

DETECTING INTERREGIONAL PATTERNS IN TOURISM SEASONALITY OF GREECE: A PRINCIPAL COMPONENTS ANALYSIS APPROACH

Dimitrios TSIOTAS

Adjunct Lecturer, Department of Regional and Economic Development, Agricultural University of Athens, Greece, Nea Poli, Amfissa, 33100, Greece, Department of Planning and Regional Development, University of Thessaly, Pedion Areos, Volos, 38334, Greece, tsiotas@aua.gr
tsiotas@uth.gr

Thomas KRABOKOUKIS

Ph.D. candidate, Department of Planning and Regional Development, University of Thessaly, Pedion Areos, Volos, 38334, Greece
tkrabokoukis@uth.gr

Serafeim POLYZOS

Professor, Department of Planning and Regional Development, University of Thessaly, Pedion Areos, Volos, 38334, Greece
spolyzos@uth.gr

Abstract

Tourism seasonality is a complex phenomenon incorporating a temporal, a spatial, and a socioeconomic (ontological) dimension. This paper builds on principal component analysis (PCA) to provide an integrated methodological framework for studying all three dimensions of tourism seasonality. The proposed method classifies the seasonal patterns of tourism demand of the Greek prefectures into regional groups, which are examined in terms of their geographical and socioeconomic characteristics. The study aims to configure distinguishable seasonal profiles in terms of their socioeconomic attributes. The proposed method is applied to monthly data of tourism overnight stays for the period 1998-2018 and detects seven principal components described by diverse socioeconomic attributes. The overall analysis proposes a useful tool for tourism management and regional policy, it advances PCA to be used as a tool of regional classification, and it incorporates a multivariate consideration based on the socioeconomic evaluation of the principal components. The proposed methodology develops an integrated framework dealing with complexity describing socioeconomic research and particularly tourism seasonality.

Keywords: regional development; seasonal classification; spatiotemporal patterns; pattern recognition.

JEL classification: C18, C38, O52, R10, R58, Z30

1. Introduction

A major aspect in the research of regional science is related to the spatial asymmetry observed in the development of regions, countries, and generally of geographical areas (Polyzos, 2019). The uneven dynamics emerging in space induce inequalities affecting the economic growth, the opportunities for economic development, the quality of the environment, and even the culture and the mentality of societies evolving in time and space (Charles Edwards and Bell, 2013; Romao and Saito, 2017; Batista et al., 2019). Within this framework, the spatial dimension of economic phenomena has become a default variable in contemporary economic analysis (Mastronardi and Cavallo, 2020), in an extent to consider economics and regional science an integrated discipline. This is because, in economic research, which is multivariable covering diverse aspects of socioeconomic life, such as stock-markets (Patatoukas, 2020), energy (Zaman et al., 2016), productivity (Romao and Nijkamp, 2019), entrepreneurship (Hundt and Sternberg, 2016), trade (Brakman and Van Marrewijk, 2017), the web economy (Li et al., 2018), transportation (Cascetta et al., 2015), and other (Kalantzi et al., 2016; Romao et al., 2017; Kummu, M. et al., 2018), the spatial dimension is a common variable controlling either directly or indirectly these socioeconomic aspects. Taking into account that space suggests a default economic-variable, many traditional aspects of economic analysis that were mainly defined within a temporal context, such as productivity (Romao and Nijkamp, 2019), labor (Giannakis and Bruggeman, 2017), energy

(Zaman et al., 2016), tourism (Tsiotas, 2017; Batista et al., 2019), even the web economy (Li et al., 2018), are revisited.

A characteristic case of such reconsideration regards tourism, which suggests a major component for many economies worldwide (Charles Edwards and Bell, 2013; Kalantzi et al., 2016; Batista et al., 2019; Polyzos, 2019). In tourism economics, a main concern of research is dealing with the seasonality of this phenomenon (Butler, 1994; 2001, Gil-Alana, 2010; Polyzos et al., 2013; Kalantzi et al., 2016; Tsiotas, 2017; Ferrante et al., 2018), which is defined as the unequal distribution of demand along the year (Butler, 2001; Batista et al., 2019). Literature research has shown that tourism seasonality is multivariable and is affected by the type of the tourism product (Cuccia και Rizzo, 2011), the climate (Butler, 2001; Fang and Yin, 2015), the social configuration (Almeida and Kastenzholz, 2019), the political regime (Fernandez-Morales et al., 2016), and other factors (Lee et al., 2008). The majority of relevant research mainly focuses on the study of the temporal dimension of tourism seasonality by examining the causes, impacts, and policy implications (Koenig-Lewis and Bischoff, 2005; Duro, 2016), as well as the temporal trends and patterns of demand (Connell et al., 2015; Ferrante et al., 2018; Batista et al., 2019; Duro and Turrion-Prats, 2019). However, all these temporal considerations have an immanent spatial dimension, which is related to the diversity caused by the effect of space and the geographical location of different tourism destinations (Romao and Saito, 2017; Batista et al., 2019). This brings up in the academic dialogue about tourism more avenues of research, such as the study of the competitiveness (Liu et al., 2018; Choe et al., 2019; Gomez-Vega and Picazo-Tadeo, 2019; Niavis and Tsiotas, 2019) and synergy (Niavis and Tsiotas, 2018, 2019; Tsiotas et al., 2019) between tourism destinations, along with the effect of geographical scale, either at the level of neighborhood (Duro, 2016), or at the regional (Romao et al., 2017), international (Batista et al., 2019), and worldwide level (Duro and Turrion-Prats, 2019). For instance, in Europe, tourism is unevenly distributed due to different geographic and socio-economic factors, such as the coastal, insular, and mainland morphology of countries, their cultural background, level of transport integration, and more (Batista et al., 2019). Further, Mediterranean countries are described by a growth-tendency in visitor arrivals that is simultaneously related to a significant increase of seasonality, unlike other competitive destinations, such as the Asia Pacific region that is described by growing demand with a simultaneous decrease of seasonality (Duro and Turrion-Prats, 2019).

A fundamental issue in quantitative studies is the measurement of tourism seasonality (Lundtorp et al., 2001), which is implemented by using a specific variable within a certain time period (e.g. monthly), regardless of their patterns (Porhallsdottir and Olafsson, 2017; Ferrante et al., 2018). The most common variables for measuring tourism seasonality are the number of visitors, arrivals, and overnight stays, while, in terms of economic impacts, income-defined variables are also used (Lundtorp et al., 2001; Porhallsdottir and Olafsson, 2017). Seasonality is also subjected to sensitivity due to subjectivity in the variables' selection (Martin et al., 2019). For instance, the rate and the intensity of seasonality, the seasonal peak factor (S), the maximal utilization constrained by seasonality (MUS), and the seasonality underutilization factor (SUF) were applied to measure seasonality in Cyprus, Sicily (Italy), Madeira (Portugal), and Hiiumaa (Estonia), in the context of increasing seasonality by developing winter tourism (Ruggieri, 2015). Other common indicators used for seasonality measurement in tourism seasonality studies are the seasonality range and ratio, the coefficient of seasonal variation, the seasonality span, the seasonality underutilization factor, and the share of seasonality (Koenig-Lewis and Bischoff, 2005; Duro, 2016). However, due to the complexity describing the phenomenon of tourism seasonality, more composite indicators appeared in literature, such as the Gini coefficient, the Theil index, and the coefficient of variation (CV) (Koenig-Lewis and Bischoff, 2005). These measures can be decomposed to sub-indices and thus include measurements at different scales (Fernandez-Morales et al., 2003; Cisneros-Martinez and Fernandez-Morales, 2015; Duro, 2016; Porhallsdottir and Olafsson, 2017; Rossello and Sanso, 2017). Despite their effectiveness, Gini coefficient and Theil index cannot sufficiently capture periodical (cyclical) structures (Lo Magno et al., 2017; Ferrante et al., 2018), they are restricted to annual computations (Karamustafa and Ulama, 2010), and they provide restricted information about regional seasonality (Cisneros-Martinez and Fernandez-Morales, 2014). Also, they are sensitive to scale around the average (Duro and

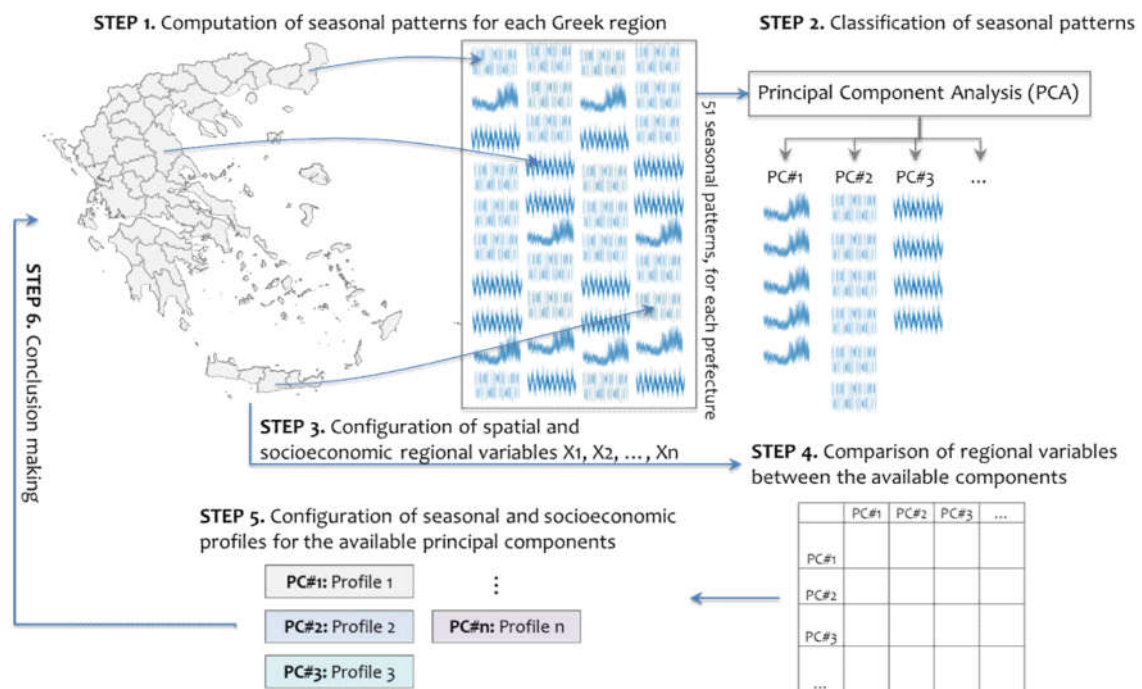
Turrion-Prats, 2019) and to subjectivity in variable's selection (Martin et al., 2019). Duro (2016) attempted a joint consideration of the Gini coefficient, the Theil index and the Coefficient of Variation (CV) to overcome the restrictions of their single use (Koenig-Lewis and Bischoff, 2005). However, in case studies conducted in Spain (Fernandez-Morales, 2003) and Iceland (Porhallsdottir and Olafsson, 2017), the authors showed that these indicators are highly correlated.

