tour Res* 63:112–127. <https://doi.org/10.1016/J.ANNALS.2017.01.008>
- Higham J, Hinch T (2002) Tourism, sport and seasons: the challenges and potential of overcoming seasonality in the sport and tourism sectors. *Tour Manag* 23:175–185. [https://doi.org/10.1016/S0261-5177\(01\)00046-2](https://doi.org/10.1016/S0261-5177(01)00046-2)

- Himdhara Environment Action and Research Collective (2016) Dried & dusted state of the rivers – Himachal Pradesh. State of the rivers report – Himachal Pradesh India Rivers Week 2016 1–64
- Hyndman RJ, Athanasopoulos G (2018) Forecasting: principles and practice
- IBEF (2023) Himachal Pradesh. Government of India 1–46. <https://doi.org/10.4324/9781003476900-16>
- Intarapak S, Supapakorn T, Vuthipongse W (2022) Classical forecasting of international tourist arrivals to Thailand. *J Stat Theory Appl* 21:31–43. <https://doi.org/10.1007/s44199-022-00041-5>
- Jangra R, Kaushik SP (2018) Analysis of trends and seasonality in the tourism industry: the case of a cold desert destination-Kinnaur, Himachal Pradesh. *Afr J Hosp Tour Leisure* 7:1–16
- Jiang Q (2023) Dynamic multivariate interval forecast in tourism demand. *Curr Issue Tour* 26:1593–1616. <https://doi.org/10.1080/13683500.2022.2060068>
- Joyeux R, Milunovich G, Rigg J (2012) Forecasting demand for Australian passports. *Asia Pac J Tour Res* 17:100–119. <https://doi.org/10.1080/10941665.2011.613205>
- Kayral IE, Sari T, TandoğanAktepe NŞ (2023) Forecasting the tourist arrival volumes and tourism income with combined ANN architecture in the post COVID-19 period: the case of Turkey. *Sustainability* 15:15924. <https://doi.org/10.3390/su152215924>
- Ke G, Hu Y, Huang X et al (2016) Epidemiological analysis of hemorrhagic fever with renal syndrome in China with the seasonal-trend decomposition method and the exponential smoothing model. *Sci Rep* 6:1–7. <https://doi.org/10.1038/srep39350>
- Koerner B, Sushartami W, Spencer DM (2023) An assessment of tourism policies and planning in Indonesia. *Tour Recreat Res*. <https://doi.org/10.1080/02508281.2023.2214030>
- Kohler T, Pratt J, Dabarbeix B, et al (2012) Sustainable mountain development green economy and institutions. From Rio 1992 to Rio 2021 and beyond
- Kolková A, Rožehnal P (2022) Hybrid demand forecasting models: pre-pandemic and pandemic use studies. *Equilib Quart J Econ Econ Policy*. 17:699–725. <https://doi.org/10.24136/eq.2022.024>
- Konarasinghe U (2017) Decomposition techniques on forecasting tourist arrivals from Western European countries to Sri Lanka. *SSRN Electr J*. <https://doi.org/10.2139/ssrn.2910291>
- Kriti K (2016) Forecasting foreign tourist arrivals in India using different time series models. *Int J Bus Manag* 15:38–43
- Kumar A, Singh G (2019) Seasonal effect on tourism in India. *J Financ Econ* 7:48–51. <https://doi.org/10.12691/jfe-7-2-1>
- Kumar S, Singh V, Saroha J (2023) Interpretation of land use/land cover dynamics with the application of geospatial techniques in Sarbari Khad watershed of Himachal Pradesh, India. *GeoJournal* 88:2623–2633. <https://doi.org/10.1007/s10708-022-10769-3>
- Lee Y, Huang C (2007) Sustainability index for Taipei. *Environ Impact Assess Rev* 27:505–521. <https://doi.org/10.1016/j.eiar.2006.12.005>
- Li C, Ge P, Liu Z, Zheng W (2020) Forecasting tourist arrivals using denoising and potential factors. *Ann Tour Res* 83:102943. <https://doi.org/10.1016/j.annals.2020.102943>
- Lidiema C (2017) Modelling and forecasting inflation rate in Kenya using SARIMA and Holt-Winters triple exponential smoothing. *Am J Theor Appl Stat* 6:161. <https://doi.org/10.11648/j.ajtas.20170603.15>
- Liu Y, Wen L, Liu H, Song H (2024) Predicting tourism recovery from COVID-19: a time-varying perspective. *Econ Model* 16:106706
- Makoni T, Chikobvu D (2018) Modelling and forecasting Zimbabwe’s tourist arrivals using time series method a case study of Victoria falls rainforest. *Southern Afr Bus Rev*. <https://doi.org/10.25159/1998-8125/3791>
- Makoni T, Chikobvu D, Sigauke C (2021) Hierarchical forecasting of the Zimbabwe international tourist arrivals. *Stat Optim Inf Comput* 9:137–156. <https://doi.org/10.19139/soic-2310-5070-959>
- Manisha K, Singh I, Chetry V (2023) Investigating and analyzing the causality amid tourism, environment, economy, energy consumption, and carbon emissions using Toda-Yamamoto approach for Himachal Pradesh, India. *Environ Dev Sustain*. <https://doi.org/10.1007/s10668-023-04252-3>
- Martín Martín JM, Salinas Fernández JA, Rodríguez Martín JA, del Ostos Rey MS (2020) Analysis of tourism seasonality as a factor limiting the sustainable development of rural areas. *J Hosp Tour Res* 44:45–75. <https://doi.org/10.1177/1096348019876688>
- Ministry of Tourism (2023) India tourism statistics
- Mohamed IE (2010) Simulating time series analysis using SAS - Part III - error correction model (ECM). Nesug 1–7

