Forecasting Hotel Prices in Selected Middle East and North Africa Region (MENA) Cities with New Forecasting Tools

Mohammed Al Shehhi, Andreas Karathanasopoulos*

Dubai Business School, University of Dubai, Dubai, UAE
Email: *andreas.kara@hotmail.com

Abstract
The purpose of this paper is to understand the potential of traditional and non-traditional statistical techniques to predict dynamic hotel room prices. Four forecast models were employed: the simple moving average, the autoregressive integrated moving average (ARIMA), the radial basis function (RBF), and the support vector machine (SVM). This research is based on an empirical study of data obtained from the company Smith Travel Research (STR). The economic predictors were obtained from other reliable sources such as the World Bank and the World Tourism Organization. This study agreed with existing literature on the ability of machine learning to predict hotel room prices precisely. Given the complexity of the hotel industry, the effect of external economic predictors was tested in the model. The challenge lay in dealing with the mixed frequencies observed in the collected data. This research is designed to add an innovative approach to the existing literature on machine learning in the hotel industry in the Middle East and North Africa (MENA) region. Some of the machine learning techniques used in this study constitute a contribution to the research conducted in this region. This creates a bridge between many academic disciplines such as computer science, economics, and marketing. Small hotel operators should benefit from this research when setting strategies as well as in using the model to set their relative room prices.

Keywords
Hotel Forecasting, Analytics, Machine Learning

1. Introduction

According to the World Tourism Organization, tourism is considered a sector of
Tourism accounts for 10% of world GDP. It contributes to world societies by promoting cultures, as well as by adding more jobs to the economy. In 2016, 1.2 billion tourists were recorded worldwide and this number is expected to grow to 1.8 billion (50% growth) by 2030 [1]. This research paper focuses on one aspect of the tourism and hospitality sector, which is accommodation and, more specifically, hotels. Practitioners as well as researchers have been motivated to study the dynamic prices of hotel rooms and understand the determinants of these changes and, accordingly, be able to forecast prices more effectively. Akm [2], Drew et al. [3], Hassani et al. [4], Padhi and Aggarwal [5], Yang et al. [6], Youn and Gu [7], Jovanovic et al. [8], Uysal [9], and Magnini et al. [10] have all forecasted hotel or tourism demand using neural networks jointly with support vector analysis, the autoregressive integrated moving average (ARIMA), logistic regression, fuzzy goal programming, and decision trees.

This study introduces a novel approach to hospitality forecasting in this region and more specifically to hotel rooms’ average daily rates (ADR) based on eight major cities in the Middle East and North Africa (MENA) region. These methods include linear models (simple moving average and ARIMA) and non-linear models (radial basis function (RBF) and the support vector machine (SVM). Research on hotels and advanced machine learning forecasting techniques is very limited in this region. This study adds to the literature in this field.

The rest of the paper is organized as follows: the second section reviews the historical literature on Tourism and Hospitality studies and the different models used for forecasting; the third section covers the problem definition and research objective of this study; the fourth section highlights the conceptual framework which would provide good visuals for the study objective and hypothesis; the fifth is the methodology section which covers the data collection part, a detailed data analysis, a list of the key variables and the models used in the research; the sixth section covers the statistical performance evaluation and measures which were employed to select the best model. Finally, a conclusion section covers a summary of the study outcomes and provides recommendation for future studies.

2. Literature Review

In the tourism and hospitality industry, it is very important to understand the variables affecting demand and eventually the performance of the industry (or a particular hotel or restaurant). In the past, managers and decision-makers relied on simple forms of data analysis (simple linear regression or multiple regression) to investigate the influence of those variables. More recently, big data analytics has unlocked the potential for studying complex business situations to understand the correlation between variables or causes and eventually forecast performance. Machine learning techniques have made it possible to analyze tens or hundreds (even thousands) of data variables that are either stored or live-streamed to help shape a time bound decision-making process.