Although seasonality affects almost every tourism destination (Corluka et al., 2016), the complex relationship between seasonality and space has not yet been studied in a comprehensive quantitative context (Connell et al., 2015; Corluka et al., 2016; Cisneros-Martinez et al., 2017; Batista et al., 2019; Martin et al., 2019). This is because, on the one hand, many studies are focusing on the multivariable determinants of seasonality (Andriotis, 2005; Gil-Alana et al., 2010; Ferrante et al., 2018; Duro and Turrion-Prats, 2019), but this is mainly done for a single destination, while, on the other hand, many other studies are interested in tourism geography but without deeply examining the temporal tourism patterns emerging in space (Terkenli, 2005; Ahas et al., 2007). Exceptions to this double consideration (Polyzos et al., 2013; Charles Edwards and Bell, 2013; Connell et al., 2015; Romao et al., 2017; Batista et al., 2019) are not enough to configure a comprehensive context for this debate and therefore these works currently highlight the demand of integration between the temporal and spatial dimensions of tourism seasonality. Aiming to serve this demand, this paper introduces a novel approach for studying temporal patterns of tourism seasonality and next classifying them into regional groups. The proposed method builds on Principal Component Analysis (PCA) to classify the 51 (NUTS III) regions in Greece into groups (principal components), which are configured according to their seasonal patterns in terms of visitor-arrivals recorded for the period 1998-2018. The study focuses on the case Greece, which is a coastal country with a mixed mountainous, land, coastal, and insular morphology, consisting of more than 55km² mountainous areas, more than 16,000 km of coastline and more than 1,350 islands, islets, and rocky islands, of which over 230 are inhabited (Tsiotas, 2017). Continental Greek regions occupy 13% of the national population, whereas insular ones occupy 12% (Tsiotas, 2017). The rich geomorphology of Greece has led to a composite tourism product (Kalantzi et al., 2016), which is diversely distributed along with the various tourism destinations of the country (Tsiotas, 2017; Polyzos, 2019). Within this context, the proposed method offers a quantitative tool for measuring and classifying the dynamics of the Greek regions in accordance with their seasonal patterns. This consideration can provide insights about how the geographical distribution of tourism seasonality in Greece can be organized along with the regional space and therefore it may contribute to the configuration of more effective and sustainable tourism development strategies leading the Greek tourism destinations to the desired regional balance.

The remainder of this paper is organized as follows: Section 2 is a brief literature review on tourism seasonality, highlighting its temporal and spatial aspects. Section 3 describes the methodological framework of the study, the available data, and the available variables participating in the analysis. Section 4 presents the results of the analysis and discusses them within the context of regional science and tourism development. Finally, at Section 4 conclusions are given.

2. Methodological Framework

The study aims to provide a methodological framework for studying temporal patterns of tourism seasonality and next classifying them into regional groups. The further purpose of the study is to detect commonalities of tourism seasonality between the Greek regions and to classify them into seasonal profiles. To do so, the proposed methodology builds on the principal component analysis (PCA) (Wold et al., 1987; Norusis, 2008), which is an established technique of dimension reduction useful in various applications. The methodological framework consists of six steps, as it is shown in Fig.1.

Figure 1. The conceptual diagram illustrating the methodological framework of the study

At the first step, the seasonal patterns of the Greek regions are computed on data referring to the monthly number of overnight stays (including both foreign and domestic visitors) per prefecture, for the period 1998-2018. The available data were granted upon request by the Hellenic Statistical Authority (ELSTAT, 2019a) to be used under an exclusive license, for the purpose of this study. At this step, 51 seasonal variables were created, each corresponding to a Greek prefecture (codes and names of the variables are shown see in the Appendix). All seasonal variables are of length 252, namely, they consist of 252 monthly scores composing the period 1998-2018. At the second step, a PCA (Wold et al., 1987; Norusis, 2008) is applied to the available 51 seasonal variables, which are classified into principal components that are coherent groups in terms of variability. In general, PCA is used to reduce the dimension of a set of possibly correlated (source) variables, by converting them into a set of linearly uncorrelated ones, which are called principal components (Norusis, 2008). For n in number available variables, the procedure applies an orthogonal transformation to them, which can be considered as fitting a p -dimensional ellipsoid ($p \leq n$) to the data. Each axis of the ellipsoid corresponds to a principal component. When some ellipsoid axes are relatively small, then the variance along them is also small and therefore the dimension of the available set of variables can be reduced by removing these axes from the dataset. The computational algorithm of the PCA (Wold et al., 1987; Norusis, 2008) is described as follows: first, to find the axes of the ellipsoid, the data are centered on the origin by subtracting the average of each variable from the dataset. Next, the algorithm computes the covariance matrix of the data, the eigenvalues, and the eigenvectors of the covariance matrix. Next, each of the orthogonal eigenvectors is normalized to a unit vector, which configures an axis of the ellipsoid fitted to the data. The total number (p) of the resulting principal components represents an uncorrelated orthogonal basis of the p -dimensional ellipsoid, on which each (of the n in number) source-variable can be projected to. The proportion of the variance each eigenvector captures is calculated by dividing its eigenvalue by the sum of the total eigenvalues. In the PCA algorithm, the resulting principal components are arranged in ascending order, according to which the first has the largest possible variance, the second one the second largest variance, and so on. Provided that not all principal components contribute the same to the total variance, their number can be reduced under a desired loss of information. In the PCA, the choice of the optimum number of principal components is facilitated by plotting them to a scree-plot, which displays an ascending sequence of the components according to the size of their eigenvalues (Norusis, 2008). The final number of principal components is then determined at the point where including more components adds insignificant variance to the total variance is currently

explained. The PCA is broadly used for data reduction in a variety of applications (Kim et al., 2002; Mudrova and Prochazka, 2005; Vyas and Kumaranayake, 2006; Acharya et al., 2012), but is particularly popular in primary research conducted with the use of questionnaires (Norusis, 2008). The principal components resulting by the PCA configure uncorrelated variables of a certain semiology, where items (source variables) within each component are relevant, first, to the extent they best describe the variability of their component and, secondly, to the extent they compose the semiology of their component (Norusis, 2008). In this study, each principal component includes seasonal variables corresponding to Greek prefectures. Within this context, variables included in each principal component express the prefectures that have relevant seasonal patterns during the period 1998-2018.

At the third step of the methodological framework, socioeconomic and geographical (spatial) variables are computed at the regional scale. These variables are of length 51 and include scores of the Greek prefectures for a set of various socioeconomic attributes extracted from the literature, as it is shown in the Appendix. Among these variables, two are included as measures of tourism seasonality, namely the Gini and Relative Seasonality index. The Gini coefficient (Fernandez-Morales et al., 2003; Kulendran and Wong, 2005; Cisneros-Martinez and Fernandez-Morales, 2015; Duro, 2016; Porhallsdottir and Olafsson, 2017) is a very common inequalities measure, which is computed according to the formula:

$$G = 1 - 2 \int_0^1 L(x) dx \quad (1)$$

where $L(x)$ is the mathematical expression of the Lorentz curve (Polyzos, 2019). The Gini coefficient is a stable inequalities measure that is not affected by extreme values (Lundtorp, 2001; Duro, 2016; Duro και Turrion-Prats, 2019). However, its seasonality performance was submitted to criticism about its effectiveness in capturing cyclical structures (Lo Magno et al., 2017; Ferrante et al., 2018). On the other hand, the Relative Seasonality Index (RSI) was proposed by Lo Magno et al. (2017) as a measure of seasonality. This index was defined within the context of the transportation problem, formulated as the problem of minimizing the cost of eliminating seasonality by transferring units from high to low season periods (Lo Magno et al., 2017; Ferrante et al., 2018). The mathematical expression of the RSI is described as follows:

$$S_R(\mu, C) = \frac{\sum_{i \in A} \sum_{j \in B} c_{ij} x_{ij}}{\mu \max_{i \in M} \left\{ \sum_{j \in M} c_{ij} \right\}} \quad (2)$$

where x_i is the i -th observation of variable x , μ is the average value of the available observations, c is the total cost for eliminating seasonality, A is the set of high-season time periods, B is the set of low-season time periods, and M is the set of all possible observed time-patterns.

At the fourth step of the methodological framework, the principal components resulted from the PCA are compared in terms of their socioeconomic and geographical attributes that are shown in the Appendix. This approach builds on the formulation of error-bars of 95% confidence interval for the mean-values (Walpole et al., 2012), which are constructed for each principal component and then are being compared. Cases, where error-bars do not overlay, imply that average values (corresponding to the principal components' groups) are statistically different, under a 95% certainty (Walpole et al., 2012; Tsiotas, 2019). Therefore, comparisons between principal components for every socioeconomic and geographical variable (shown in the Appendix) are expected to reveal the groups (i.e. principal components) with maximum and minimum performance per available attribute. Within this context, at the fifth step of the methodological framework, the results of the comparisons are tabulated to configure seasonal and socioeconomic profiles of the available principal components. This approach develops a classification of attributes determining each principal component and thus it defines the conceptual framework of each principal component in a broader than the seasonal context. The overall approach is expected to provide a tool of quantitative analysis useful for the regional policy and tourism management. The results of

the analysis and the overall approach are discussed at the sixth step of the methodological framework.

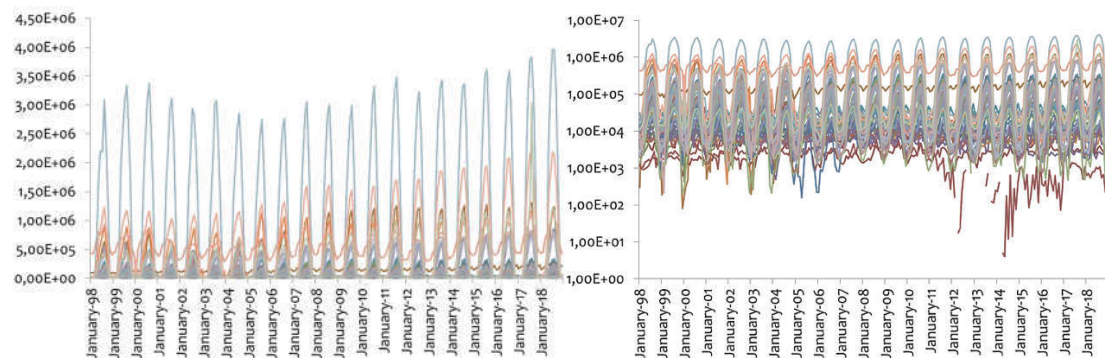
3. Results and Discussion

3.1. Principal component analysis

The available 51 seasonal variables participating in the PCA are plotted in the line-plots shown in Fig.2. Each seasonal variable corresponds to a Greek prefecture and has 252 monthly scores of tourism overnight stays, for the period from Jan 1998 to Dec 2018. As it can be observed, all variables are described by discrete seasonal patterns. In these patterns, we can observe differences in scale (height of oscillation) and trend (e.g. some patterns show increasing trend). Within this context, the PCA is applied to reduce the dimension of this dataset and to organize these diverse patterns into classes (principal components). The further purpose of the analysis is to detect socioeconomic and geographical attributes describing the principal components and to shape a profile describing each group.

The results of the PCA are shown in Fig.3 consisting of four sub-plots. The first (Fig.3a) is the PCA's scree plot showing an ascending sequence of the components according to the size of their eigenvalues (Norusis, 2008). This plot indicates the point after which including more components adds insignificant variance to the total variance is currently explained. According to this plot, 7 principal components can be extracted from the total of 51 available seasonal variables. These principal components explain an amount of ~85% (84.86%) of the total variance. The second sub-plot (Fig.3b) illustrates the coefficients included in the PCA's component matrix (Norusis, 2008), shown on a color scale instead of in absolute numbers. These PCA coefficients illustrate the level at which a source variable is correlated to the resulting principal components, in a context similar to the coefficient of correlation (Walpole et al., 2012). To reduce the complexity of this figure (which includes $51 \times 7 = 357$ coefficients of correlation), the next pair of sub-plots apply maximum (Fig.3c) and minimum (Fig.3d) filters to the information of Fig.3b. In particular, Fig.3c shows with which principal component the Greek prefectures (source variables) are most positively correlated, while Fig.3d shows with which principal component they are most negatively correlated. As it can be observed, the Greek prefectures are all included in the first five components in the maximum coefficients' plot of Fig.3c, whereas are scattered throughout all seven principal components in the minimum coefficients' plot of Fig.3d.

