A complex network analysis of inbound tourism in Sicily

Davide Provenzano\textsuperscript{a}, Rodolfo Baggio\textsuperscript{b}

\textsuperscript{a} Department of Economics, Statistics and Business, University of Palermo
E-mail: davide.provenzano@unipa.it

\textsuperscript{b} Master in Economics and Tourism, Bocconi University and National Research Tomsk Polytechnic University
E-mail: rodolfo.baggio@unibocconi.it

Abstract

In this paper the complex dynamics of inbound tourism in Sicily are analyzed for the period 1998-2017. The Horizontal Visibility Graph Algorithm is used to transform the overnight stays time series into a network whose topology is investigated by standard network analysis. Discontinuities in the domestic and international tourism demand were identified in order to detect signals of change and the timing of the directional change in tourism growth. The network degree distribution confirms the complex structure of the destination and reveals the random, and thus more unpredictable nature of the international tourism demand in Sicily, compared to a more stable domestic segment. Some policy implications are drawn.

Keywords: Tourism; network analysis; horizontal visibility graph algorithm; time series; island; Sicily

1 Introduction

Sicily is an autonomous region located in the south of the Italian Peninsula and the largest island in the Mediterranean Sea. Because of the past dominations by the Phoenicians, the Greeks, the Romans, the Vandals, the Ostrogoths, the Byzantine Empire, the Emirate of Sicily, the Normans and the Kingdom of Aragon, Sicily has a rich and unique art, music, literature, cuisine, and architecture. The island is also known for its stunning golden beaches, quaint fishing villages, and, of course, for the capital, Palermo, famous for its street food and its extraordinary architecture. Sicily welcomes more than 4.4 million of tourists per year (ISTAT, 2017), both international (45%) and domestic (55%) with a positive trend reported since 2010 due to the crises in North Africa and the Middle East, lately characterized by political turmoil and terrorism. In 2017, one million more tourism overnight stays were recorded with 10 to 20 million hotel rooms booked by both residents and visitors (EUROSTAT, 2018).

Yet, such unexpected upturn has showed a lack of policy formulation and agenda setting in local tourism. The shortage of adequate transportations and tourist facilities, and a wrong distribution of public support have been often pointed out as the main bounds to the full exploitation of the enormous potentialities tourism offers for increasing the employment rates and the development of the local economy. Besides a process of seasonal adjustment of
tourism flows in Sicily, it is strongly felt the need of a regional planning based on the structural and dynamic characteristics of the tourism demand in Sicily. In this study we tested the hypothesis that the structure of tourism in Sicily has complex traits that make the recourse to linear models in demand analysis flawed. As a matter of fact, linear models can be usefully applied for explaining present configurations or predict future ones if the system remains stable over time. On the contrary, when the system shows dynamics attributable to a complex structure, linear models experience serious issues. Once revealed the complex structure of tourism demand in Sicily, our second concern was for the chaotic, deterministic or stochastic dynamics showed by the system under study, whose investigation is a fundamental prerequisite for any governance action. We also looked for discontinuities in the evolution over time of the destination, called turning points (Coshall, 2000), which could help detect the symptoms of a directional change in the tourism growth. In fact, the research of turning points has a high practical value because “tourism-related firms are keen to know not only the overall trends of tourism demand, but also the timing of the directional change in tourism growth” (Song & Li, 2008, p. 214). Turning points also “helps both territories and businesses to increase the consciousness of their strategies and permits to align managerial levers to changes” (Baggio & Sainaghi, 2016, p. 24).

Hence, the following research questions were central to our study:

- Q1: Does tourism demand in Sicily show traits and dynamics typical of complex systems?
- Q2: Provided that the complex nature of the system under study is demonstrated, is its behavior characterized by chaotic, deterministic or stochastic dynamics over time?
- Q3: Does the time series of overnight stays in Sicily exhibit any symptom of a directional change in the tourism trend?
- Q4: Which interventions should be implemented by policy-makers, destination marketing organizations (DMOs), and, in general, practitioners to govern the tourism system in Sicily coherently with the results of the analyses conducted?

Following such questions, our research framework was developed through four steps. After a similarity analysis, the time series of overnight stays in Sicily were mapped into a graph and then analyzed by using standard network techniques (Baggio et al., 2010b; da Fontoura Costa et al. 2007; Newman, 2010; Scott et al., 2011). In particular, the predictability of the future behavior of tourism demand in Sicily was inspected by calculating the shape of the nodal degree distribution. The third act was a modularity analysis to uncover sets of nodes more densely connected with each other than with nodes outside the group. These communities are a common feature of many real systems and are important to detect discontinuities in the time series. Lastly, some policy implications were drawn.

Our analysis is strongly rooted in network science for the benefits proven by the use of the network concept in tourism research (Scott et al., 2007; Scott et al., 2008; Baggio et al., 2010a; Baggio et al. 2010b; Benckendorff & Zehrer, 2013). The combination of network analysis and the horizontal visibility graph algorithm (HVG) used to map time series into networks is able to provide a number of interesting insights in spite of its simplicity and straightforward implementation.