Hospitality and tourism studies using machine learning to predict demand have gained momentum lately. Simple neural networks (NN) (or artificial neural networks) have been used to forecast hotel demand and prices. For instance, Kusum et al. [11] used machine learning models to predict hotel demand in India. Similarly, Al-Masri et al. [12] applied machine learning algorithms to forecast hotel demand in Jordan. These studies demonstrate the potential of using machine learning techniques in the hospitality industry to improve decision-making processes and enhance business performance.
networks (ANN)) have been the most-used machine learning technique. Philips et al. [11], Pattie and Snyder [12], Govers et al. [13], and Law [14] have all used NN or ANN to forecast tourism or hotel demand using historical tourists’ arrival data or room occupancy data. Some researchers have combined neural networks with other ML techniques, either to pick the best model or to improve the performance of the model.

Others have explored different machine learning techniques as well. Vu et al. [15], Tkaczynski et al. [16], Toral et al. [17], Dolnicar and Leisch [18], Brochado et al. [19], and Geetha et al. [20] have used clustering while dealing with market or customer segmentations and consumer behaviors. Hadavandi et al. [21], Yu and Schwartz [22] and Sohrabi et al. [23] have used fuzzy systems to predict hotel demand using arrival data. Fuzzy systems are also widely used in planning and decision-making in retail and banking. Li and Sun [24] used support vectors to predict firm failure using financial and non-financial data, while Chen and Wang [25] used the support vector technique to forecast the demand. Pantano et al. [26] used tourist attraction characteristics and the random forest method to predict tourist response, while Shapoval et al. [27] used inbound visitors numbers and the decision tree technique to develop effective destination marketing.

Rong et al. [28], Xiang et al. [29], Li et al. [30], Yang et al. [31], Guo et al. [32], Versichele et al. [33], Sun et al. [34], Athanasopoulos et al. [35], Schwartz and Hiemstra [36], Yang et al. [37], Tussyadiah and Wang [38], Pereira [39], Liu et al. [40], Lim et al. [41], Wu et al. [42], Zhang and Zhang [43], and Chiu et al. [44] have all introduced new “unconventional” variables while studying hospitality performance. The results are evidence that machine learning techniques are extremely powerful in enhancing the accuracy of analysis and prediction. With a very globalized industry such as the hospitality and tourism industry, machine learning has indeed proven its potential.

3. Problem Definition and Research Objective

Prior studies have focused mostly on tourism demand in countries, that is, at the macro level. Tourism arrivals have been extensively used as a forecasting component and most of the studies are country-specific. Macroeconomic factors take longer time to respond to changes of what is known as the lag effect. On the other hand, firm specific microeconomic factors which represent the fundamental factor model are much faster to respond to changes and are in the control of the hotel managers/owners. They often lead to significant characterization of the dependent variable under investigation [45]. The objective of this study is to firstly combine macro and micro elements while studying dynamic hotel prices. Several cities from the MENA region, which, it is assumed, go through the same economic effects, are included in the study. In addition, this study explores the benefits of using machine learning techniques in forecasting.

4. Research Framework and Proposed Hypotheses

Hotel practitioners have used inventory planning and pricing as major inputs in
M. Al Shehhi, A. Karathanasopoulos

their revenue management systems, which has led to successful hotel performance [39]. Lee [46] found while measuring hotel room rates that prices were affected by both internal (hotel-specific) and external (economic) factors.

Driven by literature findings, hotel attributes have a direct effect on hotel performance. This leads to the first hypothesis:

H1: Big data on hotels leads to better price prediction.

Moreover, various economic factors were investigated in hotel performance studies. This has led to the inclusion of economic factors as a moderating effect to be tested in this study:

H2: Economic factors moderate the relation between hotels’ attributes and hotel room pricing.

The following diagram (Figure 1) provides a visual representation of the study direction.

5. Methodology

Data collection and analysis

The daily hotel data used in this research paper came from STR and covered eight cities in the MENA region (at our request): Dubai, Jeddah, Manama, Muscat, Kuwait, Beirut, Amman, and Sharm El Sheikh. The hotels were split into three categories (Luxury and Upper Upscale Class, Upscale and Upper Midscale Class, Midscale and Economy Class). However, due to some missing data and the need for uniform analysis across all eight cities, it was decided to deal only with the Luxury and Upper Upscale class since data were available for this class.
for all cities. The sample contained 2800 observations of daily room sales covering the period between January 2010 and August 2017. The data set was split in the following way (Table 1).

The data split is done to test the models’ performance in predicting unknown observations (test data is part of unsupervised learning) after determining all network/model parameters using the training data.