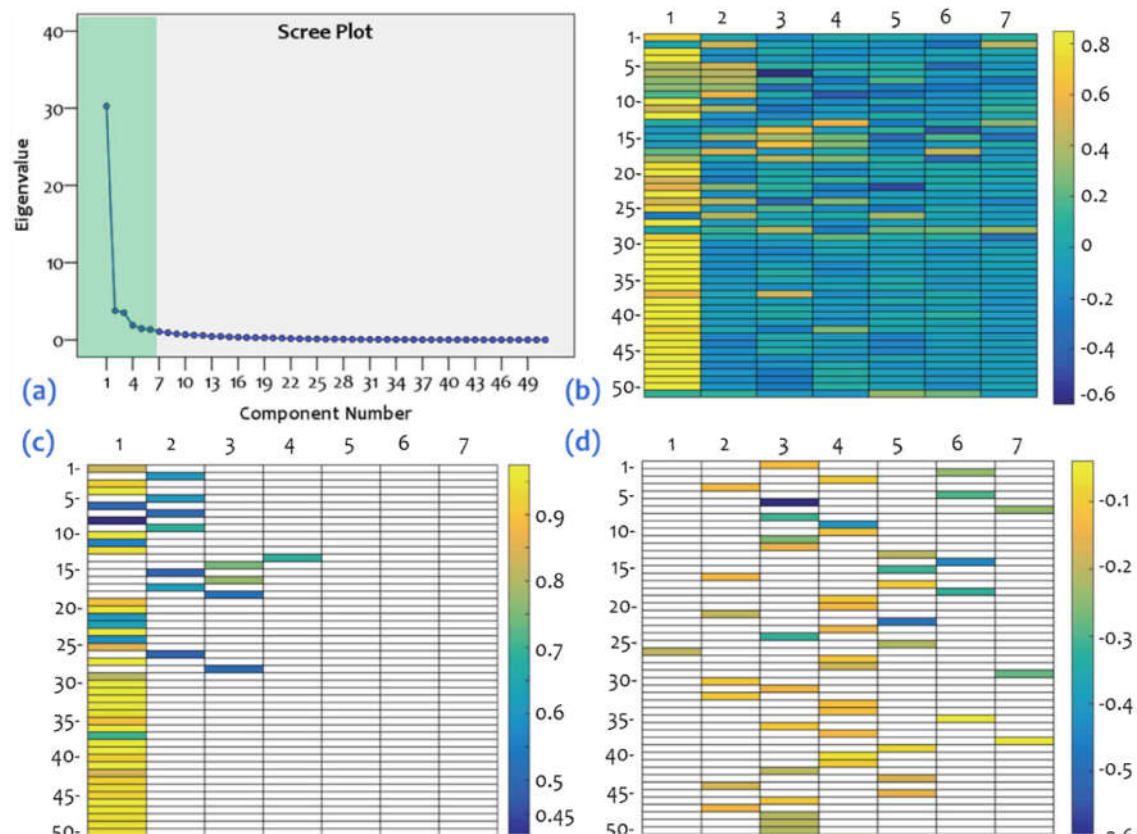
Figure 2. Line plots of all available 51 seasonal variables shown in (left) metric and (right) log scale. Each variable represents a seasonal pattern of tourism overnight stays for a Greek prefecture. Seasonal variables include 252 monthly scores for the period from Jan 1998 to Dec 2018. The available data were granted upon request by the Hellenic Statistical Authority (ELSTAT, 2019a) to be used under an exclusive license, for the purpose of this study



In particular, the first principal component (PC#1) includes maximum coefficients of the prefectures of Rodopi (1), Evros (3), Kavala (4), Thessaloniki (6), Kilkis (8), Pieria (10), Chalkidiki (12), Thesprotia (19), Fthiotida (25), Evoia (27), Fokida (29), Rethymno (50) (prefecture names and coding is shown in the Appendix). These prefectures are distributed throughout the country (except the north-west part of Greece), as it is shown in Fig.4. In particular, the spatial distribution of the prefectures composing PC#1 forms a cluster at the

north-east country, another one at the north, one more at the coastal central part of Greece, another one at west Greece, and a major island cluster at the Aegean sea (at the east part of the country). On the other hand, the prefecture with the minimum (and negative) coefficient included in the first principal component is Viotia (26) located in central Greece. At next, for standardization purposes, principal components defined by the max-value filter of Fig.3 will be denoted as $PC\#i(+)$, with $i=1, \dots, 5$, whereas those defined by the max-value filter of Fig.3 will be denoted as $PC\#i(-)$, where $i=1, \dots, 7$.

Figure 3. Results of the principal component analysis (PCA) applied to the socioeconomic and geographical variables of Table A2, where (a) is the scree plot showing the eigenvalues in accordance to the principal components, (b) is the heat-plot with the total PCA coefficients, (c) is the heat-plot with the maximum PCA coefficients, and (d) is the heat-plot with the minimum PCA coefficients. Columns in cases (b), (c), and (d) express the (6) principal components, whereas rows the available variables

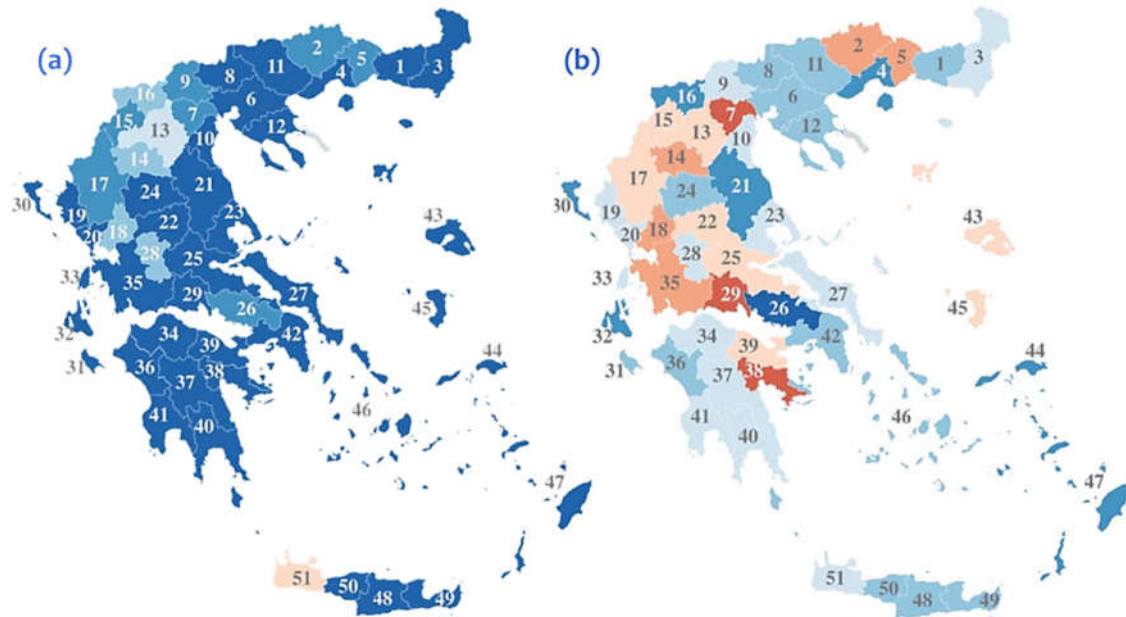


The second principal component ($PC\#2(+)$) includes maximum coefficients (Fig.3c) of the prefectures of Drama (2), Xanthi (5), Imathia (7), Pella (9), Kastoria (15), Ioannina (17), and Viotia (26), as it is shown in Fig.4. These prefectures form three clusters located in the border-arc of the country extended from the west to the east part of Greece, whereas the prefecture of Viotia (26) is located in the central Greece. On the other hand, the prefectures with the minimum coefficients (Fig.3d) included in this component ($PC\#2(-)$) are Kavala (4), Florina (16), Larissa (21), Kerkyra (30), Kefallonia (32), Samos (44), and Dodecanese and they are distributed throughout the Greek periphery.

Next, the third principal component ($PC\#3(+)$) includes maximum coefficients (Fig.3c) of the prefectures of Grevena (14), Florina (16), Arta (18), and Evrytania (28), which are located at the north-west part of the country (Fig.4). These prefectures form a cluster located in the north-west country, another one in the north, whereas the prefecture of Viotia (26) is located in the central Greece. On the other hand, the prefectures with the minimum coefficients (Fig.3d) included in this component ($PC\#3(-)$) are Thessaloniki (6), Kilkis (8), Serres (11), and Chalkidiki (12), at the north, Rodopi (1), at the north-east, Trikala (24), at the central, Zakynthos (31) and Ilia (36), at the west, Attiki (42), Cyclades (46), at the central Aegean, and Heraklion (48) and Rethymno (50), at the island of Crete.

The fourth principal component (PC#4(+)) includes maximum coefficient (Fig.3c) of the prefecture of Kozani (13), at north-west Greece (Fig.4), whereas the prefectures with the minimum coefficients (Fig.3d) included in this component (PC#4(-)) are Evros (3), Pella (9), Pieria (10), Thesprotia (19), Preveza (20), Karditsa (22), Magnesia (23), Evia (27), Evrytania (28), Lefkada (33), Achaia (34), Arkadia (37), Lakonia (40), Mesinia (41), and Chania (51). The geographical distribution of these prefectures forms a heterogeneous pattern scattered throughout the Greek domain.

Figure 4. Regions with the maximum (positive) and minimum (negative) coefficients included in each component



Principal Component	Participation to components' variability	
	Maximum (positive values)	Minimum (negative values)
PC#1	1,3,4,6,8,10-12,19-25,27,29-50	26
PC#2	2,5,7,9,15,17,26	4,16,21,30,32,44,47
PC#3	14,16,18,28	1,6,8,11,12,24,31,36,42,46,48-50
PC#4	13	3,9,10,19,20,23,27,28,33,34,37,40,41,51
PC#5	51	13,15,17,22,25,39,43,45
PC#6	-	2,5,14,18,35
PC#7	-	7,29,38

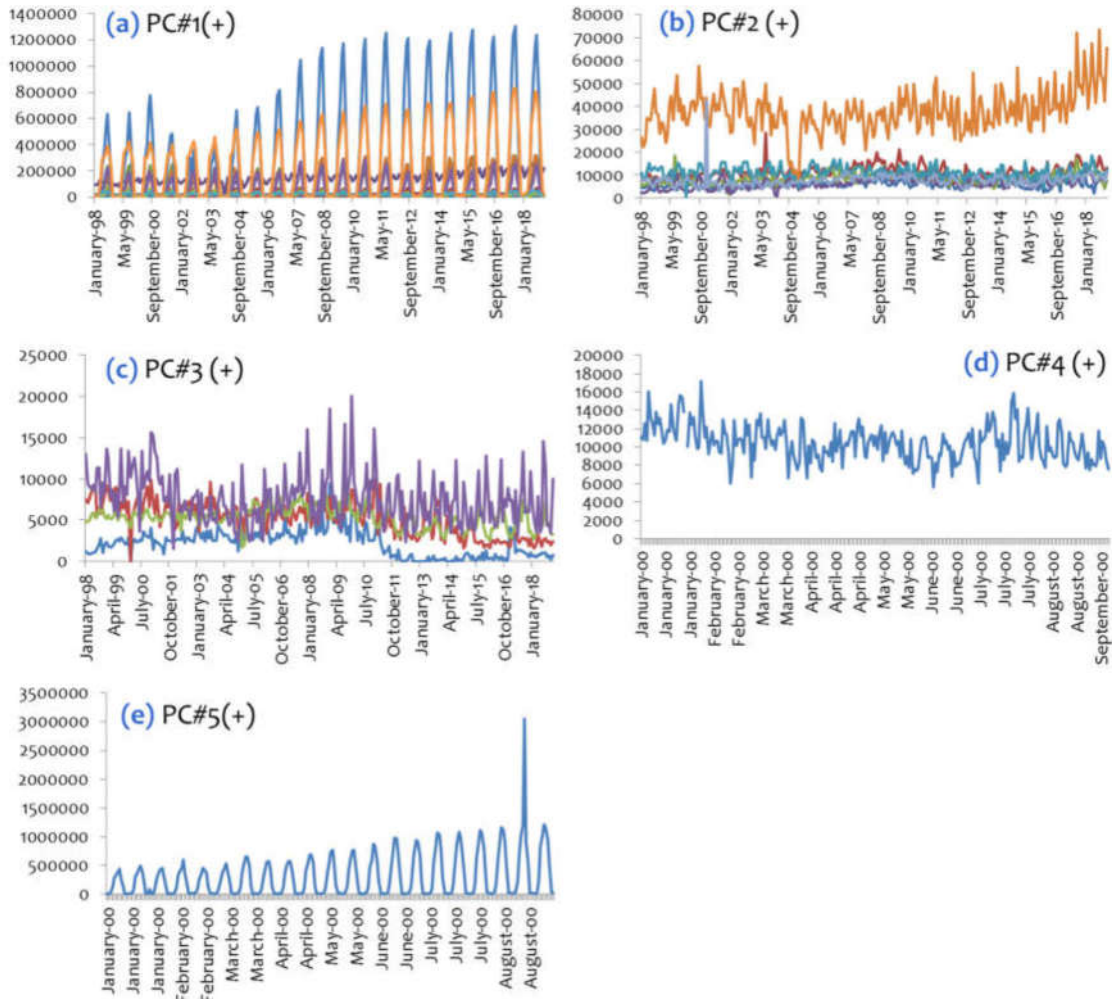
Nomenclature of prefectures shown in the Appendix

