ANALYSIS AND FORECAST OF CROATIAN TOURISM DEMAND SEASONALITY

Krešo Mihalinčić
Iva Mrša Haber

Abstract
Purpose – Investigate the issue of seasonality of Croatian tourism demand.
Design – We have established the seasonality of Croatian tourism as a whole by measuring monthly overnight stays (“overnights”) using advanced automated as well as original methods. We also investigated land vs. coast contributions to Croatian hospitality industry.
Methodology – Seasonal ARIMA (SARIMA) was utilized to analyze above data and forecast future demand in terms of the “overnights” time series.
Approach – Having proven the standard SARIMA automated methods deficient, we developed an alternative approach and proved its reliability by a comparison to actual data.
Findings – The smoothing nature of ARIMA forecasting algorithm leads to a mitigation of seasonal effects in its forecast. Therefore, an alternative approach is called for.
Originality of the research – Due to the apparent shortcomings of ARIMA forecasting we developed an alternative model that pays due attention to the strong seasonal effects of tourism demand common to Mediterranean countries. Nonetheless, we anchored our approach in some existing concepts (e.g. random walk). Still, our concept of “tiers” that mimic standard quarters but preserve the information on crucial differences between peak and shoulder months is, for all our research, original. Our results demonstrate clearly that we have found a fine balance between automated methods and expert judgment.
Keywords tourism, demand, Croatia, ARIMA modeling, seasonality, forecasting

1. INTRODUCTION

Croatian tourism dates back to the mid-19th century. Croatian coast has since been one of the most visited tourist destinations on the Mediterranean. After the war years of 1991 to 1995 more attention has been paid to the development of continental tourism. Nonetheless, coastal tourism remains the main driver of Croatian tourism and Croatian economy as a whole. The total contribution of travel and tourism to Croatian GDP in 2016 was 24.7% and their total contribution to employment was 23.4%, according to Travel and Tourism Economic Impact – Croatia, 2017 (WTTC, 2017).

Now, this dependence on the coastal tourism comes at a price of high seasonal effects. Butler (Butler, 2014) defines seasonality as a “temporal imbalance in the phenomenon of tourism, which may be expressed in terms of dimensions of such elements as numbers of visitors, expenditure of visitors, traffic on highways and other forms of transportation, employment, and admissions to attractions”. They emphasize a great economic importance of measuring seasonality; in particular: “The tourism industry – more specifically hotels – has a strong desire for seasonal extension. The main reason is overcapacity outside the peak season.”
Long-term international tourism demand forecasting serves to avoid inappropriate investments and improves the planning of economic growth (Baldigara and Mamula, 2012). Still, the unreliability of tourism demand forecasts is rarely listed among the negative effects of seasonality despite the importance of the quality of forecasts for the planning of economic activities. (Vergori, 2017). Vergori also shows that the stronger the seasonality is, the less reliable the forecasts are. In this paper, we devise an alternative methodology to improve on the generally accepted and widely used seasonal ARIMA approach, cf. Apergis, N., Mervar, A., Payne, J.E., (2017).

As we will see, apart from the peak summer months, tourism demand remains largely dormant in winter. While typical of any Mediterranean country, this issue affects Croatia to a relatively high degree, cf. Kozić (2013), leaving Croatian economy especially vulnerable to general economic trends and political developments. In the second part of section 2 we review some general concepts of seasonal phenomena and provide a historic overview of the related research. We begin with a brief look at our present paper.

Section 3 presents an immediate, descriptive analysis of our underlying time series representing monthly overnight stays between 2010 and 2014 broken down by coast and continent. In further text we denote these monthly overnight stays as „overnights“. We introduce several relevant measures that appropriately quantify the observed phenomena. In order to better reflect the in-year seasonal effects we introduce a concept of “tiers”. These tiers are essentially subtler groupings of months that better reflect monthly seasonal information than the often used standard quarters.

In section 4 we follow the common seasonal ARIMA methodology. This approach was used in Baldigara and Mamula (2012) to analyze Croatian quarterly seasonality. In contrast, we focus on monthly granularity. The tiers from section 3 now clearly illustrate deficiencies of ARIMA forecast at this level.

Motivated by these deficiencies, in section 5, we develop an alternative model to forecast tourism demand that pays special attention to the observed nature of monthly seasonality. The concluding comparison of our forecast to the corresponding data from The Croatian Institute for Statistics demonstrates a remarkable correspondence between our forecast and the actual results. In particular, it preserves the important monthly seasonal relationships affecting the whole of the Mediterranean area.