Data for the period 1998-2017, regarding tourism overnight stays in Sicily at regional level (Fig. 1, a and b), were provided by the “Assessorato Regionale del Turismo dello Sport e
Acturally, overnight stays play an important role in the definition of tourism demand as such quantity is associated to tourist spending and depends on the perceived characteristics of the destination (Sainaghi, 2012).

Results revealed that tourism demand in Sicily shows dynamics partly attributable to a random series so as to make any policy intervention based on past experience potentially inappropriate. We found that in the last 20 years the historical evolution of tourism in Sicily has been characterized by two turning points in late 2000 and early 2011 for the domestic and international tourism flows, respectively.

The proposed study contributes to the existing literature in three ways. First of all, the case of Sicily adds to the small number of investigations successfully applying HVG to tourism demand analysis and it could be used as a reference point for evaluating and discussing other studies conducted in a comparable framework. Secondly, our analysis confirms the complex structure of tourism systems and destinations against the prevalent recourse of academics and practitioners to linear models in demand analysis (Baggio & Sainaghi, 2016; Sainaghi & Baggio, 2017). Finally, despite the huge flow of tourists hosted every year (Sicily is the ninth region in Italy for tourism), to the best of our knowledge, the structural and dynamic characteristics of tourism in the island have never been studied before.
The rest of the paper is organized as follows. The next section provides a brief overview of the literature dealing with tourism in Sicily and complex network analysis. Section 3 describes in detail the methodology used in the study, its peculiarities and advantages compared to other methods. The forth section presents an application of the Horizontal Visibility Graph algorithm to the monthly series of stays in the Sicilian accommodations and debates the outcomes of the analysis. Section 5 discusses policy implications and concludes. Finally, limitations of the proposed investigation and possible future developments are presented.

2. Literature review

Literature dealing with tourism in Sicily is limited to a few studies investigating the causes and impact of seasonal patterns in tourism demand, and the mobility of tourists in the island. In particular, tourism demand is analyzed in De Cantis et al. (2011) for the effects of seasonal variations on the bed occupancy rate of Sicilian accommodation establishments. The seasonality of tourist presence is also investigated in Cuccia and Rizzo (2011) for several destinations in Sicily, selected according to their different degree of cultural attractiveness. The authors investigate whether the cultural attractiveness of tourism destinations is able to mitigate the seasonal patterns of tourism demand. Finally, the relationship between seasonality patterns and the development of the tourism sector is addressed in Volo (2007, 2010), for a better understanding of the evolution of international and domestic tourism in Sicily.

The spatial distribution of tourists represents the second main strand of research for tourism in Sicily. In particular, mobility of incoming tourism has been investigated to analyze the competitiveness of Sicilian tourism destinations (Oliveri et al., 2012), to quantify the main implications of multi-destination trips in Sicily both on tourism statistics and on destination management (Parroco et al., 2012a, 2012b), to define a territorial network of tourism demand in Sicily (D’Agata et al., 2013), and to explore the issue of unobserved tourism and estimate the actual magnitude of tourism flows in the island (De Cantis et al., 2015).

Two isolated case studies have identified and examined which factors have helped rural communities to successfully develop agritourism (Privitera, 2010), and the main determinants of tourism in Sicily (Provenzano, 2010).

Seasonal indices and measures (Kuznets, 1933; Lim & McAleer, 2001; Lo Magno et al., 2017) are the preferred statistical methods for measuring seasonality of inbound tourism in the cited manuscripts. In Oliveri et al. (2012), and D’Agata et al. (2013) network analysis (Beaumont & Dredge, 2010; Timur & Getz, 2008) is applied to demonstrate that Sicilian tourism develops around few attractive destinations. Finally, System Dynamics (Forrester, 1961; Sterman, 2000) completes the list of the methodological frameworks used to analyze tourism flows in Sicily.

In the field of tourism demand analysis, the most used techniques are linear or linearized models with a clear prevalence of regression methods. (Frechtling, 2001). Yet, linear (or linearized) models are not able to describe and study the nature of complex systems. Thus, new approaches and paradigms to deal with complexity are not only welcome but needed.
In the last years, the growing awareness of the complex nature of tourism systems and destinations made of non-linear relationships, self-organization behaviors, emergence of modular and hierarchical structures, and robustness or fragility regarding some events has favored the use of complexity science methods. On the heels of the applications in economics, mathematics, physics, and sociology (Ellis, 2005; Holling, 2001; Mainzer, 2005; Perona, 2007; Prigogine & Hiebert, 2008, and references therein) the use of complexity science has already provided an effective contribution in the study of destination development (Cole, 2009; Faulkner, 2002; Warnken, et al., 2003; Zahra & Ryan, 2007), the management and the effects of crises and disasters (Crandall et al., 2010; Faulkner & Vikulov, 2001; Laws & Prideaux, 2005; Prideaux et al., 2003; Ritchie, 2004; Scott & Laws, 2005), the forecast of future demand (Faulkner & Russell, 2001; Faulkner & Valerio, 1995), the development of entrepreneurship (Russell & Faulkner, 1999, 2004), the structure of the networks between tourism companies (Tinsley & Lynch, 2001), the management of hospitality businesses (Edgar & Nisbet, 1996), or in maritime transportation systems and aviation networks (Fremont, 2007; Hu & Zhu, 2009; Ducruet et al., 2010; Tsiotas & Polyzos 2014).