Next, the fifth principal component (PC#5(+)) includes maximum coefficient (Fig.3c) of the prefecture of Chania (51), located in the island of Crete (Fig.4), whereas the prefectures with the minimum coefficients (Fig.3d) included in this component (PC#5(-)) are Kozani (13), Kastoria (15), and Ioannina (17), at north-west Greece, Karditsa (22) and Fthiotida (25), at central Greece, Korinthia (39), at the region of Peloponnesus, and Lesvos (43) and Chios (45) and the east Aegean sea. Next, the sixth (PC#6) and seventh (PC#7) principal components do not include maximum coefficients (Fig.3c). The prefectures with the minimum coefficients (Fig.3d) included in PC#6(-) are Drama (2) and Xanthi (5), at the north-east, and Grevena (14), Arta (18), and Aitolokarnania (35), at central-west Greece. Finally, the prefectures with the minimum coefficients (Fig.3d) included in PC#7(-) are Imathia (7), at the north, Fokida (29), at coastal central Greece, and Argolida (38), at the region of Peloponnesus.

To examine the seasonal patterns of these principal components, we construct the line-plots shown in Fig.5 and Fig.6. The first of these figures (Fig.5) shows the seasonal patterns'

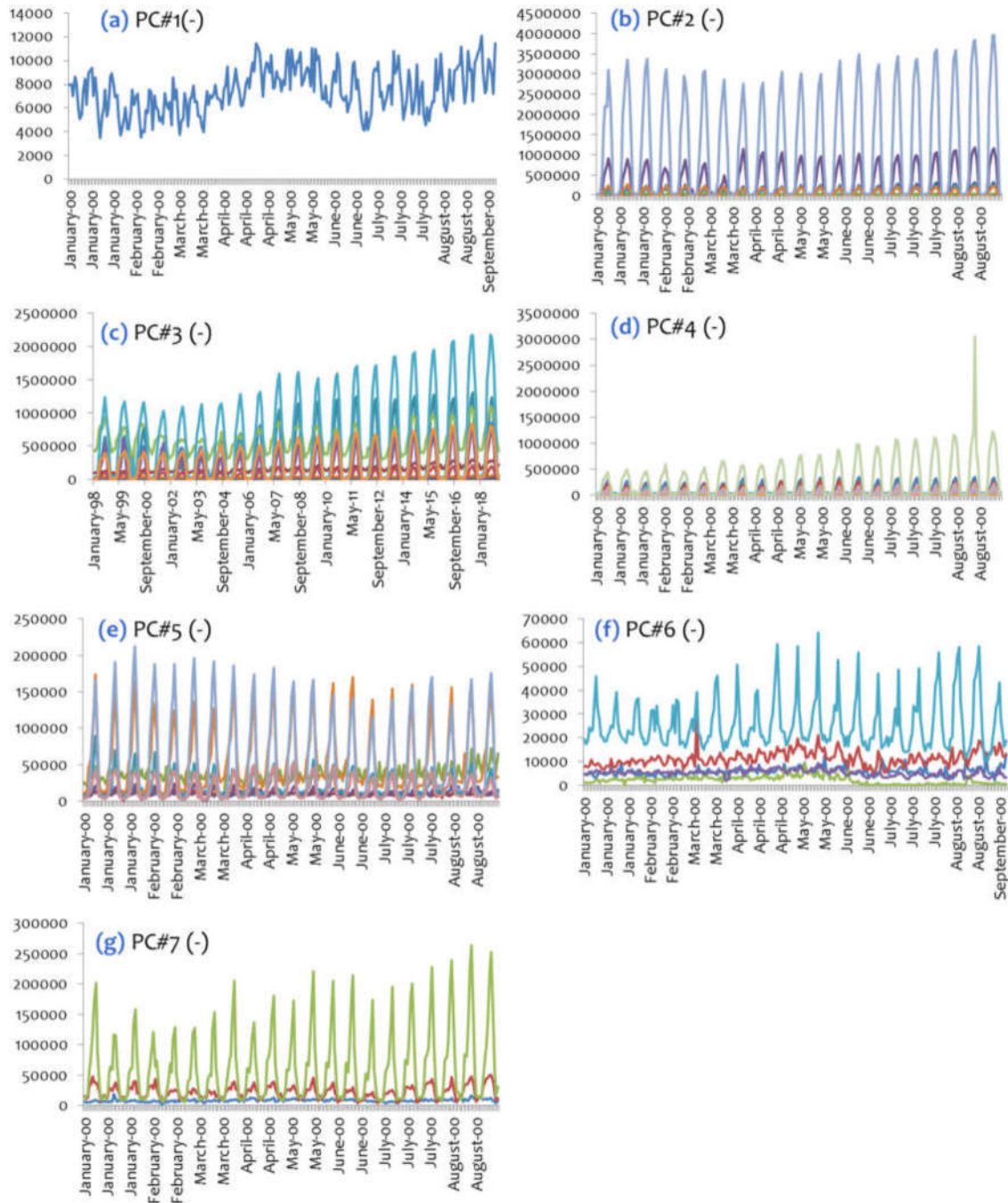
grouping according to the maximum-value filtering (Fig.3c), whereas the second is constructed in accordance with the minimum-value filtering shown in Fig.3d.

Figure 5. Line plots with the available 51 seasonal patterns of the Greek prefectures that are related positively (+) to principal components PC#*i*, with *i*=1,...,5. All displayed components are defined by the max-value filter grouping of Fig.3c



As it can be observed in Fig.5, principal components PC#1(+) and PC#5(+) include prefectures with more discrete periodical (cyclical) patterns, whereas prefectures included in the other components (PC#2(+), PC#3(+), and PC#4(+)) have more noisy patterns. Further, according to Fig.6, principal components PC#2(-), PC#3(-), PC#4(-), PC#5(-), and PC#7(-) include prefectures with more discrete periodical (cyclical) patterns, whereas prefectures included in components PC#1(-) and PC#6(-) have more noisy patterns.

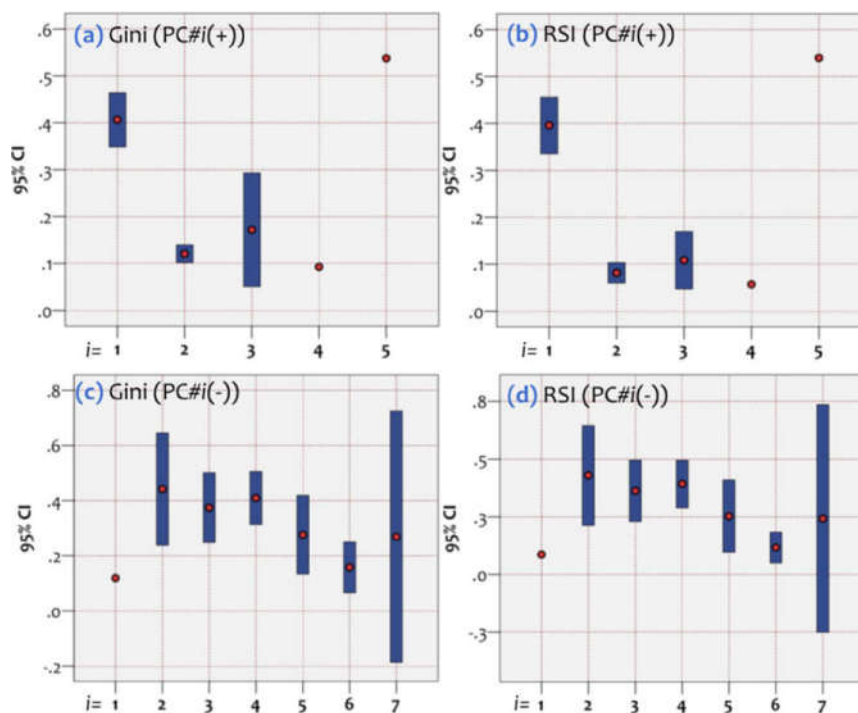
Figure 6. Line plots with the available 51 seasonal patterns of the Greek prefectures that are related negatively (+) to principal components PC#i, with $i=1, \dots, 5$. All displayed components are defined by the min-value filter grouping of Fig.3d



Further, in order to examine the seasonal patterns of these principal components in a less complex way, we construct the error-bars shown in Fig.7. These plots show comparative 95% confidence intervals (CIs) for the mean-values of the available principal components that are computed on the scores of the Gini coefficient (Duro, 2016; Porhallsdottir and Olafsson, 2017) and the Relative Seasonal Index (RSI) (Lo Magno et al., 2017; Ferrante et al., 2018), as they were previously described and defined in relations (1) and (2). As it can be observed in Fig.7a and Fig.7b, principal component PC#5(+) defined by the max-value filter (including the prefecture of Chaneon - 51) has the highest seasonality captured by both indices. The component with the second maximum seasonality is PC#1(+), whereas the pairs of components PC#2(+)-PC#3(+) and PC#3(+)-PC#4(+) are considered to have statistically equal seasonality. This result is in line with Fig.5, where it was observed that principal components PC#1 and PC#5 include prefectures with more discrete cyclical patterns.

On the other hand, none of the principal components $PC\#i(-)$, $i=1,\dots,7$ shows a statistically significant max value, in terms of the min-value filter defined in Fig.3d. This observation illustrates that the max-value filter defined in Fig.3c is more determinative than the max-value filter defined in Fig.3d, in the description of the seasonal patterns of the Greek prefectures. Finally, error-bars in Fig.7 may provide loose and indirect insights about the performance of the two examined indices in capturing seasonality. As it can be observed, the RSI produces CIs of shorter length than the Gini coefficient. Within the context that both seasonality indices are lie on the same scale, this observation implies that the variability within each interval is smaller in the case of RSI and therefore this composite index is capable of producing more homogenous measurement. Consequently, the RSI can be loosely considered as a more effective measure of seasonality than the Gini coefficient.

Figure 7. Error-bars of 95% confidence intervals (CIs) showing the mean-values of (a) the Gini coefficient and (a) the Relative Seasonal Index (RSI) of Lo Magno et al. (2017), both defined by the max-value (+) filter of Fig.3, and (c) the Gini coefficient and (d) the RSI, both defined by the min-value (-) filter of Fig.3



3.2. Socioeconomic determination of seasonal profiles

At this step of the analysis, the principal components resulted from the PCA are compared in socioeconomic and geographical terms. Comparisons build on error-bars of 95% confidence interval for the mean-values, which are constructed for each principal component (similarly to the analysis shown in Fig.7). The results of this comparative approach are shown in detail in the Appendix, where statistically “minimax” performance is illustrated for each component. To facilitate conclusion making, the available variables (see Appendix) are organized into the thematic categories “Geographical”, “Seasonality”, “Transport Infrastructures”, “Demographics”, “Productivity”, “Tourism”, “Environmental”, and “Cultural”.