2. LITERATURE REVIEW

Research on tourism demand has grown rapidly from the last decade of the 20th century, cf. Jelušić (2017). Nonetheless, Witt and Witt (1995) list 114 papers on tourism demand forecasting models developed before 1995. Still, “there is no scientific theory on tourism seasonality,” according to Baum and Lundtorp (2001). As our main concern in this paper is seasonality, we focus on this aspect of tourism demand analysis with an additional emphasis on the Mediterranean area.
Jelušić also notes that tourism demand models for quantitative forecasting can be divided into time series and econometric models. It is important to note that econometric models, in one way or another, eliminate seasonality from their typically regression-based analyses, cf. Thomas and Wallis (1971). For an analysis of seasonality per se, therefore, it is natural to select the time series approach, e.g. Baldigara and Mamula (2013). In fact, recent research has shown that no more than a single variable is truly needed for seasonality analyses. For instance, observations of Bigović (2011) concerning seasonality in Montenegro suggest that holistic approaches are not required to understand seasonality; one measure suffices.

Works of Syriopoulos (1995) and Syriopoulos and Sinclair (1993) as well as Papatheodorou (1999) use econometric models to analyze the demand for international tourism in the Mediterranean region. According to Kulendran with Witt and Wong (Kulendran and Wong, 2005, also Kulendran and Witt 2003), the research on tourism demand and seasonality in particular has moved to more advanced time series analyses such as ARIMA and conditional volatility models. They emphasize the importance of the alignment of a selected model to the nature of seasonality; however, there does not seem to be a clear method for such selection.

Song and Li (2008) reviews the studies on tourism demand modeling and forecasting since 2000. They show that there was no one model that consistently outperformed other models in all situations. This is further confirmed by Li (2005) and Song and Witt (2003); a unique model across different origin countries was shown inadequate although ARIMA variations are adequate in some situations. Song, Li and Witt (2010) expressly state that for a seasonal analysis and attempts to improve the number of overnights in the off-season, one is best advised to use univariate models that clearly reflect seasonality. This was also shown by García-Ferrer and Queralt (1997) in case of Spanish tourism demand.

Searching through WOS, SCOPUS and Science Direct databases, we have found few articles on Croatian tourism demand forecasting. Šergo, Matošević and Zanini-Gavranic (2016) studied the seasonality patterns and determined their impact across numerous countries in Europe between 1995-2014 via censored panel regression. Baldigara and Mamula (2012) investigates tourism statistics in Croatia and forecasts future challenges for Croatian tourism via ARIMA methodologies. Apergis, Mervar and Payne (2017) show that SARIMA models with Fourier transformations outperform other models in 20 Croatian counties and the City of Zagreb.

Kozić (2013) studies seasonal patterns of the Euro-Mediterranean region comparable to Croatia in terms of geographic location, climate and type of tourism. Apart from showing that the model of our paper applies to such countries as well as to Croatia, his work indicates that a relatively high seasonality in Croatia leaves ample room for improvement of hospitality offering. Pletikosa (Pletikosa, 2015) gives a low overall rating to Croatian tourism in comparison to other Mediterranean countries and suggests the need for investments in order to increase the competitiveness of Croatian tourism, both at macro and micro level.

The summary of our results will be seen in the conclusion of our article.
3. DESCRPTIVE ANALYSIS

As mentioned in the introduction, our main measure \( x \) is “overnights” denoting the number of overnight stays booked in Croatian hotels in a given month. The time span is January 2010 to December 2014. These data are broken down by continent and coast. Obvious seasonal effects in these data (Chart 1) combined with the natural economic interest in the growth of tourism demand suggest an introduction of a measure “Year over Year (Y/Y) Growth by Month”, defined as

\[
\text{Year/Year Growth by Month} = \frac{x_{n+12} - x_n}{x_n}.
\]

Thus, we calculate the percentage change between the overnights in a given month of a given year and the same month one year prior.

Chart 1: Year/Year Growth by Month at the Continent

In the years prior to 2010 continental overnights reach some 200,000 levels in the peak months. However, in 2010, August overnights grow by 250% to 680,000 and soon stabilize at approximately 1,000,000. Thus, 2010 is a turning point in our time series; in fact, from that year on, continental contribution to total overnights stabilizes at 5% in the peak months (compared to approx. 0% before). Similar arguments speak for a stabilization of data since 2010 at the coast as well (Chart 2). Later on, we will prove these data stationary.
Often, as in Baldigara and Mamula, analyses are performed at the quarterly level. Having observed strong differences between months, we introduce the concept of “tiers” that mimic standard quarters but better reflect monthly information. The nature of data makes the following categorization sensible.