- Mohd Lip N, Jumery NS, Ahmad Termizi FA et al (2020) Forecasting international tourist arrivals in Malaysia using Sarima and Holt-Winters model. *J Tour Hosp Environ Manag* 5:41–51. <https://doi.org/10.35631/jthem.518004>
- Mustafa H, Ahmed F, Zainol WW, Enh AM (2021) Forecasting the impact of gross domestic product (Gdp) on international tourist arrivals to Langkawi, Malaysia: a postcovid-19 future. *Sustainability* (Switzerland). <https://doi.org/10.3390/su132313372>
- Panigrahi BK, Das S, Nath TK, Senapati MR (2018) An application of data mining techniques for flood forecasting: application in rivers Daya and Bhargavi, India. *J Inst Eng (India): Ser B* 99:331–342. <https://doi.org/10.1007/s40031-018-0333-9>
- Peiris HRI (2016) A seasonal ARIMA model of tourism forecasting: the case of Taiwan. *J Tour Hosp Sport* 22:2312–5187. <https://doi.org/10.1080/10941661003630001>
- Pirthee M (2017) Grey-based model for forecasting Mauritius International Tourism from different regions. *Grey Syst* 7:259–271. <https://doi.org/10.1108/GS-04-2017-0008>
- Pradesh H (2020) COVID-19 : Implications on regional industry and Economy
- Reivan-Ortiz GG, Cong PT, Wong WK et al (2023) Role of geopolitical risk, currency fluctuation, and economic policy on tourist arrivals: temporal analysis of BRICS economies. *Environ Sci Pollut Res* 30:78339–78352. <https://doi.org/10.1007/s11356-023-27736-1>
- Rizal P, Asokan R (2014) Seasonality of tourism: a major constraint for the growth of tourism in the regional economy of Sikkim State, India. *Int J Curr Res* 6:7211–7218
- Shen S, Li G, Song H (2009) Effect of seasonality treatment on the forecasting performance of tourism demand models. *Tour Econ* 15:693–708. <https://doi.org/10.5367/000000009789955116>
- Song H, Li G, Witt SF, Athanasopoulos G (2011) Forecasting tourist arrivals using time-varying parameter structural time series models. *Int J Forecast* 27:855–869. <https://doi.org/10.1016/j.ijforecast.2010.06.001>
- Song H, Qiu RTR, Park J (2019) A review of research on tourism demand forecasting. *Ann Tour Res* 75:338–362. <https://doi.org/10.1016/j.annals.2018.12.001>
- Statistics Times (2021–22) Ministry of Statistics and Programme Implementation - Indian states by GDP. Available online: <https://statisticstimes.com/economy/india/indian-states-gdp.php>. Accessed 28 May 2024
- Su Z, Wen R, Zeng Y et al (2022) The influence of seasonality on the sustainability of livelihoods of households in rural tourism destinations. *Sustainability* (Switzerland). <https://doi.org/10.3390/su141710572>
- Suhartono, Lee MH, Javedani H (2011) A weighted fuzzy integrated time series for forecasting tourist arrivals. *Communications in computer and information science* 252 CCIS:206–217. https://doi.org/10.1007/978-3-642-25453-6_19