According to this analysis, the prefectures that are positively related (1,3,4,6,8,10-12,19-25,27, and 29-50) to the first principal component $PC\#1$, have rich coastal configuration (as denoted by the max-value in variable SE.7/COASTAL), high seasonality profile (max in variables SE.3/RSI and SE.4/GINI), high transport integration (maximums in variables SE.5/ROAD DENSITY, SE.6/ROAD LENGTH, and SE.12/AIRPORTS), high specialization in the primary sector (max in variable SE.18/ A_{SEC}), specialization in winter tourism activities (max-values in variables SE.34/SKI CENTERS and SE.35/SKI ROUTES LENGTH), and

high environmental and cultural resources profile (max-values in variables SE.38/PARKS, SE.42/BEACHES, SE.50/BEACHES LENGTH, SE.43/ANC MONUMENTS, SE.44/UNESCO MONUMENTS, and SE.49/CULTURAL RESOURCES). On the other hand, the prefecture that is negatively related (26) to the first component has rich geomorphological configuration (max-values in variables SE.7/COASTAL, SE.9/INLAND, SE.13/AREA, SE.22/TILLING LAND, SE.24/INLAND WATERS, SE.26/LAND AREA, SE.27/SEMI MOUNTAIN AREA), low seasonality profile (min-values in variables SE.3/RSI and SE.4/GINI), rich land transport background (max-values in variables SE.5/ROAD DENSITY, SE.6/ROAD LENGTH, and SE.10/RAIL). Due to the negative contribution of prefecture 26 (Viotia), this component has a poor demographic profile (min-values in variables SE.15/URB and SE.17/HUMAN CAPITAL), high income (max in SE.16/GDP), high secondary sector specialization (max in SE.19/B_{SEC}), low tourism profile (min-values in variables SE.39/HOTELS, SE.45/HOTEL BEDS, SE.46/ROOMS, SE.47/ROOMS BEDS, SE.48/ACCOMMODATION BEDS, SE.40/CAMPING, SE.36/RESTAURANTS, SE.29/MOUNT ACTIVITIES, SE.30/CLIMB FIELDS, SE.31/MOUNT ROUTES, SE.32/RAFTING POINTS, and SE.33/CANYONING POINTS), specialization in winter tourism activities (max in variables SE.34/SKI CENTERS and SE.35/SKI ROUTES LENGTH), high woodland parks capacity (max in variable SE.38/WOODLANDS PARKS), and high cultural resources profile (max-values in variables SE.44/UNESCO MONUMENTS and SE.49/CULTURAL RESOURCES).

Next, the prefectures that are positively related (2,5,7,9,15,17, and 26) to the second component (PC#2) have (on average) a northern location (max in variable SE.1 - LAT), high forest coverage (max in variable SE.23 - FORESTS), poor transport integration (min in variables SE.5/ROAD DENSITY and SE.11/PORTS), specialization in winter tourism activities (max in variables SE.34/SKI CENTERS and SE.35/SKI ROUTES LENGTH), and low cultural resources profile (min values in variables SE.43/ANC MONUMENTS and SE.49/CULTURAL RESOURCES). On the other hand, prefectures that are negatively related (4,16,21,30,32,44, and 47) to the second principal component have rich island configuration (max-values in variable SE.8 -ISLAND), high seasonality profiles (max-values in variables SE.3/RSI and SE.4/GINI), poor rail transport background (min-value in variable SE.10/RAIL), low income (min-value in variable SE.16/GDP), high primary (max-value in SE.18/A_{SEC}) and tertiary sector specialization (max-value in SE.20/C_{SEC}), and high rooms capacity (max values in SE.46/ROOMS and SE.47/ROOMS BEDS).

The prefectures that are positively related (14,16,18,and 28) to the third principal component (PC#3) have northern location (max in variable SE.1 - LAT), mainland geomorphology (max in variable SE.9/INLAND), low seasonality (min in variables SE.3/RSI and SE.4/GINI), poor transportation configuration (min recorded for all variables in this category), low income (min in SE.16/GDP), tourism specialization (min in SE.21/TOURISM GDP) and tourism resources background (min in variables SE.39/HOTELS, SE.45/HOTEL BEDS, SE.46/ROOMS, SE.47/ROOMS BEDS, SE.48/ACCOMMODATION BEDS, SE.40/CAMPING, SE.36/RESTAURANTS, SE.29/MOUNT ACTIVITIES, SE.30/CLIMB FIELDS, SE.31/MOUNT ROUTES, SE.32/RAFTING POINTS, and SE.33/CANYONING POINTS) Also, prefectures contributing positively (14,16,18,and 28) to this principal component have low beach environmental wealth (min in variables SE.41/BLUE FLAG BEACHES, SE.42/BEACHES, SE.50/BEACHES LENGTH, and SE.51/SAND BEACHES LENGTH) and cultural resources profile (min in variables SE.43/ANC MONUMENTS, SE.44/UNESCO MONUMENTS, and SE.49/CULTURAL RESOURCES). On the other hand, prefectures that are negatively related (1,6,8,11,12,24,31,36,42,46, and 48-50) to the third principal component (PC#3) have rich island configuration (max in SE.8/ISLAND), high seasonality (max in variables SE.3/RSI and SE.4/GINI), poor rail transport background (min in SE.10/RAIL), and low income (min in SE.16/GDP). However, they have high primary (max in SE.18/A_{SEC}) and tertiary sector specialization (max in SE.2/C_{SEC}), high tourism performance (max in variables SE.39/HOTELS, SE.45/HOTEL BEDS, SE.48/ACCOMMODATION BEDS, SE.40/CAMPING, SE.36/RESTAURANTS, SE.29/MOUNT ACTIVITIES, SE.31/MOUNT ROUTES), high beach quality (max in SE.41/BLUE FLAG BEACHES) and length (max in SE.51/SAND BEACHES LENGTH).

Next, the prefecture that is positively related (13) to the fourth principal component (PC#4) has a northern location (max in variable SE.1 - LAT), rich geomorphological configuration (max in variables SE.13/AREA, SE.22/TILLING LAND, SE.23/FORESTS, SE.24/INLAND WATERS, SE.26/LAND AREA, and SE.27/SEMI MOUNTAIN AREA), rich rail transport background (max in variables SE.10/RAIL and SE.12/AIRPORTS), high income (max in SE.16/GDP) and secondary sector specialization (max in SE.19/B_{SEC}). However, the prefecture of Kozani (13) has low tourism profile (min values in variables SE.39/HOTELS, SE.45/HOTEL BEDS, SE.46/ROOMS, SE.47/ROOMS BEDS, SE.48/ACCOMMODATION BEDS, SE.40/CAMPING, SE.36/RESTAURANTS, SE.30/CLIMB FIELDS, SE.32/RAFTING POINTS, SE.33/CANYONING POINTS, SE.34/SKI CENTERS, SE.35/SKI ROUTES LENGTH), low environmental wealth (min-values in all variables of the relevant category), and low cultural resources profile (min-values in all variables of the relevant category). On the other hand, the prefectures that are negatively related (3,9,10,19,20,23,27,28,33,34,37,40,41, and 51) to this principal component (PC#4) have rich mountainous configuration (max in variables SE.23/FORESTS and SE.28/MOUNTAIN AREA), high seasonality (max in variables SE.3/RSI and SE.4/GINI), rich airport configuration (max in SE.12/AIRPORTS), high primary (max in SE.18/A_{SEC}) and tertiary sector specialization (max in SE.20/C_{SEC}), high tourism performance (max in variables SE.39/HOTELS, SE.45/HOTEL BEDS, SE.46/ROOMS, SE.47/ROOMS BEDS, SE.48/ACCOMMODATION BEDS, SE.40/CAMPING, SE.36/RESTAURANTS, SE.29/MOUNT ACTIVITIES, SE.30/CLIMB FIELDS, SE.31/MOUNT ROUTES, SE.32/RAFTING POINTS, and SE.33/CANYONING POINTS), high environmental wealth (max in all variables of the relevant category), and ancient monuments resources profile (max in SE.43/ANC MONUMENTS).

The prefecture of Chania (51), Crete, which is positively related to the fifth principal component (PC#5) suggests a unique mixture of the island and mountainous geomorphology (max in variables SE.7/COASTAL, SE.8/ISLAND, SE.9/INLAND, and SE.28/MOUNTAIN AREA), it has high seasonality (max in variables SE.3/RSI and SE.4/GINI), rich port configuration (max in SE.11/PORTS), high tertiary sector (max in SE.20/C_{SEC}), tourism specialization (max in SE.21/TOURISM GDP), and overall tourism performance (max in variables SE.39/HOTELS, SE.45/HOTEL BEDS, SE.46/ROOMS, SE.47/ROOMS BEDS, SE.48/ACCOMMODATION BEDS, SE.40/CAMPING, SE.36/RESTAURANTS, SE.29/MOUNT ACTIVITIES, SE.30/CLIMB FIELDS, SE.31/MOUNT ROUTES, and SE.33/CANYONING POINTS), high environmental wealth (max in all variables of the relevant category), high cultural resources profile (max in variables SE.43/ANC MONUMENTS and SE.49/CULTURAL RESOURCES). On the other hand, the prefectures that are negatively related (13,15,17,22,25,39,43, and 45) to this principal component (PC#4) are northern-located (max in SE.1/LAT), they have high seasonality (max in variables SE.3/RSI and SE.4/GINI), rich airport configuration (max in SE.12/AIRPORTS), low income (min in SE.16/GDP), high primary (max in SE.18/A_{SEC}) and tertiary sector specialization (max in SE.20/C_{SEC}), and rich hotel-infrastructure background (max in SE.39/HOTELS).

The prefectures that are negatively related (2,5,14,18, and 35) to the sixth principal component (PC#6) are northern located (max in SE.1/LAT), they have low seasonality (min in variables SE.3/RSI and SE.4/GINI), poor transport background (min in variables SE.5/ROAD DENSITY and SE.12/AIRPORTS), low income (min in SE.16/GDP), high tertiary sector specialization (max in SE.20/C_{SEC}), rich rafting activities (max in SE.32/RAFTING POINTS), high woodland parks capacity (max in SE.38/WOODLANDS PARKS) and low cultural resources profile (min in variables SE.44/UNESCO MONUMENTS and SE.49/CULTURAL RESOURCES). Finally, the prefectures that are negatively related (7,29, and 38) to the seventh principal component (PC#7) have poor geomorphological configuration (min in variables SE.8/ISLAND, SE.13/AREA, and SE.27/SEMI MOUNTAIN AREA), poor transport background (min in SE.5/ROAD DENSITY, SE.6/ROAD LENGTH, and SE.12/AIRPORTS), low income (min in SE.16/GDP) and industrial specialization (min in SE.19/B_{SEC}), high tertiary sector specialization (max in SE.20/C_{SEC}), and low tourism profile (min in SE.39/HOTELS, SE.32/RAFTING POINTS, SE.33/CANYONING POINTS, and SE.35/SKI ROUTES LENGTH). The previous

observations are summarized in Table 1, which configures the socioeconomic and geographical semiology of the principal components resulted from the previous PCA.

Table 1. The socioeconomic and geographical semiology of the principal components resulted from the PCA

Principal Component	SOCIOECONOMIC AND GEOGRAPHICAL SEMIOLOGY	
	POSITIVELY RELATED ^(a)	NEGATIVELY RELATED ^(b)
PC#1	Rich coastal configuration; High seasonality; High transport integration; High primary sector specialization; Specialization in winter tourism activities; High cultural resources profile.	Rich geomorphological configuration; Low seasonality; Rich land transport background; Poor demographic profile; High income; High secondary sector specialization; Low tourism profile; Specialization in winter tourism activities; High woodland parks capacity; High cultural resources profile.
PC#2	Northern location; High forest coverage, Poor transportation integration; Specialization in winter tourism activities; Low cultural resources profile.	Rich island configuration; High seasonality; Poor rail transport background; Low income; High primary and tertiary sector specialization; High rooms capacity.
PC#3	Northern location; Mainland geomorphology; Low seasonality; Poor transportation configuration; Low tourism resources profile; Low beach environmental wealth and cultural resources profile.	Rich island configuration; High seasonality; Poor rail transport background; Low income; High primary and tertiary sector specialization; High tourism performance; High beach quality and length.
PC#4	Northern location; Rich geomorphological configuration; Rich rail transport background; High income; High secondary sector specialization; Low tourism profile; Low environmental wealth; Low cultural resources profile.	Rich mountainous configuration; High seasonality; Rich airport configuration; High primary and tertiary sector specialization; High tourism performance; High environmental wealth; High ancient monuments resources profile.
PC#5	Mixture of island and mountainous geomorphology; High mountainous coverage; High seasonality; Rich port configuration; High tertiary sector and tourism specialization; High tourism performance; High environmental wealth; High cultural resources profile.	Northern located; High seasonality; Rich airport configuration; Low income; High primary and tertiary sector specialization; Rich hotel-infrastructures background.
PC#6	n/a ^(c)	Northern located; Low seasonality; Poor transport background; Low income; High tertiary sector specialization; Rich rafting activities; High woodland parks capacity; Low cultural resources profile.
PC#7	n/a	Poor geomorphological configuration; Poor transport background; Low income and industrial specialization; High tertiary sector specialization; Low tourism profile.