Table 1: Tiers definition

<table>
<thead>
<tr>
<th>Tiers</th>
<th>Overnights</th>
<th>Months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coast</td>
<td>Continent</td>
</tr>
<tr>
<td>1</td>
<td>10 000 000 – 25 000 000</td>
<td>900 000 – 1 200 000</td>
</tr>
<tr>
<td>2</td>
<td>5 000 000 – 10 000 000</td>
<td>500 000 – 900 000</td>
</tr>
<tr>
<td>3</td>
<td>1 000 000 – 4 000 000</td>
<td>200 000 – 500 000</td>
</tr>
<tr>
<td>4</td>
<td>0 – 1 000 000</td>
<td>100 000 – 200 000</td>
</tr>
</tbody>
</table>

Source: Authors
We will refer to tiers 1 and 2 as “higher tiers” and tiers 3 and 4 as “lower tiers”. Of particular interest is the fact that both the membership of months in tiers and their order – also indicated in table - within tiers are constant, apart from a few exceptions at the continent in earlier years. We are going to measure the in-year seasonal difference at the level of tiers and express it via:

$$\frac{\text{Max} T_{n+1}}{\text{Min} T_n}$$  \hspace{1cm} (2)

Note that these measures reflect the common-sense understanding of Croatian (and Mediterranean in general) seasonality (July and August – high season, May and June – pre-season etc.) Tables 2. and 3. show the evolution of tier differences.

### Table 2: Tiers growth at the coast

<table>
<thead>
<tr>
<th>Tiers growth</th>
<th>Year</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>MinT1/MaxT1</td>
<td></td>
<td>91%</td>
<td>88%</td>
<td>89%</td>
<td>88%</td>
<td>82%</td>
</tr>
<tr>
<td>MaxT2/MinT1</td>
<td></td>
<td>36%</td>
<td>43%</td>
<td>41%</td>
<td>41%</td>
<td>47%</td>
</tr>
<tr>
<td>MaxT3/MinT2</td>
<td></td>
<td>47%</td>
<td>36%</td>
<td>42%</td>
<td>48%</td>
<td>41%</td>
</tr>
<tr>
<td>MaxT4/MinT3</td>
<td></td>
<td>33%</td>
<td>23%</td>
<td>27%</td>
<td>37%</td>
<td>26%</td>
</tr>
</tbody>
</table>

Source: Authors

There is a trend of increasing seasonality in top tier (tier 1) since 2010 as the peak month of August grows relatively to the (flat) second-strongest month of July. There is some evidence of June catching up on July. Seasonality seems to be increasing between tiers T3/T2 and T4/T3 (as ratios drop from 47% to 41% and 33% to 26%, respectively).

### Table 3: Tiers growth at the continent

<table>
<thead>
<tr>
<th>Tiers growth</th>
<th>Year</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>MinT1/MaxT1</td>
<td></td>
<td>93%</td>
<td>84%</td>
<td>85%</td>
<td>88%</td>
<td>82%</td>
</tr>
<tr>
<td>MaxT2/MinT1</td>
<td></td>
<td>64%</td>
<td>66%</td>
<td>55%</td>
<td>57%</td>
<td>60%</td>
</tr>
<tr>
<td>MaxT3/MinT2</td>
<td></td>
<td>91%</td>
<td>70%</td>
<td>74%</td>
<td>82%</td>
<td>64%</td>
</tr>
<tr>
<td>MaxT4/MinT3</td>
<td></td>
<td>90%</td>
<td>88%</td>
<td>88%</td>
<td>76%</td>
<td>87%</td>
</tr>
</tbody>
</table>

Source: Authors

As in the case of the coast, there is a trend of increasing seasonality in top tier since 2010. There is some evidence of an increasing gap between tiers 3 and 2. The trend between (and separation of) March-April due to different timing of Easter remains unclear.
An analysis of total overnights, understandably, shows the same order of tiers growth as the coast. Still, during slower months, the contribution of continental tourism to the total ranges from 30% (March) to 50% (January). This leads to a mitigation of seasonality of approx. 6% in seasonal drop when total is compared to coast alone.

4. STATIONARITY AND ARIMA FORECAST

Despite perception, possibly obtained from e.g. Chart 1, we cannot establish a clear growth trend with certainty. Looking at the critical month of August, Y/Y growth is either so small that it can easily come from measurement error (coast, 1% to 6%). For our present purposes, we dispense with mathematical rigor and define stationarity as an absence of a clear growth trend. For concepts and methods of this section, we refer once more to Baldigara and Mamula (Baldigara and Mamula, 2012). The ADF stationarity test actually proves stationarity, as does the selection of SARIMA model (2,0,3)(1,0,1) as the best fitting model (automatic ARIMA forecasting of EViews 9).

