THE SEASONAL ADJUSTMENT OF OCCUPANCY VARIABLES:  
THE ITALIAN EXPERIENCE

Lorenzo Cavallo and Maria Teresa Santoro

THEME 4: Quality of tourism: competitiveness, sustainability and seasonality

Table of contents

1. Introduction .................................................................................................................. 1
2. Main features of the survey ...................................................................................... 1
3. Preliminary analysis on the raw monthly data on occupancy ........................................ 2
4. Seasonality .................................................................................................................... 3
5. Analysis of Seasonality ............................................................................................. 4
6. The seasonal adjustment of the time series on occupancy ............................................ 11
7. Final remarks .............................................................................................................. 14
8. References .................................................................................................................. 15
9. Annex I: Time Series of raw data on Occupancy ....................................................... 16
10. Annex II: Elementary component of the time series on nights spent: Charts .............. 19

1. Introduction

The survey on the Occupancy in Collective Accommodation establishments is carried out by Istat in the frame of the Regulation (EU) n. 692/2011 concerning European statistics on tourism. It mainly collects monthly data on occupancy of tourist accommodation establishments in Italy. Although it is not expressly provided for in the Regulation n. 692/2011, the seasonal adjustment of variables on occupancy, traditionally strongly seasonal (at least in such countries as Italy), derives mainly from the need to make short-term analysis, with more sophisticated methods than simple ratios. Moreover it permits to properly investigate the long-term trend of the phenomena.

Before dealing with the seasonal adjustment of variables on occupancy, this paper focuses particular attention on the main component that influences the time series of tourism: the seasonal factor. It also explains the methods used to quantify and identify the seasonal pattern, the evidence that the seasonally adjusted series highlight and the discrepancies between the forecasts made by the chosen models for estimating and the real occupancy values.

2. Main features of the survey

The census survey on the Occupancy in Collective Accommodation establishments (OCA) is carried out by Istat, the Italian NSI, in the frame of the Regulation (EU) n. 692/2011 concerning European statistics on tourism. Anyway, Italy has a long tradition in collecting data on tourists flows in tourist accommodations, as it started this activity since 1957.

OCA mainly collects data on arrivals and nights spent at tourist accommodation establishments in Italy. The accommodation establishments are mainly divided into two main typologies: hotels and similar

1 ISTAT, Short-term Economic Statistics Directorate; cavallo@istat.it – masantor@istat.it
accommodation and other collective accommodation establishments. Hotels and similar accommodation are broken down by category; the other collective accommodation establishments by type: tourist camp-sites, holiday villages, tourist camp-sites and holiday villages-mixed forms, holiday dwellings (rented), farmhouses, youth hostels, holidays homes, mountain refuges, bed and breakfast and other accommodation n.e.c.

Arrivals and nights spent are broken down by category/type of accommodation, by region of residence (for residents) and by country of residence (for non-residents). Data are collected monthly at municipality level. Moreover, OCA provides also information on the average length of stay and, for hotels and similar establishments, on the occupancy rates of bed-places and of bedrooms.

The process of data collection works as follows: daily data on tourists flows are provided by the accommodation establishments (using means as electronic forms, output deriving from the software used during their activity, according to the Istat standards). Then they are summed up for each municipality on monthly basis by the local tourism bodies and sent (mainly by the Regional offices) monthly to Istat by web (uploading the files directly on the protected Istat website for data capturing).

The survey procedure, definitions, classifications and rules are defined and sent by Istat annually to the Regional offices.

OCA is the main source for evaluating tourism flows thanks to the high detail and the availability of long time series. It allows users to break down the data by several variables (NUTS, month, type of accommodation, region of residence/country of residence), providing highly disaggregated data for the analysis.

First results (arrivals, nights spent, at national level, by residents-non-residents, by type of accommodation) are disseminated monthly, after about three months, through the data warehouse of the statistics produced by Istat “I.Stat” (http://dati.istat.it) and the Eurostat database; final data (much more detailed) are disseminated yearly through many ways (data warehouse I.Stat, specific publications such as ASI yearbook and NOI ITALIA, Eurostat database, etc.).

