Università
Ca’Foscari
Venezia

Master’s Degree

in Economics and Finance

Final Thesis

An Econometric Analysis of Touristic Flows

Supervisor
Ch. Prof. Roberto Casarin

Assistant supervisor
Ch. Prof. Elisa Barbieri

Graduand
Emanuele Lopetuso
Matricolation number
877685

Academic Year
2019 / 2020

Contents

1 Social Network Analysis **5**
1.1 Introduction 5
1.2 Graph 6
1.3 Network Metrics 7
1.3.1 Local Metrics 8
1.3.2 Global Metrics 11
1.4 Applications 12

2 Structural Time Series **14**
2.1 Introduction 14
2.2 State-Space Form 17
2.2.1 Trend 17
2.2.2 Cycle 18
2.2.3 Seasonality 19
2.2.4 Daily Effect 21
2.2.5 State Vector 21
2.3 State estimation 23
2.3.1 Kalman filter 24
2.3.2 Kalman Gain 26
2.3.3 The Algorithm 27
2.3.4 Kalman Smoother 28
2.3.5 R-T-S Algorithm 28

3 Review **30**
3.1 Introduction 30
3.2 Econometric Models 31
3.2.1 Static Regression Model 32
3.2.2 DL and ADL Model 32
3.2.3 ECM 33
3.2.4 Panel Data Regression Models 34
3.2.5 Other on Econometric Models 34

3.3 Time-Series Models . . . 35
3.3.1 Basic Techniques . . . 35
3.3.2 Advanced Techniques . . . 37
3.4 Mixed Models . . . 38
3.5 Artificial Intelligence Models . . . 39
3.5.1 ANN . . . 40
3.5.2 SVR . . . 40
3.5.3 Fuzzy Time Series . . . 41
3.5.4 Rough Set Approach . . . 41
3.5.5 Grey Theory . . . 41
3.5.6 Hybrid Models . . . 42
3.6 Model Comparison . . . 42

4 Data and Methodology 45
4.1 Data . . . 45
4.1.1 Argentina . . . 46
4.1.2 Australia . . . 46
4.1.3 Belgium . . . 47
4.1.4 Brazil . . . 47
4.1.5 Canada . . . 47
4.1.6 France . . . 48
4.1.7 Germany . . . 48
4.1.8 Greece . . . 48
4.1.9 Indonesia . . . 49
4.1.10 Israel . . . 49
4.1.11 Italy . . . 49
4.1.12 Japan . . . 49
4.1.13 Mexico . . . 49
4.1.14 Netherlands . . . 50
4.1.15 New Zealand . . . 50
4.1.16 Poland . . . 50
4.1.17 Portugal . . . 50
4.1.18 South Africa . . . 51
4.1.19 South Korea . . . 51
4.1.20 Spain . . . 51
4.1.21 Thailand . . . 52
4.1.22 Turkey . . . 52
4.1.23 United Kingdom . . . 52
4.1.24 United States of America . . . 52
4.1.25 UNWTO . . . 53
4.2 Methodology . . . 53

5 Results **56**
5.1 Past Networks 56
5.1.1 Inbound Analysis 56
5.1.2 Outbound Analysis 58
5.2 Time Series Analysis 59
5.3 Backtesting 74
5.4 Future Networks 76
5.4.1 Inbound Analysis 76
5.4.2 Outbound Analysis 81

6 Covid-19 Impact **86**
6.1 Border Restrictions 87
6.2 Internal Restrictions 89
6.3 Networks 90
6.4 Conclusion 91

A Tables of Results Chapter 4 **95**

B Backtesting Time-Series **102**

C Restriction-Period Table **106**

Abstract

The tourism industry has grown rapidly in the last few decades, becoming one of the main economic activity for the greatest part of countries. This thesis aims to analyze and forecast tourist flows among 24 countries, creating a network where the vertices represent the nations and the weights of the directed edges represent the tourist flow between pairs of countries. The network is made dynamic since the observations collected are monthly, allowing me to assess the intra-annual and long-period changes of network structure. The network dynamic is forecasted by means of structural time series analysis which allows for non-stationary components such as trend and seasonality that are strongly present in touristic data. The thesis is divided into six chapters. In the first two, it is provided an introduction to the two tools used in this thesis. In chapter one, there is a brief explanation of social network theory, while in the second is presented the structural time series analysis and the Kalman filter and smoother. The third chapter is a literature review of the forecasting methods for the prediction of the touristic flows. The fourth one presents the methodology and the data used for the analysis, specifying the sources and the collection mechanism. The results are provided in chapter 5. In the last chapter is made an analysis of the mobility restrictions due to Covid-19 case, trying to forecast the impact of the limitations on the future touristic flows.