Economic variables were obtained from other credible sources such as the World Bank, the World Tourism Organization, the World Economic Forum and the US Energy Information Administration. The only challenge was that most of these data had different frequencies (monthly, yearly, and once every two years). To deal with different frequencies, international tourist arrival data were converted to daily rates by dividing the annual rate by 365, while country-level annual GDP growth percentage rate, inflation rate (average consumer prices), oil price (WTI and Brent), and index data on each country’s business environment, safety and security, health and hygiene, human resources, and labor market were all converted to daily rates by maintaining the same rate throughout the year. The aim was to use these variables and measures to gain insight into the determinants of hotel room prices that would help us, and eventually decision-makers, to predict these prices with more accuracy. Innovative ways of handling mixed data frequencies could be an opportunity for future research. This study would also be a good piece of research to validate the hotel performance determinants (HPD) model suggested by Assaf et al. [47]. The table below (Table 2) represents a list of the variables found in the literature that we utilized in our study based on data availability/accessibility.

6. Models in the Research

This research is based on predicting dynamic hotel room prices based on the selection of the best forecasting model. These models are linear (simple moving average and ARIMA) and non-linear (RBF and SVM) in form. The goal is to compare the above-mentioned models using the model performance measures to determine the best model or a combination of them.

6.1. Time Series Forecast

Using ADR values from the years, time series forecast would help in predicting the future ADR of hotels based on historical data. The main goal is to find a model with a better fit for the data, hence reducing the noise or error. The models that we used are the simple moving average and the Box–Jenkins ARIMA model. These models are widely used in tourism and hospitality research.

6.2. Simple Moving Average

The simple moving average method uses the average of previous n-periods as a forecast value [59].
Table 1. Data sample.

<table>
<thead>
<tr>
<th>Period</th>
<th>Total Observation</th>
<th>Start Date</th>
<th>End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total data set</td>
<td>2800</td>
<td>1-Jan-10</td>
<td>31-Aug-17</td>
</tr>
<tr>
<td>Training set</td>
<td>2191</td>
<td>1-Jan-10</td>
<td>31-Dec-15</td>
</tr>
<tr>
<td>Test set (out of sample)</td>
<td>609</td>
<td>1-Jan-16</td>
<td>31-Aug-17</td>
</tr>
</tbody>
</table>

Table 2. Summary table of variables used in tourism and hospitality literature.

<table>
<thead>
<tr>
<th>Variables found in the literature</th>
<th>Main area of study</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hotel performance determinants (HPD) framework:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ General economic conditions: tax rate, gross domestic product growth, exchange rate, fuel price, disposable income.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Safety and stability: number of natural disasters, crime rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Market supply and demand: number of other accommodation providers, number of competing hotel rooms, performance of the overall hotel industry, domestic tourist arrivals per capita, international tourist arrival per capita.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Regulation and support: T &amp; T government expenditure, transparency of government policies, fairness in government spending, visa requirements.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Quality of infrastructure: quality of air transport infrastructure, quality of port infrastructure, quality of road infrastructure.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Labor skills and training: ease of hiring foreign labor, quality of the educational system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hotel performance</td>
<td>[47]</td>
<td></td>
</tr>
<tr>
<td>PESTELIR model:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Political conditions: regional and local policies supporting website development, public subsidies for ICT implementation, tourism promotion.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Economic conditions: macro- and microeconomic trends determining business investments: economic cycles, changing origins of most tourists, development of new destinations.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Socio-demographic conditions: changes in consumption habits, new customer needs, new market segments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Technological conditions: technology-related innovations allowing the offer of new hotel products and services or improvements in productive processes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Ecological conditions: pollution, environmental policies regulating protected areas.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Legal conditions: legislation regarding construction, legislation for protection of personal data, legislation for protection of personal data.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Infrastructure conditions: changes in tourism infrastructure (airports, convention centers, new communication infrastructures).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Human resource conditions: basic features of the workforce to be accessed by hotels and other tourism organizations: training, motivation, languages.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hotel performance</td>
<td>[48]</td>
<td></td>
</tr>
<tr>
<td>Tourism demand</td>
<td>[49]</td>
<td></td>
</tr>
<tr>
<td>Tourism growth</td>
<td>[50]</td>
<td></td>
</tr>
<tr>
<td>Hotel profitability</td>
<td>[51]</td>
<td></td>
</tr>
<tr>
<td>Tourism demand</td>
<td>[52]</td>
<td></td>
</tr>
<tr>
<td>Tourism sector, oil prices and economic growth</td>
<td>[53]</td>
<td></td>
</tr>
</tbody>
</table>
ARIMA \((p, d, q)\) consists of the autoregressive AR\((p)\), the moving average MA\((q)\) and \((d)\), which represents the order of differencing used to achieve stationarity. ARIMA is one of the models most widely used to forecast with time series data \([60]\).