3. Preliminary analysis on the raw monthly data on occupancy

Graphic methods are a powerful tool to analyze time series, since they allow simple comparisons. These ones can give a first general idea of the main features of the series, in terms of periodicity, trend-cycle, seasonality and peak values.

The analysis of the raw monthly data relates to the following 18 time series:

1. Total arrivals in “Total accommodation establishments”;
2. Total nights spent in “Total accommodation establishments”;
3. Arrivals of residents in “Hotels and similar accommodation”;
4. Arrivals of non-residents in “Hotels and similar accommodation”;
5. Arrivals of residents in “Other accommodation establishments”;
6. Arrivals of non-residents in “Other accommodation establishments”;
7. Arrivals of residents in “Total accommodation establishments”;
8. Arrivals of non-residents in “Total accommodation establishments”;
9. Total arrivals in “Hotels and similar accommodation”;
10. Total arrivals in “Other accommodation establishments”;
11. Nights spent by residents in “Hotels and similar accommodation”;
12. Nights spent by non-residents in “Hotels and similar accommodation”;
13. Nights spent by residents in “Other accommodation establishments”;
14. Nights spent by non-residents in “Other accommodation establishments”;

2 The rapid key indicators of nights spent by residents and non-residents at tourist accommodation establishments, as listed in Section 2 of Annex I of the EU Regulation n. 692/2011, are sent to Eurostat within 8 weeks after the end of the reference period and disseminated through the Eurostat database.
15. Nights spent by residents in “Total accommodation establishments”;
16. Nights spent by non-residents in “Total accommodation establishments”;
17. Total nights spent in “Hotels and similar accommodation”;
18. Total nights spent in “Other accommodation establishments”.

The reference time taken into account is the period from 2000 to 2012.

The seasonal pattern of the time series is evident at first sight (Annex I). Time series of total arrivals and total nights spent (1 and 2 of the previous list) have a clear cyclical pattern and the values are concentrated in specific periods of the year, mainly in summer time. Moreover, the general trend of arrivals is more increasing than nights spent.

About arrivals, both, residents and, especially, non-residents, show a growing trend with different peaks (7, 8).

The presence in “Other collective accommodation establishments” of the “open air” establishments causes the concentration of arrivals in summer time, more relevant than “Hotels” (9, 10).

Regarding arrivals in “Hotels”, non-residents highlight a wider peak period (4) than residents, that, on the contrary, are concentrated in August (3). No relevant differences between residents and non-residents for arrivals in “Other collective accommodation establishments”, except other secondary peak months for the resident component (5, 6).

Nights spent by residents are more regular than non-residents (15, 16). Considering the type of accommodation, “Others” show a higher seasonal swing than “Hotels” (17, 18).

As well as for arrivals, nights spent in “Hotels” highlight a different seasonal pattern between the two components (11, 12).

Nights spent by residents are not so different for the two types of accommodation (11, 13), this is not the case for the non-resident component (12, 14).

4. Seasonality

Seasonality is largely considered as a physiological characteristic of tourism that involves the concentration of tourist flows in relatively short periods of the year in a specific destination.

Among many factors affecting tourism flows, seasonality is the most important and, for sure, one of the most visible characteristics.

Although there isn’t a generally accepted definition of seasonality in tourism, there have been numerous attempts to define the seasonal concept in the tourism industry. All the definitions are similar and can be summarized as movements in a time series during a particular time of the year that recur similarly each year.

It has been generally accepted that seasonality in tourism has two basic elements, one that may be called “natural” and the other one “institutional”, involving both the origin and the destination areas (BarOn, 1975).