Chapter 1

Social Network Analysis

1.1 Introduction

The idea behind the social network analysis is that it is possible to better understand things if viewed as an intersection of particular interactions instead as isolate entities. The primacy idea of network analysis can be found in works of influential thinkers of the past such as Heraclitus, and on the works of some of the most important sociologists such as Marx, Durkheim and Weber [1]. Simmel [2] was the first that explicitly outlined the importance of study interactions rather than single individuals. Although he developed theories of many forms, he did not formalize them in methemathical form, which were developed in more recent years. Today social network analysis has become an interdisciplinary area of studies, with application in anthropology, communication, computer science, education, economics, management science, medicine, political science, public health, psychology and other disciplines [3]. Today, Social Network analysis has progressed enormously [4], becoming a thriving research area. For instance, how outlined by S. Klassen in its speech for the inauguration of the academic year of the university of toronto, in Canada, the network based projects accounted for the fourth largest share of grants dispensed by the Social Science Research Council between 1998 and 2007, receiving the highest per-project grants. One of the strengths of this tool is that it allows both qualitative and quantitative analysis. The former is the one that has experienced the most the advances with the creation of specific network measures. There has been no comparable development on the qualitative side [5]. This chapter provides a very brief introduction to the social network analysis, explaining the mathematical formulation, the main measures and reporting some of the several fields of application.

1.2 Graph

Graph is a network of dots and lines, in particular, is an object composed of two sets, called vertex set, denoted by V, and edges set, denoted by E. The former is a non-empty set while the latter can be empty, but otherwise, its elements are two elements subsets of the vertex set [6]. Let G a graph composed of the sets $V = \{A, B, C, D\}$ and $E = \big\{\{A, C\}\big\}$. $\{A, C\}$ is an edge of the graph, which represents a relationship of such nature between the element A and the element C. The absence of other edges means that there are no links among the other vertices. The elements of V with 0 connections are called isolated. If the relationships hold only in a direction, the graph G is called directed. This means that the connection starts from A and arrives in C, but it is not mutual. For example, if the vertices represent people at a party, and there is an edge between two people if they shake hands, then this graph is undirected because any person A can shake hands with a person B only if B also shakes hands with A. In contrast, if any edge from a person A to a person B corresponds to A owes money to B, then this graph is directed because owing money is not necessarily reciprocated. If not all the edges have the same importance, it is possible to add a weight to them [7]. Such weights might represent for example costs, lengths or capacities, depending on the problem at hand. Let assume that G is a graph describing the relationships among the users of a social network, for instance Facebook. Assuming that there are three users, John, Jerry and Jack Doe, and John is "friend" with both Jack and Jerry. In Facebook the "friendship" is mutual, it is not possible to be linked with another user if he is not connected with you. For this reason the network is undirected. A graphical representation of this situation can be found in figure 1.1.

The sets are

$$G = \{V, E\} \tag{1.1}$$

$$V = \{John, Jerry, Jack\} \tag{1.2}$$

$$E = \big\{\{John, Jack\}, \{John, Jerry\}\big\} \tag{1.3}$$

Now assuming that the social network works as Instagram, where it is possible to follow a user without being followed. In this case the network is directed, and assuming that John is followed by Jack, and Jerry follows both Jack and Jerry, the resulting graphical representation is reported in figure 1.2, where the direction of the edge is displayed with an arrow.

The sets are

$$G = \{V, E\} \tag{1.4}$$

$$V = \{John, Jerry, Jack\} \tag{1.5}$$

$$E = \big\{\{Jack, John\}, \{Jerry, John\}, \{Jerry, Jack\}\big\} \tag{1.6}$$

Assuming now that we want study the interation among users, taking into account the "likes" exchanged among them. It is possible to do it adding a number to the edges. Assuming that Jack likes 30 posts of John, and Jerry likes 2 posts of John and 12 posts of Jack. To represent this scenario usually the size of edges is modified based on the weight, as showed in figure 1.3.

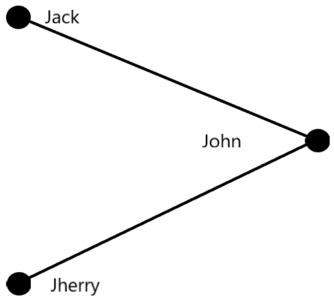


Figure 1.1: Undirected Network

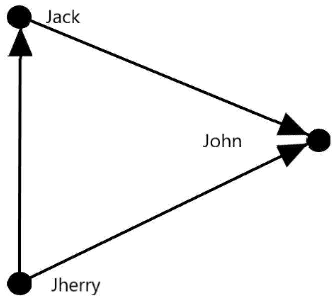


Figure 1.2: Directed Network

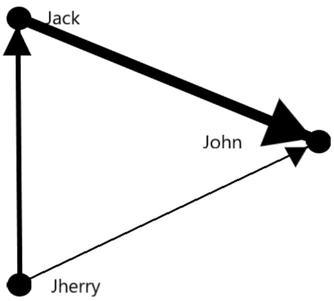


Figure 1.3: Weighted Directed Network

The resulting sets are

$$G = \{V, E\} \tag{1.7}$$

$$V = \{John, Jerry, Jack\} \tag{1.8}$$

$$E = \big\{\{Jack, John, 30\}, \{Jerry, Jack, 12\}, \{Jerry, John, 2\}\big\} \tag{1.9}$$