In the domain of complexity science, a recent methodological proposal, named Horizontal Visibility Graph (HVG), has shown remarkable characteristics of clarity and computational simplicity in examining the complex dynamics of stock market indices, exchange rates, macroeconomic indices, human behaviors, occurrence of hurricanes, and dissipation rates in turbulent systems (Chao & Jin-Li, 2012; Elsner et al., 2009; Tang et al. 2013; Wang & Wang, 2012). In spite of such suitability for the analysis of complex structures, its use in the context of tourism destinations is still limited to a few studies (Baggio, 2014; Baggio & Saineghi, 2016; Baggio & Saineghi, 2017). The Horizontal Visibility Graph (HVG) technique is the approach used in this research for the empirical analysis of tourism demand in Sicily.

3. Materials and methods

The HVG algorithm is able to provide interesting insights into the complex characteristics of the system under study, allowing for a quantitative analysis of its structural and dynamic features. It shows results very similar to those generated by most popular methods as the Lyapunov exponents, the Hurst exponent, fractal dimensions, and symbolic discretization avoiding their sophisticated procedures that are not always completely and rigorously defined. Moreover, the practice of these methods and the interpretation of their results are tasks which can be problematic for many, especially practitioners (Baggio, 2008; Baggio & Saineghi, 2011). Conversely, the HVG approach keeps its remarkable characteristics of clarity, coupled with computational simplicity, and suitability in the context of tourism. Basically, the idea is to transform a time series of \( N \) observations, \((x_1, x_2, \ldots, x_n)\), into a network of \( N \) nodes (Nuñez et al., 2012), which inherits the periodic, random, or fractal nature of the original set of observations, so that standard network analysis (Baggio et al., 2010b; da Fontoura Costa et al., 2007; Newman, 2010; Scott et al., 2011) can be used to understand and evaluate the complexity features of the system under study (Campanharo et al., 2011).

Following the HVG algorithm, numeric values in a time series (Fig. 2a) are first represented as a set of vertical bars (Fig. 2b). Any data point in the series is a node of the network under construction. Then, links are drawn among bars ‘in view’, meaning among bars that can be
connected by a horizontal segment without intersecting any other intermediate bar (Fig. 2c). Formally, for any two data points in the time series, \( i \) and \( j \), horizontal visibility exists if (and only if) every bar \( k \) between \( i \) and \( j \) has a value \( x_k \) lower than both \( x_i \) and \( x_j \) \((x_k < \inf [x_i; x_j])\). Once the nodes and the links identified as described have univocally determined the undirected network graph (Fig. 2d), some practical information can be derived for the time series representing the dynamic system under study. In particular, as analytically discussed in Lacasa & Toral (2010) and Luque et al. (2009), when in the network graph obtained by the HVG algorithm the number of nodes with degree \( h \), \( N(h) \), follows the exponential relationship, \( N(h) \sim e^{-\lambda h} \), the parameter \( \lambda_c = \ln(3/2) = 0.404 \) (Luque et al., 2009) is assumed as the threshold value that separates stable or complex behaviors from a chaotic state. More precisely, the system under study is characterized by chaotic dynamics for \( \lambda \) less than \( \lambda_c \) (\( \lambda < \lambda_c \)), shows an edge of chaos behavior in the region around \( \lambda_c \), is stable and predictable for \( \lambda \) greater than \( \lambda_c \) (\( \lambda > \lambda_c \)).

Networks also show a structure often organized in groups, called communities or modules, where nodes belonging to a group display denser connections between them than with nodes outside the group (Newman & Girman, 2004). The extent to which a network can be divided into well recognizable communities is measured by a modularity index \( Q = \sum_i (e_{ii} - a_i)^2 \), where \( e_{ii} \) is the fraction of connections between nodes belonging to the same module \( i \), and \( a_i \) is the fraction of links with at least one end node inside module \( i \). \( Q \) is normalized between 0 and 1, where 0 means absence of modules, and 1 a perfect division into completely separated groups.

Among the many techniques proposed in the literature to identify the different modules and derive the value of \( Q \) (Fortunato, 2010), we used the algorithm of Blondel et al. (2008). Such algorithm uses a resolution parameter to determine the granularity level at which communities are detected. Basically, it is an iterative optimization model that aims to determine the optimal number of partitions that maximize the index \( Q \). Here we set the resolution equal to 1 to get a standard modularity-based community detection.