The “natural” element relates to regular temporal variations in natural phenomena, particularly those ones associated with climate and the seasons of the year, and refers to weather, temperature, hours of sunlight, levels of rainfall and snowfall; these natural phenomena make the tourist areas different for seasonal potential and resources and thus they are perceived to have particular seasonal qualities. Although natural factors can make a destination unattractive to particular markets, they are not the sole reason for

3 “Open air” (tourist camp-sites and holiday villages) account in 2012 for 51.6% of the total nights spent in other collective accommodation establishments.
4 For instance, Butler (2001) described seasonality in tourism “as a temporal imbalance in the phenomenon of tourism, which may be expressed in terms of dimensions [...] as numbers of visitors”. Wall and Mathieson (2006) argued that most destination areas “experience an annual cycle of activity with a peak season and an off-season which are separated by two shoulder seasons”. Cooper, Wanhill, Fletcher, Gilbert and Fyall (2008) stated that seasonality is “the temporal fluctuations of tourism on a daily, weekly, monthly or annual basis".
variations in tourism flows. As Koenig-Lewis and Bischoff (2005) wrote, all these natural factors “are predictable as they are relatively stable in a particular destination and recur with only small changes”.

About “institutional” seasonality is primarily linked to the following three factors, mainly related to the consumers of tourism activity: holidays (school/university/industry and public/religious) and availability of leisure time; travel habits and motivations (which are affected by changing tastes, social pressure and fashion); hosting and timing of events.

Finally, it is to point out that some factors are stable over long periods (e.g. the timing of Christmas), some change at discrete intervals (e.g. vacations), some vary continuously but predictably (e.g. the timing of Easter), whilst others are unpredictable (e.g. the weather).

In order to understand the tourism seasonality, it is not enough to use only one measure: the decision about which measure has to be used depends on the research and the degree of detail required. It is necessary to quantify the phenomenon but also to explore the pattern, the amplitude of the seasonal swing and the persistence of its variations.

The first step is to analyze the monthly distribution of the phenomena, detect the peak and the bottom periods and their ratios, use appropriate graphical representations and calculate “seasonal indexes” like the seasonal range\(^5\), the seasonality ratio\(^6\), the coefficient of seasonal variation\(^7\) and the Gini coefficient\(^8\).

### 5. Analysis of Seasonality

Monthly arrivals and nights spent show a significant seasonal bias during the reference period. Arrivals are a bit more evenly spread over the year than nights spent, mainly due to the concentration of longer stays in July and August\(^9\). Both variables have their peak in August (Fig. 1).

**Figure 1. Distribution of arrivals and nights spent per month (% - average 2000-2012)**

---

\(^5\) The *seasonal range* (SR) is the difference between the highest and the lowest value of monthly indices (Koenig-Lewis & Bischoff, 2005), where \(\text{index} = 100\) is the average number of nights spent per month.

\(^6\) The *seasonality ratio* (R) is a relation between the highest number of monthly nights spent and the average number of nights spent per month.

\(^7\) The *coefficient of seasonal variation* (s) is the standard deviation of the seasonal indices, where \(\text{index} = 100\) is the average number of nights spent per month.

\(^8\) The *Gini coefficient* (G) is a statistical measure of inequality and is derived from the Lorenz curve. The Lorenz curve is a graphical representation of inequality and the Gini coefficient is a measure of that.
Arrivals in the peak month are 3.04 times higher than in the lowest month (January), while nights spent in the peak month are 6.74 times higher than in the lowest month (November)\textsuperscript{10}.

The following analysis is focused mainly on the monthly data on nights spent, variable that better quantify the volume of tourism from the supply side.

The tourist accommodation sector experiences strong seasonal fluctuations. During the reference period, the two months of July and August account for nearly the 40% of all nights spent in tourist accommodation in a year, with the III Quarter that represents more than half (51%) of all nights spent (Fig. 2).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Distribution of nights spent – two-month periods and quarters (% - average 2000-2012)}
\end{figure}

An indicator of the amplitude of the seasonal swing is the ratio between nights spent in the peak month and nights spent in the bottom month, periods of maximum divergence. This ratio is in average 6.73 during the period, with the lowest value (6.20) in 2007 and the highest in 2003 (7.21) (Fig. 3).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Ratio between nights spent in peak month / bottom month (Years 2000-2012)}
\end{figure}

The Gini coefficient ($G$), used as a synthetic indicator of seasonality, amounts to nearly 0.38, not so variable over the years (Tab. 1). Also the other seasonal indexes provide the same indication.