\[
\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}
\]

6.3. Machine Learning Models

6.3.1. RBF

RBF is one of the most widely used neural network techniques. Used for classification as well as regression, the RBF model is a feed-forward neural network that is based on three layers: input, hidden and output \([61]\). RBF models gained interest due to their advantage in achieving faster convergence with fewer errors while also being reliable (Moradkhani et al., 2004) \([62]\).

\[
f(x) = \exp\left(\frac{-x - C^2}{2\sigma^2}\right)
\]

where \(C\) represents the center and represents the width of the neuron or the radius (Wei, 2012) \([61]\).
6.3.2. SVM/SVR

SVM was introduced by Vapnik in the early 1990s. SVM is a statistical technique used widely for classification and more recently for regression (support vector regression (SVR)). Unlike other models, SVM aims to minimize the generalized error (structural risk minimization). When visualized, SVM works to maximize a hyperplane that separates two classes (or more). SVR is another version of SVM that was proposed by Drucker, Burges, Kaufman, Smola, and Vapnik [63]. SVM could also work in higher dimensions if a kernel function is applied, which allows SVM to solve non-linear equations [25].

\[ f(x) = \omega \phi(x) + b \]  

(3)

7. Data Analysis

As a first step, we ran a descriptive statistics analysis for the data which highlights the mean, max, min and standard deviation of the variables used in the study. The data used in this study represents daily observations obtained from STR which provided insights of the internal or microeconomic factors used by the industry to study the hospitality sector. Other macro-economic factors, which appeared also in several studies within the tourism and hospitality field; were obtained from different sources such as the world bank the World Tourism Organization, the World Economic Forum and the US Energy Information Administration. However, those factors where country based since capturing them at city level and on daily basis was impossible and out of the scope of this study. The table below (Table 3) provides a summary of the descriptive statistics for significant variables generated from Dubai Luxury upper data which was produced using IBM SPSS:

With skewness that is close to zero and less than 1 for most variables, this indicated that though we are dealing with very dynamic environment, yet data is normally distributed with means around zero.

Using documented steps in literature for each proposed model in this study, the data was then used in each model to produce the forecasts and to carry further analysis. The aim is to compare models and choose the best model for ADR prediction (or perhaps a combination of models) based on

\[ MAD = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|, \]

\[ MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} * 100\%, \]

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}, \]

and \[ MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]

performance criteria.

As a first step in the time series analysis, the data were plotted to check visual-
ly for any seasonal trends throughout the year (refer to Figure 2 for ADR). The aim was to regulate this seasonality or make the data stationary in order to be able to explain the data using the autoregressive model, ARIMA. Many of these cities showed acceptable stationarity in data while some (i.e. Jeddah) showed some increasing trends over a number of years, which necessitated some treatment of the data to make the mean constant. As a result, and to maintain uniformity, first-order differencing for all cities’ data was considered to make those data stationary following the 1970 Box-Jenkins method (see Figure 3) [64].

After dealing with stationarity, AIC and BIC tests were employed. A different combination of ARIMA models for each city was tested and based on the tests criteria the best model was selected for data analysis.

8. Statistical Performance

The following table (Table 4) represents the result of the models’ performance measures for each city as a measure of forecasting accuracy.