These analyses were carried out to investigate inbound tourism in Sicily. The nights spent in any accommodation in the island from 1998 to 2017 was the quantity of interest. Actually, the number of nights spent in Sicily was preferred to tourism arrivals as the former is influenced
by many of the system components: the demand side (tourists), the supply side (infrastructure available to visitors), and a number of internal and external economic factors (Ferro et al., 2003). Moreover, overnight stays are strongly related to tourist spending (Sainaghi, 2012). Therefore, measurements of the nights spent by tourists can be taken as a meaningful representation of the system under study. The relevance of overnight stays is also confirmed in many papers dealing with destination management (Beritelli & Laesser, 2014; Gonçalves & Ratsimbanianera, 2012, 2011). Table 1 shows the figures used for the analyses.

### Table 1. Monthly overnight stays in all types of tourism accommodation in Sicily by guest origin, 1998–2017 (Tot=total, Dom=domestic, Int=international).

<table>
<thead>
<tr>
<th>Month</th>
<th>Year</th>
<th>Dom</th>
<th>Int</th>
<th>Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 1998</td>
<td>1998</td>
<td>194,787</td>
<td>58,301</td>
<td>253,088</td>
</tr>
<tr>
<td>Feb 1998</td>
<td>1998</td>
<td>252,151</td>
<td>74,495</td>
<td>326,646</td>
</tr>
<tr>
<td>Mar 1998</td>
<td>1998</td>
<td>320,635</td>
<td>190,187</td>
<td>510,820</td>
</tr>
<tr>
<td>Apr 1998</td>
<td>1998</td>
<td>488,221</td>
<td>462,292</td>
<td>950,513</td>
</tr>
<tr>
<td>May 1998</td>
<td>1998</td>
<td>561,620</td>
<td>538,101</td>
<td>1,099,721</td>
</tr>
<tr>
<td>Jun 1998</td>
<td>1998</td>
<td>574,701</td>
<td>463,501</td>
<td>1,038,202</td>
</tr>
<tr>
<td>Jul 1998</td>
<td>1998</td>
<td>1,069,474</td>
<td>559,879</td>
<td>1,629,349</td>
</tr>
<tr>
<td>Aug 1998</td>
<td>1998</td>
<td>1,715,791</td>
<td>539,009</td>
<td>2,254,800</td>
</tr>
<tr>
<td>Sep 1998</td>
<td>1998</td>
<td>829,124</td>
<td>589,993</td>
<td>1,419,117</td>
</tr>
<tr>
<td>Oct 1998</td>
<td>1998</td>
<td>408,187</td>
<td>409,916</td>
<td>818,103</td>
</tr>
<tr>
<td>Dec 1998</td>
<td>1998</td>
<td>271,067</td>
<td>71,859</td>
<td>342,926</td>
</tr>
<tr>
<td>Jan 1999</td>
<td>1999</td>
<td>238,460</td>
<td>55,783</td>
<td>294,243</td>
</tr>
<tr>
<td>Feb 1999</td>
<td>1999</td>
<td>256,771</td>
<td>82,111</td>
<td>338,882</td>
</tr>
<tr>
<td>Mar 1999</td>
<td>1999</td>
<td>347,204</td>
<td>218,784</td>
<td>565,988</td>
</tr>
<tr>
<td>Apr 1999</td>
<td>1999</td>
<td>483,355</td>
<td>452,425</td>
<td>935,780</td>
</tr>
<tr>
<td>May 1999</td>
<td>1999</td>
<td>574,725</td>
<td>654,237</td>
<td>1,228,962</td>
</tr>
<tr>
<td>Jun 1999</td>
<td>1999</td>
<td>744,060</td>
<td>525,915</td>
<td>1,269,974</td>
</tr>
<tr>
<td>Jul 1999</td>
<td>1999</td>
<td>1,141,452</td>
<td>630,882</td>
<td>1,772,334</td>
</tr>
<tr>
<td>Aug 1999</td>
<td>1999</td>
<td>1,776,986</td>
<td>785,658</td>
<td>2,562,643</td>
</tr>
<tr>
<td>Sep 1999</td>
<td>1999</td>
<td>865,295</td>
<td>648,892</td>
<td>1,514,187</td>
</tr>
<tr>
<td>Oct 1999</td>
<td>1999</td>
<td>463,960</td>
<td>444,977</td>
<td>918,937</td>
</tr>
<tr>
<td>Nov 1999</td>
<td>1999</td>
<td>287,012</td>
<td>130,699</td>
<td>417,881</td>
</tr>
<tr>
<td>Dec 1999</td>
<td>1999</td>
<td>312,011</td>
<td>77,499</td>
<td>389,510</td>
</tr>
<tr>
<td>Jan 2000</td>
<td>2000</td>
<td>275,448</td>
<td>56,711</td>
<td>332,159</td>
</tr>
<tr>
<td>Feb 2000</td>
<td>2000</td>
<td>302,669</td>
<td>113,890</td>
<td>416,559</td>