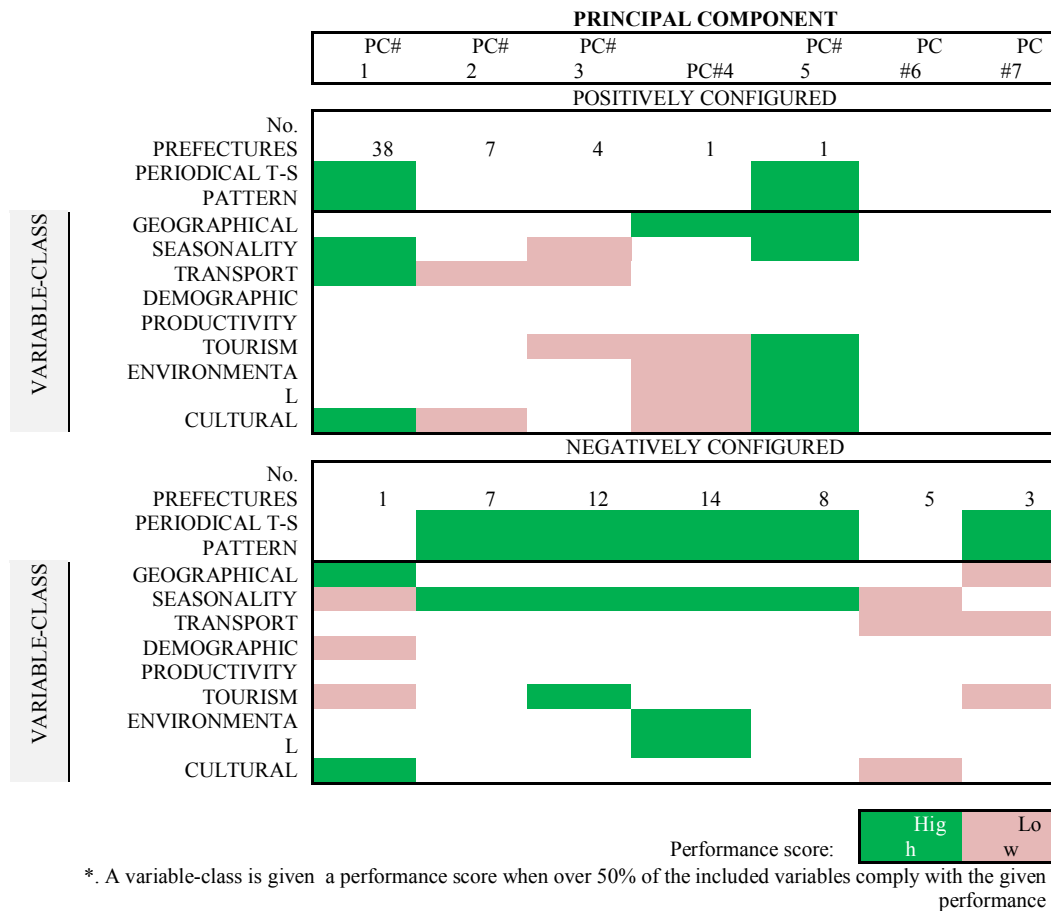
a. Defined by the max-value (+) filter of Fig.3

b. Defined by the m-value (+) filter of Fig.3

c. Not applicable

Further, to facilitate an overall assessment of the semiology of each principal component, we construct Table 2 summarizing the seasonal, geographical, and socioeconomic attributes of the principal components resulted from the PCA. In this table, a (high or low) performance score is given when over 50% of the included variables comply with the given performance. It can be noted that the discrete periodicity observed in line graphs of Fig.5 and Fig.6 is captured by the seasonal measures (RSI and Gini coefficient) used in the analysis. Further, Table 2 allows filtering the major aspects of the socioeconomic attributes describing the principal components.

Table 2. Summary of seasonal, geographical, and socioeconomic attributes^(*) of the principal components resulted from the PCA



Within this context, the first principal component (PC#1) is a component of high seasonality that mainly builds on transport integration (according to the high-performance observed in positively-configured variables) and on cultural characteristics (prefectures that are positively and negatively related to this component have high-performance observed in both positively-configured and negatively-configured variables). The PC#1 does not benefit from demographics and tourism activity (according to the low-performance observed in negatively-configured variables). The second principal component (PC#2) has low transport integration and cultural characteristics (according to the low-performance observed in positively-configured variables) and is competitive to seasonality (according to the high-performance observed in negatively-configured variables). The third principal component (PC#3) is described by low seasonality (according to the low-performance observed in positively-configured and to the high-performance observed negatively-configured variables) and transport integration (according to the low-performance observed in positively-configured variables), and it has low tourism performance (according to the low-performance observed in positively-configured and to the high-performance observed in positively-configured variables). The fourth principal component (PC#4) enjoys rich geomorphology (according to the high-performance observed in positively-configured variables), it is competitive to seasonality (according to the high-performance observed in negatively-configured variables), but it has low environmental (according to the low-performance observed in positively-configured and to the high-performance observed negatively-configured variables), cultural and tourism (according to the low-performance observed in positively-configured variables) performance. The fifth principal component (PC#5) is the privilege to enjoy rich geomorphology, environmental, and cultural welfare (according to the high-performance observed in positively-configured variables) and high tourism activation (according to the high-performance observed in positively-configured variables). The prefectures related to this component are also of high-seasonality (according to the high-performance observed in

positively-configured and the high-performance observed in negatively-configured variables). Finally, the sixth (PC#6) and seventh (PC#7) principal components are competitive to low transport integration (according to the low-performance observed in negatively-configured variables). Component PC#6 is also competitive to seasonality and to cultural features, whereas component PC#7 is competitive to geographical and tourism characteristics (according to the low-performance observed in respective negatively-configured variables). These competitive trends can interpret opposite dynamics in the configuration of the respective components.

4. Conclusions

This paper provided a methodological framework for classifying temporal patterns of tourism seasonality into regional groups. The proposed method built on principal component analysis (PCA) to classify (according to their variability) seasonal patterns of tourism demand of the Greek prefectures into regional groups, for the period 1998-2018. The resulting groups (principal components) were examined in terms of their geographical and socioeconomic characteristics aiming to configure distinguishable seasonal profiles. The analysis resulted in seven principal components and it showed that they are mainly described by distinguishable socioeconomic characteristics. In particular, the first principal component relates its seasonality to high transport integration and cultural resources, whereas the profile of the second component is an inverse (i.e. low transport integration and cultural resources) to the first one. The third principal component relates its seasonality-pattern to low transport integration and tourism activation, whereas the fourth to geomorphological privileges but to low tourism, environmental, and cultural resources. The fifth component relates its seasonality to a privilege mixture of rich geomorphology, environmental, and cultural welfare, whereas the sixth and seventh components relate their seasonality to their competitive profile to transport integration and to cultural (PC#6) and tourism resources (PC#7). The overall analysis can propose a useful tool for tourism management and regional policy, in the context that it deals with complexity in three different dimensions; one temporal related to the seasonality of tourism demand, another related to the geographical diversity of seasonal demand, and a final related to the socioeconomic determinants driving the previous dimensions. The proposed method advances PCA to be used as a tool of regional classification based only on temporal data and incorporates a multivariate consideration based on the socioeconomic evaluation applied to the resulting principal components. The proposed methodology develops an integrated framework dealing with complexity describing socioeconomic research and particularly the seasonality in tourism.

Acknowledgments

This research is co-financed by Greece and the European Union (European Social Fund - ESF) through the Operational Programme “Human Resources Development, Education and Lifelong Learning 2014-2020” in the context of the project “*Analysis and methodological approach of tourism seasonality for the Greek regions*” (MIS 5048961).



5. References

- Acharya, U. R., Sree, S. V., Alvin, A. P. C., Suri, J. S. 2012. “Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework”. *Expert Systems with Applications* 39(10): 9072-9078.
- Ahas, R., Aasa, A., Mark, U., Pae, T., Kull, A. 2007. “Seasonal tourism spaces in Estonia: Case study with mobile positioning data”. *Tourism Management* 28(3): 898-910.
- Almeida, A., Kastenholz, E. 2019. “Towards a Theoretical Model of Seasonal Tourist Consumption Behaviour”. *Tourism Planning & Development* 16(5): 533-555.

- Andriotis, K. 2005. "Seasonality in Crete: Problem or a Way of Life?". *Tourism Economics* 11(2): 207-224.
- Batista e Silva, F., Kavalov, B., Lavalle, C. 2019. *Socio-economic regional microscope series - Territorial patterns of tourism intensity and seasonality in the EU*, Publications Office of the European Union, Luxembourg.
- Brakman, S., Van Marrewijk, C. 2017. "A closer look at revealed comparative advantage: Gross-versus value-added trade flows". *Papers in Regional Science* 96(1): 61-92.
- Butler, R.W. (1994), Seasonality in tourism: Issues and implications. A. Seaton. ed. *Tourism: The state of the art*. Chichester: Wiley.
- Butler, R.W. 2001. Seasonality in Tourism: Issues and Implication, In *Seasonality in Tourism*; Baum, T., Lundtorp, S., (Eds.), Elsevier Ltd.: Oxford, UK.:5–21.
- Cascetta, E., Carteni, A., Pagliara, F., Montanino, M. 2015. "A new look at planning and designing transportation systems: A decision-making model based on cognitive rationality, stakeholder engagement and quantitative methods". *Transport Policy* 38(1): 27-39.
- Chan, F., Lim, C. 2011. "Spectral analysis of seasonality in tourism demand". *Mathematics and Computers in Simulation* 81(1): 1409–1418.
- Charles-Edwards, E., Bell, M. 2013. "Seasonal Flux in Australia's Population Geography: Linking Space and Time". *Population, Space and Place* 21(2): 103-123.
- Choe, Y., Kim, H., Joun, H. 2019. "Differences in Tourist Behaviors across the Seasons: The Case of Northern Indiana". *Sustainability* 11(16): 1-16.
- Cisneros-Martinez, J., Fernando-Morales, A. 2015. "Cultural tourism as tourist segment for reducing seasonality in a coastal area: the case study of Andalusia". *Current Issues in Tourism* 18(8): 765-784.
- Cisneros-Martinez, J., McCabe, S., Fernandez-Morales, A. 2017. "The contribution of social tourism to sustainable tourism: a case study of seasonally adjusted programs in Spain". *Journal of Sustainable Tourism* 26(1): 85-107.
- Connell, J., Page, S., Meyer, D. 2015. "Visitor attractions and events: Responding to seasonality". *Tourism Management* 46(1): 283-298.
- Corluka, G., Mikić, K., Milenkovska, A. 2016. "Classification of tourist season in coastal tourism". *UTMS Journal of Economics* 7(1): 71-83.
- Cuccia, T., Rizzo, I. 2011. "Tourism seasonality in cultural destinations: Empirical evidence from Sicily", *Tourism Management*, 32(3): 589-595.
- Duro, A. 2016. "Seasonality of hotel demand in the main Spanish provinces: Measurements and decomposition exercises". *Tourism Management* 52(1): 52-63.
- Duro, J., Turrion-Prats, J. 2019. "Tourism seasonality worldwide". *Tourism Management Perspectives* 31(1): 38- 53.
- Fang, Y., Yin, J. 2015. "National Assessment of Climate Resources for Tourism Seasonality in China Using the Tourism Climate Index". *Atmosphere* 6(1), 183-194.
- Fernandez-Morales, A. 2003. "Decomposing seasonal concentration". *Annals of Tourism Research* 30(4): 942-956.
- Fernandez-Morales, A., Cisneros-Martinez, J. D., McCabe, S. 2016. "Seasonal concentration of tourism demand: Decomposition analysis and marketing implication". *Tourism Management* 56(1): 172–190.
- Ferrante, M., Lo Magno, G., De Cantis, S. 2018. "Measuring tourism seasonality across European countries". *Tourism Management* 68(1): 220-235.
- Giannakis, E., Bruggeman, A. 2017. "Economic crisis and regional resilience: Evidence from Greece". *Papers in Regional Science* 96(3): 451-476.
- Gil-Alana, L. A. 2010. "International arrivals in the Canary Islands: Persistence, long memory, seasonality and other implicit dynamics". *Tourism Economics* 16(1): 287-302.
- Gil-Alana, L., Barros, C., Assaf, A. 2010. "Persistence in the Short- and Long-Term Tourist Arrivals to Australia". *Journal of Travel Research* 50(2): 213-229.
- Gomez-Vega, M., Picazo-Tadeo, A. 2019. "Ranking world tourist destinations with a composite indicator of competitiveness: To weigh or not to weigh?". *Tourism Management* 72(1): 281-291.
- Hellenic Statistical Authority – ELSTAT 2019a. *Number of monthly overnight stays in the Greek prefectures for the period 1998-2018*. Data granted upon request by ELSTAT (www.statistics.gr) to be used under an exclusive license, for the purpose of this study.