\textsuperscript{9}To synthesize the phenomenon, the average values over the period of each month are used.

\textsuperscript{10}Arrivals/Nights Spent in the peak month (average 2000-2012) divided by Arrivals/Nights Spent in the bottom month (average 2000-2012).
Table 1. Seasonal indexes (Years 2000-2012)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Seasonal Range</td>
<td>224.15</td>
<td>224.13</td>
<td>222.80</td>
<td>225.04</td>
<td>217.54</td>
<td>213.18</td>
<td>207.90</td>
<td>208.52</td>
<td>211.84</td>
<td>220.52</td>
<td>217.85</td>
<td>216.33</td>
<td>213.21</td>
</tr>
<tr>
<td>Seasonality Ratio</td>
<td>2.62</td>
<td>2.62</td>
<td>2.61</td>
<td>2.61</td>
<td>2.55</td>
<td>2.52</td>
<td>2.47</td>
<td>2.49</td>
<td>2.49</td>
<td>2.57</td>
<td>2.56</td>
<td>2.53</td>
<td>2.52</td>
</tr>
<tr>
<td>Coefficient of Seasonal Variation</td>
<td>73.01</td>
<td>73.09</td>
<td>72.57</td>
<td>72.76</td>
<td>70.13</td>
<td>69.34</td>
<td>68.78</td>
<td>69.77</td>
<td>69.69</td>
<td>71.72</td>
<td>71.87</td>
<td>72.04</td>
<td>70.94</td>
</tr>
<tr>
<td>The Gini coefficient</td>
<td>0.3897</td>
<td>0.3920</td>
<td>0.3905</td>
<td>0.3921</td>
<td>0.3797</td>
<td>0.3736</td>
<td>0.3729</td>
<td>0.3774</td>
<td>0.3750</td>
<td>0.3847</td>
<td>0.3845</td>
<td>0.3887</td>
<td>0.3820</td>
</tr>
</tbody>
</table>

The value of $G$ is a sign of a fairly pronounced seasonality, especially if compared to the main EU competitors. In fact, for example, in 2012 Italy has a $G$ value higher than France and Spain and lower only than Greece, a summer destination with a pronounced seasonality ($G = 0.5442$) (Fig. 4).

Figure 4. The Gini coefficient – Italy vs. main EU competitors (Year 2012)

It’s important to analyze not only the amplitude, but also the persistency and the variability of the seasonality; seasonality indexes are very suitable for the purpose.

As seen before, all the seasonality indexes show the same pattern in the period 2000-2012. Generally, during the reference period, the seasonality is in slight reduction (Tab. 1 and Fig. 5).
The seasonality factor influences more the time series of residents than of non-residents. In fact, the $SR$ values are always higher for residents, although they decrease a little over the reference period (Fig. 6).

Anyway, residents and non-residents have a different seasonal pattern. In fact, nights spent by residents are more concentrated in summer months than non-residents, which are flatter over the year. The domestic component is the main contributor to the peak of Total nights spent in tourist accommodation in
July and August. The peak month is August for residents and July for non-residents; the bottom month is, respectively, November and December.

Residents have a second peak month, December, determined by the Christmas holidays. Non-residents show higher values in the shoulder periods (Spring and Autumn) (Fig. 7).

Figure 7. Distribution of nights spent per month by residence (% - average 2000-2012)

To analyze the general pattern of the series, a useful graphical representation is the cycle plot, used to highlight the changes occurring in the same month over the years.

The Figure 8 show, in addition to the different monthly levels for the two components, the substantial stability over the time of the monthly distribution of nights spent for residents and a fluctuating pattern for non-residents. The main differences are in the period March-June: more variable for non-resident, less for residents. This is a sign of a seasonality factor in these months, more prominent for the non-resident component, and determined by the influence on the series of the moving holidays (e.g. Easter, Whitsuntide) (Fig. 8).