</tr>
<tr>
<td>Mar 2000</td>
<td>2000</td>
<td>304,735</td>
<td>225,763</td>
<td>530,508</td>
</tr>
<tr>
<td>Apr 2000</td>
<td>2000</td>
<td>613,172</td>
<td>600,151</td>
<td>1,213,323</td>
</tr>
<tr>
<td>May 2000</td>
<td>2000</td>
<td>501,532</td>
<td>656,484</td>
<td>1,184,016</td>
</tr>
<tr>
<td>Jun 2000</td>
<td>2000</td>
<td>825,171</td>
<td>607,598</td>
<td>1,432,769</td>
</tr>
<tr>
<td>Jul 2000</td>
<td>2000</td>
<td>1,171,093</td>
<td>707,069</td>
<td>1,878,162</td>
</tr>
<tr>
<td>Aug 2000</td>
<td>2000</td>
<td>1,959,473</td>
<td>654,659</td>
<td>2,614,132</td>
</tr>
<tr>
<td>Sep 2000</td>
<td>2000</td>
<td>955,760</td>
<td>731,551</td>
<td>1,687,311</td>
</tr>
<tr>
<td>Oct 2000</td>
<td>2000</td>
<td>487,539</td>
<td>562,210</td>
<td>1,050,749</td>
</tr>
<tr>
<td>Nov 2000</td>
<td>2000</td>
<td>307,996</td>
<td>721,962</td>
<td>1,032,958</td>
</tr>
<tr>
<td>Dec 2000</td>
<td>2000</td>
<td>306,316</td>
<td>107,980</td>
<td>414,296</td>
</tr>
<tr>
<td>Jan 2001</td>
<td>2001</td>
<td>267,431</td>
<td>76,196</td>
<td>343,627</td>
</tr>
<tr>
<td>Feb 2001</td>
<td>2001</td>
<td>278,928</td>
<td>129,743</td>
<td>408,671</td>
</tr>
<tr>
<td>Mar 2001</td>
<td>2001</td>
<td>394,187</td>
<td>279,415</td>
<td>673,602</td>
</tr>
<tr>
<td>Apr 2001</td>
<td>2001</td>
<td>592,727</td>
<td>655,754</td>
<td>1,248,481</td>
</tr>
<tr>
<td>May 2001</td>
<td>2001</td>
<td>582,289</td>
<td>746,812</td>
<td>1,329,101</td>
</tr>
<tr>
<td>Jun 2001</td>
<td>2001</td>
<td>647,004</td>
<td>652,733</td>
<td>1,302,737</td>
</tr>
<tr>
<td>Jul 2001</td>
<td>2001</td>
<td>1,018,879</td>
<td>780,586</td>
<td>1,800,465</td>
</tr>
<tr>
<td>Aug 2001</td>
<td>2001</td>
<td>1,358,652</td>
<td>741,645</td>
<td>2,100,295</td>
</tr>
<tr>
<td>Sep 2001</td>
<td>2001</td>
<td>1,912,012</td>
<td>661,097</td>
<td>2,573,109</td>
</tr>
<tr>
<td>Oct 2001</td>
<td>2001</td>
<td>914,259</td>
<td>724,346</td>
<td>1,638,605</td>
</tr>
<tr>
<td>Dec 2001</td>
<td>2001</td>
<td>290,256</td>
<td>48,490</td>
<td>338,746</td>
</tr>
<tr>
<td>Jan 2002</td>
<td>2002</td>
<td>254,573</td>
<td>61,775</td>
<td>316,348</td>
</tr>
<tr>
<td>Feb 2002</td>
<td>2002</td>
<td>287,676</td>
<td>412,378</td>
<td>699,054</td>
</tr>
<tr>
<td>Mar 2002</td>
<td>2002</td>
<td>403,336</td>
<td>288,981</td>
<td>692,317</td>
</tr>
<tr>
<td>Apr 2002</td>
<td>2002</td>
<td>570,540</td>
<td>557,290</td>
<td>1,137,830</td>
</tr>
<tr>
<td>May 2002</td>
<td>2002</td>
<td>665,630</td>
<td>737,139</td>
<td>1,302,769</td>
</tr>
<tr>
<td>Jun 2002</td>
<td>2002</td>
<td>862,381</td>
<td>615,124</td>
<td>1,477,505</td>
</tr>
<tr>
<td>Jul 2002</td>
<td>2002</td>
<td>1,138,688</td>
<td>728,499</td>
<td>1,867,187</td>
</tr>
</tbody>
</table>

Data from June 1999 to June 2001 and June 2009 to June 2017 (Tot=total, Dom=domestic, Int=international).
4. Empirical analysis and results

We started the study of the total (Tot), domestic (Dom), and international (Int) overnight stays in Sicily, looking for any correlation and similarity between the respective time series, coupled by twos. In particular, we investigated whether domestic and international tourism contributed to the total number of nights spent in Sicily with similar patterns of variation and showed any phase displacement. Same analysis was carried out between domestic and international overnight stays.

Results of the Tot-Dom, Tot-Int, and Dom-Int comparisons are shown in Fig. 3 where the cross-correlation function (CCF) is used to identify lags of the latter time series that might be useful predictors of the former one.