Chart 3: ARIMA forecast for 2015-total

A number of things point at inadequacy of this forecast (e.g. occurrence of negative numbers). The most striking is certainly the suspiciously low forecast level of overnights; on top of that, July has taken over the top spot from August. Based on past data, both developments are highly unlikely. A convincing summary of this forecast’s failure is provided by our tier-concept.
Table 4: Tiers according to ARIMA forecast (2015-total)

| Total tiers growth | Year       | Year       | Year       | Year       | Year       | Forecast
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2010</td>
<td>2011</td>
<td>2012</td>
<td>2013</td>
<td>2014</td>
<td>2015</td>
</tr>
<tr>
<td>MinT1/MaxT1</td>
<td>91%</td>
<td>88%</td>
<td>89%</td>
<td>88%</td>
<td>82%</td>
<td>92%</td>
</tr>
<tr>
<td>MaxT2/MinT1</td>
<td>37%</td>
<td>44%</td>
<td>41%</td>
<td>41%</td>
<td>47%</td>
<td>99%</td>
</tr>
<tr>
<td>MaxT3/MinT2</td>
<td>49%</td>
<td>38%</td>
<td>44%</td>
<td>50%</td>
<td>42%</td>
<td>99%</td>
</tr>
<tr>
<td>MaxT4/MinT3</td>
<td>41%</td>
<td>30%</td>
<td>34%</td>
<td>46%</td>
<td>32%</td>
<td>38%</td>
</tr>
</tbody>
</table>

Source: Authors

Thus, the striking differences between tiers are all but eradicated. This calls for an alternative approach of a kind highlighted in the introduction. It is worth noting that Baldigara and Mamula identify ARIMA (0,0,0) (1,1,3) as the best fit for their analysis. In simple terms, this suggests that “only seasonality matters” - that is, Year/Year relationships determine the nature of the series rather than in-year trends.

5. AN ALTERNATIVE FORECAST

Motivated by the concluding remarks of the prior section, we recall the well-known models called random walk \( x_t = x_{t-1} + g(t) \) (3) and seasonal random walk \( x_t = x_{t-12} + g(t) \). (4)

For a rigorous treatment of general time series and random walk in particular, we refer the reader to Ghysels and Osborn (Ghysels and Osborn, 2001). For our purposes, we will refer to the term \( g(t) \) as “growth assumption” and base our forecast on a variation of the seasonal random walk via:

\[
\text{overnights\,(month\,in\,2015) = overnights\,(same\,month\,in\,2014) + g(2015)}
\]

(5)

We now design a set of rules-of-thumb – referred to as “decision criteria” – to determine the \( g \)-terms above. These criteria are based on the four observations of the prior growth from 2011 to 2014. Note that, as growth is calculated from the prior year and we start with 2010, this is the longest series available. In fact, as we are interested in trends, earlier years are less relevant than the more recent ones.

The term “oscillating” will denote an absence of a clear trend. This expression occurs in literature; Chiang (Chiang, 1984) points out at a number of definitions and notes that some authors use the word “fluctuating”. For all our purposes we can use the term in its intuitive sense: the values do not show any clear increasing or decreasing trend but vary around some mean value. In this case, our growth assumption is taken as the median of the oscillating values. On the other hand, a consistent three times in the row increase or decrease in Y/Y by month growth will be considered a trend; our growth assumption then becomes the median of the trend.
Although these rules may seem somewhat arbitrary, they are anchored in general time series theory and expertly judgment. As demonstrated through the literature review in introduction, such forecasts often outperform highly sophisticated methods. Our results, as we are about to show, present another convincing contribution to such observations.

Table 5: Y/Y growth at the coast and growth assumption from 2011 to 2014

<table>
<thead>
<tr>
<th>Month</th>
<th>Y/Y Growth by Month</th>
<th>Growth Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2011</td>
<td>2012</td>
</tr>
<tr>
<td>1</td>
<td>-18%</td>
<td>25%</td>
</tr>
<tr>
<td>2</td>
<td>0%</td>
<td>-2%</td>
</tr>
<tr>
<td>3</td>
<td>-20%</td>
<td>29%</td>
</tr>
<tr>
<td>4</td>
<td>27%</td>
<td>-2%</td>
</tr>
<tr>
<td>5</td>
<td>-13%</td>
<td>25%</td>
</tr>
<tr>
<td>6</td>
<td>21%</td>
<td>-3%</td>
</tr>
<tr>
<td>7</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>8</td>
<td>6%</td>
<td>1%</td>
</tr>
<tr>
<td>9</td>
<td>13%</td>
<td>8%</td>
</tr>
<tr>
<td>10</td>
<td>14%</td>
<td>10%</td>
</tr>
<tr>
<td>11</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>12</td>
<td>6%</td>
<td>1%</td>
</tr>
</tbody>
</table>