This effect is even more evident in the Figures 9 and 10. In these graphs are showed the deviations from the average\textsuperscript{11} of the monthly distributions of nights spent in March and in April (months in which Easter falls) and, only for non-residents, in May and in June (months in which Whitsuntide falls).

There is an “Easter effect” in the series, wider and persistent for the non-resident component for which, in addition, it has to be noted also the effect of Whitsuntide, celebrated in many EU countries (Fig. 9 and 10).

\textsuperscript{11} It means that, for each year \(y\), the deviations for the month \(m\), is: \(NS_{my} / NS_{y} = \sum_{y=1}^{n} NS_{my} / \sum_{y=1}^{n} NS_{y}\) where \(NS\) are nights spent and \(1\) to \(n\) the length of the time series.

\textsuperscript{12} In time series decomposition, trend and cycle are often associated.
Figure 8. Cycle plot - Distribution of nights spent per month by residence (Year 2000-2012)

Figure 9. Distribution of nights spent in March and in April by residents (Difference from average 2000-2012)

Figure 10. Distribution of nights spent in March, in April, in May and in June by non-residents (Difference from average 2000-2012)
Seasonality by type of accommodation is less pronounced for “Hotels” than for “Other accommodation establishments” (Fig. 11), for the presence of “open air” establishments in this latter typology (tourist campsites and holiday villages, often closed during the winter period). The SR values of nights spent in “Hotels” are substantially stationary over the period, while for “Others” they are decreasing.

In the reference periods, the peak for “Hotels” is steadily lower than for “Others” (Fig. 12). Typically, seasonal bias for hotels and similar establishments is smoothened by overnight business travels.

Figure 11. Seasonal Range by type of accommodation (Years 2000-2012)

Figure 12. Distribution of nights spent per month by type of accommodation (Average 2000-2012)

Seasonality is more evident, break down data by type of accommodation and by residence. Monthly distribution of nights spent by residents in “Hotels” show over the period, in addition to August, a second peak month in December and the Easter effect (Fig. 13). Regarding non-residents in the same type of accommodation, the time series is completely different, so as the seasonality pattern. Non-residents have a two peaks seasonality (July and September), nearly on the same level, and a strong moving holidays effect, linked to Easter and Whitsuntide (Fig. 14).

Figure 13. Distribution of nights spent by residents in “Hotels” (Years 2000-2012)
About “Others”, it has to be noted, for the resident component, the second peak month in December, lower than in “Hotels” (Fig 11 and 13 in Annex I).

To have a complete view of the seasonal factor affecting the main time series of nights spent, it is represented in the charts in Annex II.

6. The seasonal adjustment of the time series on occupancy

Seasonal adjustment (SA) is a statistical method for removing the seasonal component of a time series. It allows to decompose a raw series of data in its elementary components in order to analyze and to determine monthly changes and long-term trends, not more influenced by seasonal factors.

A monthly time series is made up of three components:

$$Y_t = T_t + S_t + I_t$$

where $t$ is the month, $T_t$ is the trend-cyclical component$^{12}$, $S_t$ is the seasonal component and $I_t$ is the irregular component.

Trend is a slow variation over several years, generally due to structural reasons. Cycle is a periodic fluctuation, characterized by alternating periods of higher and lower rates of change; in the majority of cases, it is connected to the fluctuations of the overall economic activity.

Seasonal variation, as previous seen, represents the effect of natural factors, climatic and institutional events, which recur more or less regularly each year.

The irregular component represents unforeseeable movements linked to any type of events. In general, they are unpredictable and of stable nature.

The decomposition of a time series makes it possible to identify the trend-cycle, the seasonal variation and the irregular fluctuations. By removing the seasonal component, it is easier to focus on the other components.

There are several methods for seasonal adjustment$^{13}$. Istat produces and disseminates seasonally adjusted data on occupancy from the end of the 90s using “TRAMO-SEATS” programs, where TRAMO stands for “Time series Regression with ARIMA”$^{14}$ noise, Missing values and Outliers” and SEATS for “Signal Extraction in ARIMA Time Series”$^{15}$.

---

$^{12}$ In time series decomposition, trend and cycle are often associated.