![Figure 3. The cross-correlation function and model statistic.](image)

The plot reported in Fig. 3 shows a clear positive correlation for both national and international tourists in Sicily and the total number of overnight stays for \( h = 0, \pm 1, \) and \( \pm 2 \). For such lags, a positive correlation is also found between Dom and Int nights spent in Sicily. Note instead that the correlations for \( h = \pm 4, \pm 5, \) and \( \pm 6 \) are negative, indicating that an above (below) average value of Dom or Int nights spent in Sicily is likely to lead to a below (above) average value of Tot overnight stays about 4, 5, and 6 months later, respectively. Same correlation holds between Dom and Int.

The information provided by the cross-correlation analysis is enriched with a measure of similarity between the time series considered, known as the cosine similarity metric.

Given two non-zero vectors, \( \mathbf{x} \) and \( \mathbf{y} \), the cosine similarity is:

\[
sim(\mathbf{x}, \mathbf{y}) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2} \sqrt{\sum_i y_i^2}} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} = \cos \theta,
\]
where, $\langle x, y \rangle$ denotes the inner product between $x$ and $y$, and $||x||$ stays for the norm of the vector $x$. Cosine similarity has an interpretation as the cosine of the angle $\theta$ between the two vectors, $x$ and $y$. Therefore, the value of $sim(x, y)$ is bounded between -1 and 1. The more the two vectors have a similar orientation, the more $sim(x, y)$ is close to 1. Two vectors oriented at 90° relative to each other have a similarity of 0, and two vectors diametrically opposed have a similarity of -1, independent of their magnitude.

Results reported in Table 2 show a clear similarity between the time series as the cosine value is very close to 1 for the three comparisons.

<table>
<thead>
<tr>
<th></th>
<th>Tot-Dom</th>
<th>Tot-Int</th>
<th>Dom-Int</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cosine Similarity</td>
<td>0.973</td>
<td>0.956</td>
<td>0.864</td>
</tr>
</tbody>
</table>

Figure 4. The HVG networks.
Said briefly, total, domestic, and international overnight stays in Sicily apparently showed the same behavior over the time horizon considered.

The networks obtained with the HVG algorithm are shown in Fig. 4 (a, b, and c), where the dotted line indicates the point in which the network breaks in two parts.

First of all we answer question: Q1 - Does tourism demand in Sicily show traits and dynamics typical of complex systems? The characteristics of the networks obtained with the HVG algorithm were investigated by first looking at the statistical distribution of the nodal connections (the degree distribution). An exponential or power law distribution of nodal connections is typical of systems that exhibit complex or even chaotic dynamics. In fact, these distributions show long tails that are representative of self-organization and self-similarity features of the system under study.

As shown in Fig. 5a, the three time series analyzed map into networks whose nodal distribution follows an exponential curve. Therefore, the analysis of the degree distribution gives outcomes consistent with the idea of a complex structure for the tourism demand in Sicily. The values of the $\lambda$ exponent are reported in Table 3. The comparison of such values with the critical value $\lambda_c$ gives us some information about the structural properties of the system under study: stochastic ($\lambda > \lambda_c$) or chaotic ($\lambda < \lambda_c$). The higher the slope of the exponential degree distribution, the higher the stability and predictability of the system.

![Figure 5. Cumulative degree distribution for the Sicilian overnight stays HVGs (a) and comparison (b) of the degree exponents $\lambda$ with the null models ($\lambda_c$ = critical value. Vertical bars are 95% confidence intervals).](image)

As shown in Fig. 5a, the three time series analyzed map into networks whose nodal distribution follows an exponential curve. Therefore, the analysis of the degree distribution gives outcomes consistent with the idea of a complex structure for the tourism demand in Sicily. The values of the $\lambda$ exponent are reported in Table 3. The comparison of such values with the critical value $\lambda_c$ gives us some information about the structural properties of the system under study: stochastic ($\lambda > \lambda_c$) or chaotic ($\lambda < \lambda_c$). The higher the slope of the exponential degree distribution, the higher the stability and predictability of the system.

Table 3. Degree distribution exponents for the HVGs examined and the null models with 95% confidence interval (critical value $\lambda_c = 0.405 \pm 0.020$); Rnd = random; fBm = fractional Brownian motion; Lrnz = Lorenz).

<table>
<thead>
<tr>
<th>Series</th>
<th>$\lambda$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tot</td>
<td>0.669</td>
<td>0.082</td>
</tr>
<tr>
<td>Dom</td>
<td>0.707</td>
<td>0.137</td>
</tr>
<tr>
<td>Int</td>
<td>0.496</td>
<td>0.066</td>
</tr>
<tr>
<td>Rnd</td>
<td>0.410</td>
<td>0.038</td>
</tr>
<tr>
<td>fBm</td>
<td>0.684</td>
<td>0.070</td>
</tr>
<tr>
<td>Lrnz</td>
<td>0.301</td>
<td>0.060</td>
</tr>
</tbody>
</table>
Conversely, the lower the value of the exponent, the more the series displays chaotic, and therefore less predictable dynamics. To better interpret our results, a random series (Rnd), a fractional Brownian motion (fBm; the series was generated with Hurst exponent $H = 0.5$), and a series calculated from the well-known chaotic Lorenz map (Lrz) (Parker & Chua, 1989) were used as reference networks. The generation of these networks was repeated ten times and then results averaged. The values of the $\lambda$ exponent are compared with the null models in Fig. 5b where the value $\lambda_c$ and the 95% confidence interval are also represented.