Source: Authors

We focused on a good growth assumption (based on economic data) from table 5. On the continent, we were mostly faced with oscillating data and median-growth assumptions. Coastal and continental forecasts are summarized below at chart 4, followed by chart 5 showing actual overnights.
Chart 4: Growth model forecast 2015; total

Source: Authors

Comparison of chart 4 and chart 5 shows a high correspondence of our forecast and actual development of the year 2015. An astute reader should immediately notice that a continuation of our procedure to 2016 produces results of the same quality.

Source: Croatian State Institute for Statistics, viewed 22 August, https://www.dzs.hr/
6. CONCLUSION

Let us highlight the three key points of our paper:

- Croatian tourism demand is characterized by strong seasonal effects that are effectively measured in terms of our concept of tiers. The growth of continental tourism has not yet significantly contributed to a mitigation of this behavior.
- Such pronounced seasonality of demand calls for a seasonal approach to forecasting. Some widely used automated techniques – in particular seasonal ARIMA – have to be used with caution and may require expert judgment.
- We have developed such an “expert judgment” approach that can be used by tourism professionals without use of advanced analytical tools and demonstrated its validity.

We conclude the paper by relating our results and methods to the existing literature.

Stevenson (Stevenson, 2007) warns that, while the widely used ARIMA is a useful tool for establishing seasonality in past data, it is not always a reliable tool for forecasting. Montgomery, Jennings and Kulachi (2008) agree. But this does not pertain to ARIMA alone. Armstrong (2001) generally states that “there is a poor correspondence between statistical fit and forecast accuracy for time-series data” for any modelling tool. Armstrong also demonstrates that so-called “naïve approaches” often outperform sophisticated tools (Armstrong, 2001). Indeed, such authorities for automated forecasting as Gilliland (Gilliland, et al. 2016) support this view, emphasizing that “naïve” is not to be taken literally (except for the basic “no-change” assumption). Quite to the contrary, an expert that understands the nature of seasonality is encouraged to devise an appropriate model of his own. Bonilla, Bonilla and Altamira (2006) did just that for Spanish seasonality. García-Ferrer and Queralt, (1997) suggest that predicted annual growth rates as a better alternative to more sophisticated methods. The results of this paper are demonstrably another successful attempt in this direction. On the sum: ARIMA is generally accepted as one of the best methodologies for establishing the existence of seasonality in time series. However, its forecasting power – as that of any automated tool – is questionable.

A review of literature shows that univariate analyses are most appropriate for the analyses of seasonal time series. Nonetheless, there is no tool that consistently outperforms others in all situations. It is this understanding of a specific nature of seasonality that determines the appropriate methodology. Therefore, informed “naïve approaches” are often the way to go.

Our approach in this paper is anchored in the concepts on which advanced numerical methods were built (seasonal random walk). We shall demonstrate the forecasting strength of our method in an out-of-sample analysis. As the foregoing review has shown, the similarity of Croatian tourism demand to other Mediterranean countries makes this methodology applicable to the whole area.
A general recommendation (Butler, 2014) is to give up on attempts to entirely eliminate seasonality but rather work through adaptation and mitigation (utilizing smaller peaks like Christmas and Easter, enriching pre- and post-seasonal offering). Our results as well as Kozić’s point at a very strong and hardly changing seasonality in Croatia. Combined with the aforementioned Pletikosa’s overall low Croatian rating among the Mediterranean countries, this strongly suggests an investment into a mitigation of the overwhelming fallouts from Croatian seasonality. A logical point to start may be to investigate how the strongest performers in the region deal with this issue.

REFERENCES


284
K. Mihalinčić, I. Mrša Haber: ANALYSIS AND FORECAST OF CROATIAN TOURISM DEMAND


Krešo Mihalinčić, PhD, Lecturer
University or Rijeka
Faculty of Tourism and Hospitality Management
Department of Quantitative Economics
Primorska 42, 51410 Opatija, Croatia
Phone: +385-51/294-684
E-mail: kresom@fthm.hr

Iva Mrša Haber, PhD, Assistant
University or Rijeka
Faculty of Tourism and Hospitality Management
Department of Quantitative Economics
Primorska 42, 51410 Opatija, Croatia
Phone: +385-51/294-684
E-mail: ivamh@fthm.hr