- Hellenic Statistical Authority – ELSTAT 2019b. *Population and social conditions*, available at the URL: <https://www.statistics.gr/el/statistics/pop> [accessed: 03/06/2020].
- Hellenic Statistical Authority – ELSTAT 2019c. *Economy, Indices*, available at the URL: <https://www.statistics.gr/el/statistics/eco> [accessed: 03/06/2020].
- Hellenic Statistical Authority – ELSTAT 2019d. *Environment and energy*, available at the URL: <https://www.statistics.gr/el/statistics/env> [accessed: 03/06/2020].
- Hundt, C., Sternberg, R. 2016. “Explaining new firm creation in Europe from a spatial and time perspective: A multilevel analysis based upon data of individuals, regions and countries”. *Papers in Regional Science*, 95(2): 223-257.
- Hylleberg, S. 1992. *General introduction in Modelling seasonality*. Oxford: Oxford University Press.
- Kalantzi, O., Tsiotas, D., Polyzos, S. 2016. “The contribution of tourism in national economies: evidence of Greece”. *European Journal of Business and Social Sciences (EJBSS)* 5(5): 41–64.
- Kim, K. I., Jung, K., Kim, H. J. 2002. “Face recognition using kernel principal component analysis”. *IEEE signal processing letters*, 9(2): 40-42.
- Koenig-Lewis, N., Bischoff, E. 2005. “Seasonality research: the state of the art”. *International Journal of Tourism Research* 7(4): 201-219.
- Kulendran, N., Dwyer., L. 2010. *Seasonal Variation versus Climate Variation for Australian Tourism: Executive Summary*, Gold Coast, Australia: CRC for Sustainable Tourism Pty Ltd.
- Kulendran, N., Wong, K. 2005. “Modeling Seasonality in Tourism Forecasting”. *Journal of Travel Research* 44(2): 163-170.
- Kummu, M. 2018. “Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015”. *Scientific Data*, 5(1): 1-15.
- Lee, C., Bergin-Seers, S., Galloway, G., O’Mahony, B., McMurray, A. 2008. *Seasonality in the tourism industry – Impacts and strategies*. Australia: Sustainable Tourism Cooperative Research Center.
- Li, J., Xu, L., Tang, L., Wang, S., Li, L. 2018. “Big data in tourism research: A literature review”. *Tourism Management*, 68(1): 301-323.
- Liu, Y., Li, Y., Parkpian, P. 2018. “Inbound tourism in Thailand: Market form and scale differentiation in ASEAN source countries”. *Tourism Management*, 64(1): 22-36.
- Lo Magno, L., Ferrante, M., De Cantis, S. 2017. “A new index for measuring seasonality: A transportation cost approach”. *Mathematical Social Sciences*, 88(1): 55-65.
- Lundtorp, S. 2001. Seasonality in Tourism: Issues and Implication, In *Seasonality in Tourism*; Baum, T., Lundtorp, S., Eds.; Elsevier Ltd.: Oxford, UK, 23-50.
- Lundtorp, S., Rassing, C., Wanhill, S. 1999. “The off-season is “no season: The case of the Danish island of Bornholm”. *Tourism Economics*, 5(1): 49-68.
- Martin, J., Fernandez J.A., Martin, J. 2019. “Comprehensive evaluation of the tourism seasonality using a synthetic DP2 indicator”. *Tourism Geographies*, 21(2): 284-305.
- Mastronardi, L., Cavallo, A. 2020. “The Spatial Dimension of Income Inequality: An Analysis at Municipal Level”. *Sustainability*, 12: 1622.
- Mudrova, M., Prochazka, A. 2005. Principal component analysis in image processing, In *Proceedings of the MATLAB Technical Computing Conference*, Prague, November 2005, available at the URL: http://dsp.vscht.cz/konference_matlab/matlab05/prispevky/mudrova/mudrova.pdf [accessed: 29/1/20]
- Niavis, S., Tsiotas, D. 2018. “Decomposing the price of the cruise product into tourism and transport attributes: evidence from the Mediterranean Market”. *Tourism Management*, 67C: 98-110.
- Niavis, S., Tsiotas, D. 2019. “Assessing the tourism performance of the Mediterranean coastal destinations: a combined efficiency and effectiveness approach”. *Journal of Destination Marketing and Management* (doi:10.1016/j.jdmm.2019.100379).
- Norusis, M., 2008. *SPSS 16.0 advanced statistical procedures companion*, Prentice Hall Press.
- Patatoukas, P. N. 2020. Stock Market Returns and GDP News, *Journal of Accounting, Auditing & Finance*,

- Pegg, S., Patterson, I., & Gariddo, P. V. 2012. "The impact of seasonality on tourism and hospitality operations in the alpine region of New South Wales, Australia". *International Journal of Hospitality Management*, 31(3): 659–666.
- Polyzos, S. 2019. *Regional Development*, Athens: Kritiki.
- Polyzos, S., Tsiotas, D., Kantlis, A. 2013. "Determining the Tourism Developmental Dynamics of the Greek Regions, by using TALC Theory". *TOURISMOS: An International Multidisciplinary Journal of Tourism*, 8(2): 159-178.
- Porhallsdottir, G., Olafsson, R. 2017. "A method to analyse seasonality in the distribution of tourists in Iceland". *Journal of Outdoor Recreation and Tourism*, 19(1): 17-24.
- Romao, J., Guerreiro, J., Rodrigues, P. 2017. "Territory and sustainable tourism development: A space-time analysis on European regions". *The Region*, 4(3).
- Romao, J., Saito, H. 2017. "A spatial analysis on the determinants of tourism performance in Japanese Prefectures". *Asia-Pacific Journal of Regional Science*, 1(1): 243-264.
- Rossello, J., Sanso, A. 2017. "Yearly, monthly and weekly seasonality of tourism demand: A decomposition analysis". *Tourism Management*, 60(1): 379-389.
- Ruggieri, G. 2015. Island Tourism Seasonality. In Pechlaner, E. Smeral. E. ed. *Tourism and Leisure: Current Issues and Perspectives of Development*. Germany: Springer Gabler.
- Shen, S., Li, G., Song, H. 2009. "Effect of Seasonality Treatment on the Forecasting Performance of Tourism Demand Models". *Tourism Economics*, 15(4): 693-708.
- Terkenli, T.S. 2005. "Human activity in landscape seasonality: The case of tourism in Crete". *Landscape Research* 30(2): 221-239.
- Tsiotas, D. 2017a. "Links between network topology and socioeconomic framework of railway transport: evidence from Greece". *Journal of Engineering Science and Technology Review* 10(3): 175-187.
- Tsiotas, D. (2017b). "The imprint of tourism on the topology of maritime networks: evidence from Greece". *Anatolia: An International Journal of Tourism and Hospitality Research* 28(1): 52–68.
- Tsiotas, D. 2019. "Detecting different topologies immanent in scale-free networks with the same degree distribution". *Proceedings of the National Academy of Sciences* 116(14): 6701-6706.
- Tsiotas, D., Niavis, S., Sdrolas, L. 2018. "Operational and geographical dynamics of ports in the topology of cruise networks: the case of Mediterranean". *Journal of Transport Geography* 72: 23–35.
- Tsiotas, D., Polyzos, S. 2017. "The topology of urban road networks and its role to urban mobility". *Transportation Research Procedia* 24C: 482–490.
- Vyas, S., Kumaranayake, L. 2006. "Constructing socio-economic status indices: how to use principal components analysis". *Health policy and planning* 21(6): 459-468.
- Walpole, R. E., Myers, R. H., Myers, S. L., Ye, K. (2012), *Probability & Statistics for Engineers & Scientists*, ninth ed., New York, Prentice Hall Publications.
- Wang, W., Xing, W., Yang, T., Shao, Q., Peng, S., Yu, Z., Yong, B. 2013. "Characterizing the changing behaviors of precipitation concentration in the Yangtze River Basin, China". *Hydrological Process* 27(24): 3375–3393.
- Wold, S., Esbensen, K., Geladi, P. 1987. "Principal component analysis". *Chemometrics and intelligent laboratory systems* 2(1-3): 37-52.
- Zaman, K., Shahbaz, M., Loganathan, N., Raza, S.A. 2016. "Tourism development, energy consumption and Environmental Kuznets Curve: Trivariate analysis in the panel of developed and developing countries", *Tourism Management* 54(1): 275-283.
- Zhang, L., Qian, Y. 2003. "Annual distribution features of precipitation in China and their interannual variations". *Journal of Meteorological Research* 17(2): 146–163.

Appendix

Table A1. The seasonal variables participating in the analysis correspond to the 51 Greek prefectures

Variable Code	Prefecture	Var. Code	Prefecture	Var. Code	Prefecture	Var. Code	Prefecture
1	RODOPI	14	GREVENA	27	EVIA	40	LAKONIA
2	DRAMA	15	KASTORIA	28	EVRYTANIA	41	MESEENIA
3	EVROS	16	FLORINA	29	FOKIDA	42	ATTIKI
4	KAVALA	17	IOANNINA	30	KERKYRA	43	LESVOS
5	XANTHI	18	ARTA	31	ZAKEENTHOS	44	SAMOS
6	THESSALONIKI	19	THESPOZIA	32	KEFALONIA	45	CHIOS
7	HMATHIA	20	PREVEZA	33	LEFKADA	46	CYCLADES
8	KILKIS	21	LARISSA	34	ACHAIA	47	DODECANESE
9	PELLA	22	KARDITSA	35	AITOLOAKARNANIA	48	HERAKLION
10	PIERIA	23	MAGNESIA	36	HELEIA	49	LASITHI
11	SERRES	24	TRIKALA	37	ARKADIA	50	RETHYMNO
12	CHALKIDIKI	25	FTHIOTIDA	38	ARGOLIDA	51	CHANIA
13	KOZANI	26	VIOTIA	39	KORINTHIA		

Table A2. The socioeconomic and geographical variables participating in the analysis