For a comparative analysis, our results are also discussed with reference to the $\lambda$ exponent calculated in Sainaghi and Baggio (2017) for the cases of Livigno and eight Trentino-Alto Adige top destinations (Canazei, Pinzolo, and Riva del Garda from Trento, and Badia, Castelrotto, Merano, Scena, Selva di Val Gardena from Bolzano).

We are then ready to answer to the research question: **Q2 - Provided that the complex nature of the system under study is demonstrated, is its behavior characterized by chaotic, deterministic or stochastic dynamics over time?** All the time series of Sicilian tourism demand are above the chaos threshold line. Values for the $\lambda$ exponent range from 0.496 (international demand) to 0.707 (domestic demand) and are quite close to the range of values calculated for the nine destinations in the North of Italy. Therefore, our results support the hypothesis that “tourism destinations are complex systems but normally do not show chaotic traits” (Sainaghi and Baggio, 2017, p.373).

The degree distributions for the total and domestic tourism demand, shown in Fig. 5a, look very similar, and the $\lambda$ values obtained, 0.707 and 0.669, put the two systems in the region of a fractional Brownian motion series characterized by a predictable and quite stable behavior (Fig. 5b). These values are close to the $\lambda$ exponent calculated for the lake destination of Riva del Garda (0.720), whose overnight stays are mainly concentrated in summer as for the case of Sicily. Yet, the greater stability of the system has a different origin: the domestic market for Sicily and the international one for Riva del Garda. From the inspection of Fig.5a, it is evident the difference between the Int and the Dom tourism demand in the tail of the respective distributions. Looking at the nodal degree distribution, although away from the boundary of chaos, the $\lambda$ exponent calculated for the international component of tourism in Sicily is not in favor of the stability of the network. Actually, the international tourism demand shows dynamics in part attributable to a random series, so as to make its behavior more difficult to predict. This information is fundamental in the planning and development of marketing strategies for tourism by the municipal or regional authorities. Indeed, the non-linear relationships and the self-organization behavior that make a tourism destination a complex system could easily push it towards a chaotic state if not well governed. Badia is the only destination in the Bolzano province showing a structure similar to the international tourism in Sicily. Yet, the case in not discussed in detail in Sainaghi and Baggio (2017) and, therefore, any comparative analysis is precluded.

Our third research question is: **Q3 - Does the time series of overnight stays in Sicily exhibit any symptom of a directional change in the tourism trend?** This brings us back to the groups of nodes displayed with different colors in Fig. 4. These groups are the communities identified in the three time series by the modularity analysis previously described. For a
network derived from a time series, nodes in a community represent periods with the same
dynamics or belonging to the same business/economic cycle.
For the Tot, Dom, and Int networks, the algorithm provided a division into modules with a $Q$
value equal to 0.847, 0.856, and 0.845, respectively, indicating a good separation between the
communities identified.
Therefore, in the time horizon 1998-2017, two different historical periods were found to be
separated by a major change in the dynamics of Sicilian tourism demand.
To confirm such inspection, we looked for a structural break in the overnight stays, that is for
a change in the trend or level of the time series, which, however, did not affect the stationarity
of the series before and after the break. This occurrence signals the presence of some change
or transition in the dynamics and corresponds to an attitude of resilience of the system under
study that adapts to varying environmental conditions without undergoing disruptive
structural modifications.
The econometric literature offers several procedures whose aim is to check for stationarity
when structural breaks are present. For the case of Sicilian tourism, we used the Lee &
Strazicich (2003)’s test, which also provides an estimate of the period where major breaks
occur.
Fig. 6 shows the results of the structural breaks analysis that are in agreement with the visual
inspection of the networks (the series showed in Fig. 6 were filtered in order to better show
the general trend).