$^{13}$ For example: X-12-ARIMA, TRAMO/SEATS or STAMP.

$^{14}$ The acronym ARIMA stands for “Auto-Regressive Integrated Moving Average”. ARIMA models are the widely-used approach to time series modelling, first popularized by Box and Jenkins. The general notation for the order of a seasonal ARIMA model with both seasonal and non-seasonal factors is ARIMA$(p, d, q)\times(P, D, Q)s$. The term $(p, d, q)$ refers to the order of the autoregressive $(p)$, integrated $(d)$ and moving average $(q)$ of the non-seasonal part of the ARIMA model; the
Modelling is made through Demetra+\(^{16}\), an application that provides user-friendly tools for checking the quality of the SA results.

The procedure of SA concerns the following 8 elementary series:

1. Arrivals of residents in “Hotels and similar establishments”.
2. Nights spent by residents in “Hotels and similar establishments”.
3. Arrivals of non-residents in “Hotels and similar establishments”.
4. Nights spent by non-residents in “Hotels and similar establishments”.
5. Arrivals of residents in “Other accommodation establishments”.
6. Nights spent by residents in “Other accommodation establishments”.
7. Arrivals of non-residents in “Other accommodation establishments”.
8. Nights spent by non-residents in “Other accommodation establishments”.

From the 8 elementary adjusted series, 10 aggregated series are derived, using the indirect approach\(^{17}\):

- 4 series of the totals by “residence of the guests” (2 for arrivals and 2 for nights spent),
- 4 of the totals by “type of accommodation” (2 for arrivals and 2 for nights spent) and
- 2 of the general totals (Total Arrivals and Total Nights spent):

9. Arrivals of residents (1 + 5).
10. Nights spent by residents (2 + 6).
11. Arrivals of non-residents (3 + 7).
12. Nights spent by non-residents (4 + 8).
13. Arrivals in “Hotels and similar establishments” (1 + 3).
14. Nights spent in “Hotels and similar establishments” (2 + 4).
15. Arrivals in “Other accommodation establishments” (5 + 7).
16. Nights spent in “Other accommodation establishments” (6 + 8).
17. Total Arrivals (1 + 3 + 5 + 7).
18. Total Nights spent (2 + 4 + 6 + 8).

The modelling approach consists in two steps: a pre-adjustment step, which removes the “deterministic” component of the series and the decomposition part itself. It’s used an automated model selection, after the analysis of the principal diagnostics and appropriate adjustments.

Since, as seen before, there is a great difference in time series on occupancy by residence, different calendars and regressors for moving holidays are used for modeling. In particular:

- different regression variables for “trading days”;
- different duration of the Easter regression effect;
- a specific regressor for moving holidays (Whitsuntide) for non-residents.

For all the analyzed series the models are “airline”\(^{18}\), except for the series of arrivals of residents (1 and 5)\(^{19}\).

The trend-cyclical component plus the irregular component compose the final seasonally adjusted series.

---

\(^{15}\) Gomez and Maravall, 1996.

\(^{16}\) Demetra+ covers the recommendations of ESS guidelines in area of seasonal adjustment. The software has been developed by the National Bank of Belgium in cooperation with Eurostat.

\(^{17}\) SA of aggregated time series can be performed in two ways, through: by direct or by indirect SA. Direct approach is performed when all the time series, including aggregates, are seasonally adjusted on a single basis. Indirect approach involves seasonally adjusting subcomponents and combining them to obtain seasonally adjusted aggregate.

\(^{18}\) The “airline model” is an ARIMA(0,1,1)(0,1,1). It was proposed by Box and Jenkins (1976) for modelling the well-known monthly series of airline passengers.

\(^{19}\) Respectively ARIMA(2,1,0)(0,1,1) and ARIMA(0,1,1)(0,1,0).
The graphical analysis of the seasonally adjusted series shows a general growing trend both for Total arrivals and Total nights spent over the reference period (Annex I, 1 and 2). Nevertheless, it’s clear the effect, more evident for nights spent, of the economic crisis: this is characterized by two main phases, one from the second semester of 2008 up to the first semester of 2009 and the other starting from the end of 2011 up to nowadays. The growing trend of Total arrivals is mainly influenced by “Others” than “Hotels”.