- Sun S, Wei Y, Tsui KL, Wang S (2019) Forecasting tourist arrivals with machine learning and internet search index. *Tour Manag* 70:1–10. <https://doi.org/10.1016/j.tourman.2018.07.010>
- Thushara SC, Su JJ, Bandara JS (2019) Forecasting international tourist arrivals in formulating tourism strategies and planning: the case of Sri Lanka. *Cogent Econ Financ* 7:1–32. <https://doi.org/10.1080/23322039.2019.1699884>
- UNWTO (2004) Excerpt from indicators of sustainable development for tourism destinations: a guidebook. world tourism organization 111–116
- Valipour M (2015) Long-term runoff study using SARIMA and ARIMA models in the United States. *Meteorol Appl* 22:592–598
- Vergori AS (2012) Forecasting tourism demand: the role of seasonality. *Tour Econ* 18:915–930. <https://doi.org/10.5367/te.2012.0153>
- Wang H, Hu T, Wu H (2023) Tourism demand with subtle seasonality: recognition and forecasting. *Tour Econ* 29:1865–1889. <https://doi.org/10.1177/13548166231184300>
- WTTC (2022) Travel & tourism economic impact 2022 : global trends. World travel & tourism council travel and tourism economic impact 1–36
- WTTC (2023) Economic impact 2023: travel and tourism India. The economic impact of global travel & tourism
- Wu BQ, Jf Wu, Shi XT et al (2019) Visiting probability model: a new method for tourist volume forecasting. *Asia Pac J Tour Res* 24:1155–1168. <https://doi.org/10.1080/10941665.2019.1666153>
- Wu DCW, Ji L, He K, Tso geoffrey KF (2021) Forecasting tourist daily arrivals with a hybrid Sarima-Lstm approach. *J Hosp Tour Res* 45:52–67. <https://doi.org/10.1177/1096348020934046>
- Wu P, Li G, Wen L, Liu H (2023) Tourism demand forecasting with multiple mixed-frequency data: a reverse mixed-data sampling method. *J Travel Res*. <https://doi.org/10.1177/00472875231203397>

- Yin L (2020) Forecast without historical data: objective tourist volume forecast model for newly developed rural tourism areas of China. *Asia Pac J Tour Res* 25:555–571. <https://doi.org/10.1080/10941665.2020.1752755>
- Yollanda M, Devianto D (2020) Hybrid model of seasonal ARIMA-ANN to forecast tourist arrivals through Minangkabau international airport. <https://doi.org/10.4108/eai.2-8-2019.2290473>
- Zhu M, Wu J, Wang YG (2021) Multi-horizon accommodation demand forecasting: a New Zealand case study. *Int J Tour Res* 23:442–453. <https://doi.org/10.1002/jtr.2416>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Keerti Manisha^{1,2}  · Inderpal Singh¹ 

✉ Keerti Manisha
ar.keertimanisha@gmail.com; keerti@nitt.edu

Inderpal Singh
ipsingh05@gmail.com; ipsingh@nith.ac.in

¹ Department of Architecture, National Institute of Technology Hamirpur, Hamirpur, Himachal Pradesh, India

² Department of Architecture, National Institute of Technology Tiruchirappalli, Tiruchirappalli, Tamil Nadu, India