![Figure 6. Turning point (vertical line) of the total, domestic, and international tourists’ stays.](image)

Actually, a transition point occurred at a time corresponding to the year 2000 for the domestic
tourism, and to the year 2010 for the international one. The transition found in the
international series coincides with a sharp increase in the international tourism, due to a
change in the tourism routes and destinations in the Mediterranean areas. In 2010, the political
instability and the social disorders in many North African and Middle East countries (Tunisia,
Egypt, Libya, Morocco, Algeria, and so on) moved the flow of international tourists towards
the Sicilian destinations whose tourism offer is comparable to that of the countries cited for
the quality (food, climate, seaside, architecture, and so on) of the journey. Yet, such growth in
the Int series was quite totally counterbalanced by the decrease in the domestic tourism (Dom)
started in 2007, so as to make the total number of nights (Tot) spent in Sicily in 2010 look
stable. This is the reason why the total overnight stays time series shows a break in 2000, but
not for the year 2010. The domestic transition in 2000 corresponds instead to the last peak of national visitors in Sicily before Italy entered a phase of stagnation, characterized by extremely low growth rate for all the following years. In fact, although the adoption of the Euro and the double recession in 2008-09 and 2012-13 have frequently been blamed for the still struggling economy of Italy, real per capita incomes in Italy stagnated between 2000 and 2005, well before the financial crisis (Banca d’Italia, 2018; Papadia, 2017; Romei, 2018). Hence, the breaks in the series correspond to clear changes in the basic dynamics of the system under study.

5. Policy implications and concluding remarks
The investigation carried out with reference to the Sicilian inbound tourism has initially displayed an apparently similar structure between the total, the domestic, and the international overnight stays time series. In fact, the cross-correlation analysis and the cosine coefficients did not reveal any significant difference in the traits of the time series under study. The real nature of the Sicilian destination came out with the use of the HVG algorithm, which revealed the complex structure of the local tourism system characterized by an exponential nodal distribution (degree distribution). This result has already some relevant policy implications. In fact, when dealing with complex systems, local managers and practitioners should abandon the performance measurement systems anchored to past values and financial or operational criteria in favor of non-linear indices.

A further and deeper analysis of the degree distribution displayed different dynamics for the domestic and international overnight stays in Sicily. In particular, the domestic component of the tourism demand showed a quite stable structure characterized by an attitude of resilience of the system. Instead, the $\lambda$ value for the international tourism overnight stays unveiled dynamics in part attributable to a random series so as to make its behavior more difficult to predict. This result should be considered an ‘alarm bell’ for the destination managers and the local firms and push them to implement policies able to give more stability to the market. If not properly managed, the system could tend towards the chaos area and experience unpredictable structural changes leading to the dispelling of the market position gained and a rapid destination decline.

Our third research question brought us to discover a structure for tourism in Sicily made of two communities (modules) corresponding to major changes in the basic dynamics of the system in the years 2000 and 2011, respectively.

The first turning point could be detected just looking at the graph as the downturn is evident only after year 2006. The change in the domestic trend starts in 2000 but becomes clear in the representation of the time series after six years. The Sicilian policymakers would have probably benefited from the anticipated detection of such change in the tourism trend. They could have reached an agreement with low-cost air companies to support tourism in Sicily in spite of the period of stagnation. Instead, the turning point recorded for the year 2011 represents the beginning of a positive trend in the international market that should have promoted investment in the tourism structures of the island. In other words, the random nature of the international tourism demand in Sicily coupled with the information about its ‘rejuvenation’ (Butler, 1980) phase, revealed by the break in the time series, should have
made clear the need for new and ‘ad hoc’ policy interventions and communication plans able to promote investments in Sicily so as to make the international flow of tourists more regular over time.

The proposed study reveals further evidence of the necessity to provide policymakers and DMOs with the information needed for decision making and policy implementation to be effective. Conversely, the complex nature of tourism is very often neglected by destination managers that in the planning and government process make use of traditional metrics and/or linear indicators, implicitly assuming a continuity in the system behavior. Destination structure and dynamics are instead characterized by non-linear relationships, self-organizing behaviors, modular and hierarchical structures, and a more or less attitude to resilience. Thus, there are not rules-of-thumb valid for any destination marketing and management. For the case of Sicily, a lack of a clear structural analysis of inbound tourism has prevented to solve the infrastructure deficiencies and the instability of the international tourism segment that, year by year, are pointed out as a limit to the full economics exploitation of tourism in the island.

6. Limitations and further research

For the nature of the methods used, and the results recalled in this study, there are no particular limitations to mention for the proposed analysis, even if, probably, a longer series or a higher frequency series (weekly data, for example) could be able to provide more details on the general dynamic behavior of the whole system.

Obviously, the main findings of our investigation are influenced by the structure and evolution of domestic and international tourism demand in Sicily. It would be interesting to see whether the complex traits shown by the time series used are confirmed when tourism demand is investigated by region and country of origin, or different turning points and structural characteristics can be found.

Acknowledgements

The authors wish to thank the three anonymous reviewers for their positive and helpful comments. We gratefully thank the “Assessorato Regionale del Turismo dello Sport e dello Spettacolo – Dipartimento Regionale del Turismo dello Sport e dello Spettacolo” and, in particular, Lia Giambrone who supported and provided us with the tourism data.

R.B. acknowledges the financial support of the Ministry of Education and Science of the Russian Federation in the framework of the Competitiveness Enhancement Program of the Tomsk Polytechnic University.

References


Romei, V. (2018). Why Italy’s economy is stagnating. https://www.ft.com/content/b3c85b34-e10a-11e8-a6e5-792428919cee.