8.1. Conventional Techniques

By employing the ARIMA and simple moving average techniques to forecast future room rates, the study found that the simple moving average performed poorly

![Figure 2](image2.png)

**Figure 2.** Average daily rates for Luxury and Upper Upscale hotels in selected cities in the MENA region (StataSE13).

![Figure 3](image3.png)

**Figure 3.** D1 of average daily rates for Luxury and Upper Upscale hotels in selected cities in the MENA region (StataSE 13).
### Table 3. Descriptive statistics for Dubai Luxury & upper sample data.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADR</td>
<td>2800</td>
<td>154.98</td>
<td>683.45</td>
<td>275.23</td>
<td>74.56</td>
<td>0.830</td>
<td>1.530</td>
</tr>
<tr>
<td>WeekDay</td>
<td>2800</td>
<td>1.00</td>
<td>7.00</td>
<td>4.00</td>
<td>2.00</td>
<td>0.000</td>
<td>-1.250</td>
</tr>
<tr>
<td>Occupancy</td>
<td>2800</td>
<td>36.10</td>
<td>98.27</td>
<td>75.73</td>
<td>12.91</td>
<td>-1.114</td>
<td>1.055</td>
</tr>
<tr>
<td>RevPar</td>
<td>2800</td>
<td>58.35</td>
<td>668.45</td>
<td>214.79</td>
<td>83.49</td>
<td>0.598</td>
<td>1.126</td>
</tr>
<tr>
<td>Demand</td>
<td>2800</td>
<td>9973.53</td>
<td>39681.71</td>
<td>24739.01</td>
<td>6055.69</td>
<td>-0.165</td>
<td>-0.599</td>
</tr>
<tr>
<td>Inflation Rate Avg. Consumer Prices</td>
<td>2800</td>
<td>0.70</td>
<td>4.10</td>
<td>1.72</td>
<td>1.08</td>
<td>1.181</td>
<td>0.341</td>
</tr>
<tr>
<td>Oil WTI</td>
<td>2800</td>
<td>26.19</td>
<td>115.39</td>
<td>76.22</td>
<td>23.49</td>
<td>-0.364</td>
<td>-1.373</td>
</tr>
<tr>
<td>Oil Brent</td>
<td>2800</td>
<td>26.01</td>
<td>128.14</td>
<td>83.49</td>
<td>28.46</td>
<td>-0.273</td>
<td>-1.488</td>
</tr>
<tr>
<td>Business Environment</td>
<td>2800</td>
<td>5.14</td>
<td>5.90</td>
<td>5.47</td>
<td>0.32</td>
<td>0.480</td>
<td>-1.557</td>
</tr>
<tr>
<td>Safety and Security</td>
<td>2800</td>
<td>5.13</td>
<td>6.60</td>
<td>5.66</td>
<td>0.69</td>
<td>0.634</td>
<td>-1.592</td>
</tr>
<tr>
<td>Health and Hygiene</td>
<td>2800</td>
<td>4.88</td>
<td>5.40</td>
<td>5.09</td>
<td>0.19</td>
<td>0.164</td>
<td>-1.442</td>
</tr>
<tr>
<td>Human Resources and Labour Market</td>
<td>2800</td>
<td>5.15</td>
<td>5.65</td>
<td>5.47</td>
<td>0.22</td>
<td>-0.625</td>
<td>-1.565</td>
</tr>
<tr>
<td>ICT Readiness</td>
<td>2800</td>
<td>4.00</td>
<td>6.10</td>
<td>5.10</td>
<td>0.72</td>
<td>-0.511</td>
<td>-1.062</td>
</tr>
</tbody>
</table>