About arrivals, residents (7) grow steadily over the period up to the IV quarter of 2011, when they start to decrease mainly due to the economic crisis. The same trend is also evident for nights spent (15).

On other hand, non-residents have a more fluctuating trend: it’s evident the first phase of the crisis and the significant growth, starting from the second part of 2009 (8, 16). It has to be noted that the effect of the economic crisis seems to affect in its first phase only the non-resident component, while in its second phase only residents.

Arrivals in “Others” increase continuously over the period, being less influenced by the first phase of the crisis than “Hotels” (9, 10). Nights spent by type of accommodation have the same pattern as arrivals, but with a flatter trend (17, 18).

The trend of residents, growing in the reference period with a decrease starting from the end of 2011, is the same for the two type of accommodation both for arrivals and nights spent. Arrivals grow more than nights spent, especially in “Others” (3, 5, 11, 13).

The trend of Total arrivals and Total nights spent by non-resident is influenced mainly by “Hotels” (4, 12). In “Others” there is no evidence of the economic crisis, with values growing steadily over the period (6, 14).

Decomposition is primarily useful for studying time series data and exploring their changes over the time, but, in addition, it can be used successfully also for forecasting.

In fact, the modelling at the base of SA allow to estimate the next future values of the series, by composing the expected trend component and the seasonal component. It is usually assumed that the seasonal component is unchanging, or changing extremely slowly, and so it is forecast by simply taking the last year of the estimated component.

According to chosen models, the results of the forecasting fit quite well with the provisional real monthly values. Focusing only on nights spent in Total and by residence, the forecasts present higher accuracy for residents (+1.9% is the difference of forecast vs. provisional real value for Total nights spent in the period Jan-Nov 2013) than for non-residents (-3.6%). For the total nights spent the accuracy is very good (+0.1%). More relevant differences arise from the analysis of the results per month, with a slight decrease in accuracy mainly in the shoulder months (months with the lowest values of the phenomenon during the year). Although the model doesn’t catch exactly the effect of the crisis for the resident component in the first part of the year, being the forecasts higher than the real values, the opposite occurs in the last months, probably a sign of a tiny recovery. For non-residents a residual seasonal component, linked mainly to the moving holidays, seems not yet perfectly detected by the model (Fig. 15 e 16).
To have a complete picture of the main decomposed time series in their elementary components (Trend, Seasonal and Irregular), in Annex II there are the related charts with also graphical representations of the forecasts.

7. Final remarks

The importance of SA, especially on occupancy variables, strongly affected by seasonality, is first of all to supply a powerful tool to analyse short-term trend of the phenomena. The decomposition of a time series in its elementary components allows to better understand the data, their changes over the time, and, moreover, to forecast.

In fact the simple ratios do not allow deep studies on seasonality influenced time series and forecasting is very useful in particular for economic planning matters.
For these reasons, the next step will be to seasonally adjust the occupancy variables of the “Open air” component separately from the other accommodation establishments, because of the very strong seasonality of the first ones respect to the others. Moreover, this break down will allow a full correspondence to the accommodation breakdown as requested by the Regulation (EU) n. 692/2011 concerning European statistics on tourism.

8. References


1. Arrivals - Total

2. Nights spent - Total

3. Arrivals of residents in “Hotels”

4. Arrivals of non-residents in “Hotels”

5. Arrivals of residents in “Others”

6. Arrivals of non-residents in “Others”
7 Arrivals of residents – Total

8 Arrivals of non-residents - Total

9 Arrivals in “Hotels” – Total

10 Arrivals in “Others” - Total

11 Nights spent by residents in “Hotels”

12 Nights spent by non-residents in “Hotels”
10. Annex II: Elementary component of the time series of nights spent: Charts

1. Total nights spent

2. Nights spent by residents
3. Nights spent by non-residents

4. Nights spent in “Hotels”
5. Nights spent in “Others”