Code	Variable's Symbol	Description	Source
SE.1	LAT	The latitude of the geographical center of the region.	(Google, 2020)
SE.2	LONG	The longitude of the geographical center of the region.	(Google, 2020)
SE.3	RSI	The Relative Seasonal Index of each prefecture computed according to relation (2)	(own elaboration)
SE.4	GINI	The Gini coefficient of each prefecture computed according to relation (2)	(own elaboration)
SE.5	ROAD DENSITY	The road density of each prefecture, defined by the fraction road length/area (km/km ²).	(Tsiotas, 2017a)
SE.6	ROAD LENGTH	The road length of each prefecture (measured in km).	(Tsiotas, 2017a)
SE.7	COASTAL	Indicator variable, returning one (1) to coastal regions and zeros (0) elsewhere.	(ELSTAT, 2020)
SE.8	ISLAND	Indicator variable, returning one (1) to island regions and zeros (0) elsewhere.	(ELSTAT, 2020)
SE.9	INLAND	Indicator variable, returning one (1) to inland (mainland) regions and zeros (0) elsewhere.	(ELSTAT, 2020)
SE.10	RAIL	The length of the rail network included in each prefecture.	(Tsiotas, 2017b)
SE.11	PORTS	The number of ports included in each prefecture.	(ELSTAT, 2020)
SE.12	AIRPORTS	The number of airports included in each prefecture.	(ELSTAT, 2020)
SE.13	AREA	The geographical area of each prefecture (km ²).	(ELSTAT, 2020)
SE.14	POP	The population of each prefecture, according to the 2011 national census.	(ELSTAT, 2020b)
SE.15	URB	Level of urbanization of each prefecture, defined by the proportion of the capital city's population to the total population of the prefecture.	(ELSTAT, 2020b)
SE.16	GDP	The Gross Domestic Product of each prefecture.	(ELSTAT, 2020b)
SE.17	Human Capital	Indicator defined by the proportion of labor force (between 18 and 65 years old) to the total population of the prefecture.	(Polyzos, 2019)
SE.18	A _{SEC}	The prefecture's specialization to the primary sector (% of the GDP).	(ELSTAT, 2020c)
SE.19	B _{SEC}	The prefecture's specialization to the secondary sector (% of the GDP).	(ELSTAT, 2020c)
SE.20	C _{SEC}	The prefecture's specialization to the tertiary sector (% of the GDP).	(ELSTAT, 2020c)
SE.21	TOURISM GDP	The prefecture's specialization to tourism sector (% of the GDP).	(ELSTAT, 2020c)
SE.22	TILLING LAND	The proportion of the tilling land's area to the total area of the prefecture.	(ELSTAT, 2020d)
SE.23	FORESTS	The proportion of the forests' area to the total area of the prefecture.	(ELSTAT, 2020d)
SE.24	INLAND WATERS	The proportion of the inland waters' area to the total area of the prefecture.	(ELSTAT, 2020d)
SE.25	INDUSTRIAL AREA	The proportion of the industrial areas to the total area of the prefecture.	(ELSTAT, 2020d)
SE.26	LAND AREA	The proportion of the land (non-mountainous) areas to the total area of the prefecture.	(ELSTAT, 2020d)
SE.27	SEMI MOUNTAIN AREA	The proportion of the semi-mountain areas to the total area of the prefecture.	(ELSTAT, 2020d)
SE.28	MOUNTAIN AREA	The proportion of the mountain areas to the total area of the prefecture.	(ELSTAT, 2020d)
SE.29	MOUNT ACTIVITIES	The number of mount activities (walking paths, mount sports, climb fields, etc.) in each prefecture.	(ELSTAT, 2020d)

Code	Variable's Symbol	Description	Source
SE.30	CLIMB FIELDS	The number of climb fields in each prefecture.	(ELSTAT, 2020d)
SE.31	MOUNT ROUTES	The number of mountain routes in each prefecture.	(ELSTAT, 2020d)
SE.32	RAFTING POINTS	The number of rafting points in each prefecture.	(ELSTAT, 2020d)
SE.33	CANYONING POINTS	The number of canyoning points in each prefecture.	(ELSTAT, 2020d)
SE.34	SKI CENTERS	The number ski centers in each prefecture.	(ELSTAT, 2020d)
SE.35	SKI ROUTES LENGTH	The length of the ski routes in each prefecture (measured in km).	(ELSTAT, 2020d)
SE.36	RESTAURANTS	The number of restaurants in each prefecture.	(ELSTAT, 2020d)
SE.37	NATURA AREA	The geographical area of Natura parks (areas) in each prefecture.	(ELSTAT, 2020d)
SE.38	WOODLANDS PARKS	The number of woodland parks in each prefecture.	(ELSTAT, 2020d)
SE.39	HOTELS	The number of hotels in each prefecture.	(ELSTAT, 2020c)
SE.40	CAMPING	The number of camping sites in each prefecture.	(ELSTAT, 2020c)
SE.41	BLUE FLAG	The number of beaches granted a blue flag in each prefecture.	(ELSTAT, 2020c)
SE.42	BEACHES	The number organized beaches in each prefecture.	(ELSTAT, 2020c)
SE.43	ANC MONUMENTS	The number ancient monuments sites in each prefecture.	(ELSTAT, 2020c)
SE.44	UNESCO MONUMENTS	The number of UNESCO monuments sites in each prefecture.	(ELSTAT, 2020c)
SE.45	HOTEL BEDS	The number of hotel beds (bed capacity) in each prefecture.	(ELSTAT, 2020c)
SE.46	ROOMS	The number of rooms to let (non-hotel accommodation) in each prefecture.	(ELSTAT, 2020c)
SE.47	ROOMS BEDS	The number of rooms' beds (non-hotel accommodation capacity) in each prefecture.	(ELSTAT, 2020c)
SE.48	ACCOMODATION BEDS	The number of other type of accommodation beds in each prefecture.	(ELSTAT, 2020c)
SE.49	CULTURAL RESOURCES	The number of cultural resources sites in each prefecture.	(ELSTAT, 2020d)
SE.50	BEACHES LENGTH	The length of beaches in each prefecture.	(ELSTAT, 2020d)
SE.51	SAND BEACHES LENGTH	The length of sand beaches in each prefecture.	(ELSTAT, 2020d)

*. All variables have length 51, including scores corresponding to the Greek prefectures

Table A3. “Minimax” comparative table showing the principal components’ performance, according to the available socioeconomic and geographical attributes

Code	Variable	PRINCIPAL COMPONENTS										
		POSITIVELY DEFINED ^(a)					NEGATIVELY DEFINED ^(b)					
		PC#1	PC#2	PC#3	PC#4	PC#5	PC#1	PC#2	PC#3	PC#4	PC#5	PC#6
GEOGRAPHICAL												
SE.1	LAT		MAX	MAX	MAX	MIN	MIN				MAX	MAX
SE.2	LONG			MIN	MIN	MAX						
SE.7	COASTAL	MAX	MIN	MIN	MIN	MAX	MAX			MIN	MIN	
SE.8	ISLAND		MIN	MIN	MIN	MAX	MIN	MAX	MAX		MIN	MIN
SE.9	INLAND		MAX	MAX	MIN		MAX	MIN	MIN			
SE.13	AREA	MIN		MIN	MAX	MIN	MAX					MIN
SE.22	TILLING LAND				MAX	MIN	MAX			MIN		
SE.23	FORESTS		MAX		MAX	MIN	MIN		MAX			
SE.24	INLAND WATERS				MAX	MIN	MAX			MIN		
SE.26	LAND AREA	MAX		MIN	MAX	MIN	MAX			MIN	MIN	
SE.27	SEMI MOUNTAIN AREA		MIN	MIN	MAX	MIN	MAX	MIN				MIN
SE.28	MOUNTAIN AREA	MIN				MAX	MIN			MAX		MAX
SEASONALITY												
SE.3	RSI	MAX		MIN	MIN	MAX	MIN	MAX	MAX	MAX	MAX	MIN
SE.4	GINI	MAX		MIN	MIN	MAX	MIN	MAX	MAX	MAX	MAX	MIN
TRANSPORT												
SE.5	ROAD DENSITY	MAX	MIN	MIN	MIN	MAX	MAX					MIN
SE.6	ROAD LENGTH	MAX		MIN		MAX	MAX					MIN
SE.10	RAIL				MAX	MIN	MAX	MIN	MIN	MIN	MIN	
SE.11	PORTS		MIN	MIN	MIN	MAX						
SE.12	AIRPORTS	MAX		MIN	MAX	MAX	MIN			MAX	MAX	MIN
DEMOGRAPHIC												
SE.14	POP			MIN		MAX						
SE.15	URB	MAX		MAX	MIN	MAX	MIN	MAX				
SE.17	HUMAN CAPITAL	MAX		MIN		MAX	MIN			MAX	MAX	
PRODUCTIVITY												
SE.16	GDP			MIN	MAX		MAX	MIN		MIN	MIN	MIN
SE.18	A _{SEC}	MAX			MIN	MIN	MIN	MAX	MAX	MAX	MAX	MAX
SE.19	B _{SEC}				MAX	MIN	MAX	MIN	MIN	MIN	MIN	MIN
SE.20	C _{SEC}				MIN	MAX	MIN	MAX	MAX	MAX	MAX	MAX
SE.21	TOURISM GDP			MIN		MAX	MIN			MAX		
SE.25	INDUSTRIAL AREA	MIN		MIN	MAX	MIN	MAX					MIN

Code	Variable	PRINCIPAL COMPONENTS											
		POSITIVELY DEFINED ^(a)					NEGATIVELY DEFINED ^(b)						
		PC#1	PC#2	PC#3	PC#4	PC#5	PC#1	PC#2	PC#3	PC#4	PC#5	PC#6	PC#7
TOURISM													
SE.39	HOTELS		MIN	MIN	MIN	MAX	MIN		MAX	MAX	MAX		MIN
SE.45	HOTEL BEDS		MIN	MIN	MIN	MAX	MIN		MAX	MAX		MIN	
SE.46	ROOMS		MIN	MIN	MIN	MAX	MIN	MAX		MAX	MIN	MIN	
SE.47	ROOMS BEDS		MIN	MIN	MIN	MAX	MIN	MAX		MAX	MIN	MIN	
SE.48	ACCOMODATION BEDS		MIN	MIN	MIN	MAX	MIN		MAX	MAX	MIN	MIN	
SE.40	CAMPING	MAX	MIN	MIN	MIN	MAX	MIN		MAX	MAX	MIN	MIN	
SE.36	RESTAURANTS		MIN	MIN	MIN	MAX	MIN		MAX	MAX		MIN	
SE.29	MOUNT ACTIVITIES	MIN		MIN		MAX	MIN		MAX	MAX			
SE.30	CLIMB FIELDS				MIN	MAX							
SE.31	MOUNT ROUTES	MIN		MIN		MAX	MIN		MAX	MAX			
SE.32	RAFTING POINTS	MAX			MIN	MIN	MIN	MIN		MAX		MAX	MIN
SE.33	CANYONING POINTS		MIN	MIN	MIN	MAX	MIN	MIN		MAX			MIN
SE.34	SKI CENTERS	MAX	MAX		MIN	MIN	MAX	MIN	MIN	MIN			
SE.35	SKI ROUTES LENGTH	MAX	MAX		MIN	MIN	MAX	MIN	MIN	MIN	MIN	MIN	MIN
ENVIRONMENTAL													
SE.37	NATURA AREA				MIN	MAX	MIN			MAX			
SE.38	WOODLANDS PARKS	MAX	MAX		MIN	MAX	MAX					MAX	MIN
SE.41	BLUE FLAG BEACHES		MIN	MIN	MIN	MAX	MIN		MAX	MAX		MIN	
SE.42	BEACHES	MAX	MIN	MIN	MIN	MAX	MIN			MAX			
SE.50	BEACHES LENGTH	MAX	MIN	MIN	MIN								
SE.51	SAND BEACHES LENGTH		MIN	MIN	MIN	MAX	MIN	MAX	MAX	MAX			
CULTURAL													
SE.43	ANC MONUMENTS	MAX	MIN	MIN	MIN	MAX	MIN			MAX			
SE.44	UNESCO MONUMENTS	MAX		MIN	MIN	MIN	MAX		MIN	MIN	MIN	MIN	
SE.49	CULTURAL RESOURCES	MAX	MIN	MIN	MIN	MAX	MAX						MIN

a. Defined by the max-value (+) filter of Fig.3 b. Defined by the m-value (+) filter of Fig.3