### Table 4. Performance for all cities.

<table>
<thead>
<tr>
<th>Forecast</th>
<th>ARIMA Simple Average</th>
<th>RBF</th>
<th>SVM</th>
<th>Forecast</th>
<th>ARIMA Simple Average</th>
<th>RBF</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amman</td>
<td>MAD 0.0547 0.0675 0.0392 0.0351</td>
<td>MAD 0.0653 0.0772 0.0423 0.0412</td>
<td>Manama</td>
<td>MAD 0.0623 0.0712 0.0543 0.0459</td>
<td>MAD 0.0765 0.0843 0.0663 0.0431</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSE 0.0531 0.06873 0.0345 0.0443</td>
<td>MSE 0.0865 0.0976 0.0652 0.0612</td>
<td></td>
<td>MSE 0.0588 0.0741 0.0342 0.0312</td>
<td>MSE 0.0234 0.0386 0.0123 0.0201</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 156.34% 181.77% 142.67% 123.45%</td>
<td>RMSE 189.61% 187.56% 155.41% 123.71%</td>
<td></td>
<td>RMSE 146.98% 156.72% 145.97% 123.91%</td>
<td>RMSE 205.54% 286.34% 190.75% 103.65%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAPE 0.029 0.0376 0.0287 0.0221</td>
<td>MAPE 0.0433 0.0541 0.0412 0.0421</td>
<td></td>
<td>MAPE 0.0345 0.0441 0.0296 0.0257</td>
<td>MAPE 0.0551 0.0589 0.0421 0.0382</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beirut</td>
<td>MAD 0.0623 0.0712 0.0543 0.0459</td>
<td>MAD 0.0765 0.0843 0.0663 0.0431</td>
<td>Muscat</td>
<td>MAD 0.0435 0.0534 0.0356 0.0345</td>
<td>MAD 0.0543 0.0567 0.0435 0.0421</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSE 0.0588 0.0741 0.0342 0.0312</td>
<td>MSE 0.0234 0.0386 0.0123 0.0201</td>
<td></td>
<td>MSE 0.0653 0.0851 0.0552 0.0433</td>
<td>MSE 0.0786 0.0854 0.0321 0.0211</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE 146.98% 156.72% 145.97% 123.91%</td>
<td>RMSE 167.42% 151.62% 123.98% 124.23%</td>
<td></td>
<td>RMSE 204.87% 234.76% 201.76% 197.61%</td>
<td>RMSE 151.62% 123.98% 124.23% 124.23%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAPE 0.0345 0.0521 0.0289 0.0222</td>
<td>MAPE 0.0421 0.0444 0.0315 0.0301</td>
<td></td>
<td>MAPE 0.0345 0.0521 0.0289 0.0222</td>
<td>MAPE 0.0421 0.0444 0.0315 0.0301</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Continued

<table>
<thead>
<tr>
<th></th>
<th>Jeddah</th>
<th>Sharm</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD</td>
<td>0.0734</td>
<td>0.0612</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0652</td>
<td>0.0665</td>
</tr>
<tr>
<td>RMSE</td>
<td>276.66%</td>
<td>323.54%</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.0665</td>
<td>0.0885</td>
</tr>
</tbody>
</table>

according to the performance measures, while ARIMA was a significantly better predictor. From the above table, it appears that Amman, Dubai, and Sharm El-Shaikh have the lowest errors in models compared to other cities, while Manama produced the highest errors. Overall, ARIMA performed better than the simple moving average method in term of forecasting accuracy using conventional techniques.

8.2. Innovative Techniques

One of the contributions of this paper is the use of innovative machine learning tools to forecast room prices. Both RBF and SVM were utilized for prediction, which resulted in significant improvements in performance. The inclusion of external economic factors could also be one reason why these models outperformed the conventional models.

When comparing the forecasting accuracy of different models and for the eight different cities, it was found that SVM and RBF performed better than ARIMA or the simple moving average. The results show the machine learning technique’s superiority in prediction compared to conventional forecasting models.

9. Conclusions

The use of innovative tools in hotel performance forecasting would help researchers as well as practitioners in planning effectively. Hotel internal attributes positively affect hotel performance and more specifically prices. External economic factors moderate the relationship between the hotel attributes and hotel performance.

The main objective of the study was to predict hotel room prices using new tools. The study shows that SVM is the leading model in “luxury and upscale” hotel room price forecasting, followed by RBF and then ARIMA, while the simple moving average is found in this study to be the inferior model.

Machine learning is insufficiently studied in the hotel and tourism sector in the MENA region, and this study adds to the academic literature. Due to their abundance and rapid development in recent years, future studies could explore other machine learning and artificial intelligence models and compare their performance against traditional models. Other performance measures such as precision and speed of returning results could be used for model evaluation as well. Given the dynamic environment within the tourism and hospitality sector,
policy makers and hotel operators could use these tools to maintain their strategic lead. SVM model or a combination of forecasting models can be utilized to forecast short-term and long-term market direction based on the model strength.

References


https://doi.org/10.1016/j.jhydrol.2004.03.027

https://doi.org/10.1016/j.knosys.2017.01.033