1.3 Network Metrics

In this section, I am going to present the most used measures to describe a network. I distinguished among global and local metrics. The former describes features of the whole graph, while the latter concentrate on describing the single node.

1.3.1 Local Metrics

Local metrics describe the importance of a node. The most important of such type of metrics are the centrality measure. Their names dervies from the fact that they assess how central (important) is a node in the network. The word "important" has different meaning, leading to several definitions of centrality. Below, it is reported a list of the most important and most used centrality measures.

Degree Historically first and conceptually simplest is degree centrality, which is defined as the number of ties that a node has. In the case of a directed network, it is possible to distinguish in-degree and out-degree value, the former is a count of the number of ties directed to the node, and the latter is the number of ties that the node directs to others. The degree centrality of a node v in a given network $G = \{V, E\}$, with $|V|$ the number of vertices and $|E|$ the number of edges, is defined as

$$C_D(v) = deg(v) \tag{1.10}$$

In an undirected network the degree can be seen as the risk to be affected by a given event, for instance, if the nodes are people and the edges the link among them, the degree can describe the exposure to some contagious disease. Instead, in a directed network, the in-degree is a measure of popularity and and outdegree a measure of gregariousness.

Closeness In a graph, the closeness centrality of a node describes how close is a node to all the other ones. This measure was proposed by Bavelas in 1950 in its work "communication patterns in task-oriented groups" [8]. It is computed as the reciprocal of the sum of the length of the shortest paths between the node and all other nodes in the graph. The shortest path among two edges is the path between two vertices in a graph such that the sum of the weights of its constituent edges is minimized. Let $d(v_1, v_i)$ the shortest path between nodes 1 and i, the closeness centrality for node 1 is

$$C_c(v_1) = \frac{1}{sum_{i=2}^{|V|} d(v_1, v_i)} \tag{1.11}$$

To allow comparison between graphs of different sizes, the closeness centrality is often used in its normalized form, obtained multiplying the previous equation by $|V| - 1$. For large graphs, this difference becomes inconsequential, so the -1 is dropped resulting in

$$C_c(v_1) = \frac{|V|}{\sum_{i=2}^{|V|} d(v_1, v_i)} \tag{1.12}$$

The measure computes how much passages in average a node must do to influence the other vertices.

Betweeness Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. It was introduced by Linton in 1977 [9]. Calling $\sigma_{i,j}$ the number of shortest path between nodes i and j, and $\sigma_{i,j}(v)$ the number of shortest path between i and j passing through v, the betweenness of v is computed as

$$C_b(v) = \sum_{i \neq j \neq v} \frac{\sigma_{i,j}(v)}{\sigma_{i,j}} \tag{1.13}$$

To resume, for a pair of vertices the shortest paths are computed, then it is determined the fraction of shortest paths that pass through the vertex in question. This process is repeated for each possible pair of vetices, then the results are summed. Betweenness centrality finds wide application in network theory. It represents the degree to which nodes stand between each other. For example, in a telecommunications network, a node with higher betweenness centrality would have more control over the network, because more information will pass through that node.

Eigenvector Centrality Eigenvector centrality, proposed for the first time by Landau [10], measures the influence that a node has in the network. The aim is to assign relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute to the score of the node in question more than equal connections to low-scoring nodes. The score x of node v is assigned by means of the following equation

$$x_v = \frac{1}{\lambda} \sum_{i=1}^{|V|} a_{v,i} x_i \tag{1.14}$$

where $a_{v,i}$ is equal to 1 if node v and i are linked, and 0 otherwise. In other words, the score is the sum of the scores given to the linked nodes, divided by a constant. Other variants of the eigenvector centrality had been proposed in the literature, the most known are the Katz centrality [11] and the Google's page Rank, used by Google to order search results [12].

There are many other metrics to rank the importance of a node in a graph, each one based on a different definition of centrality. In figure 1.4 are displayed the differences among the measures previously presented, where the closer to blue is the colour of an edge, the lower is the value of the measure, and the opposite when it approaches to red. It is strongly evident how changing the definition of centrality, the importance of nodes varies widely.

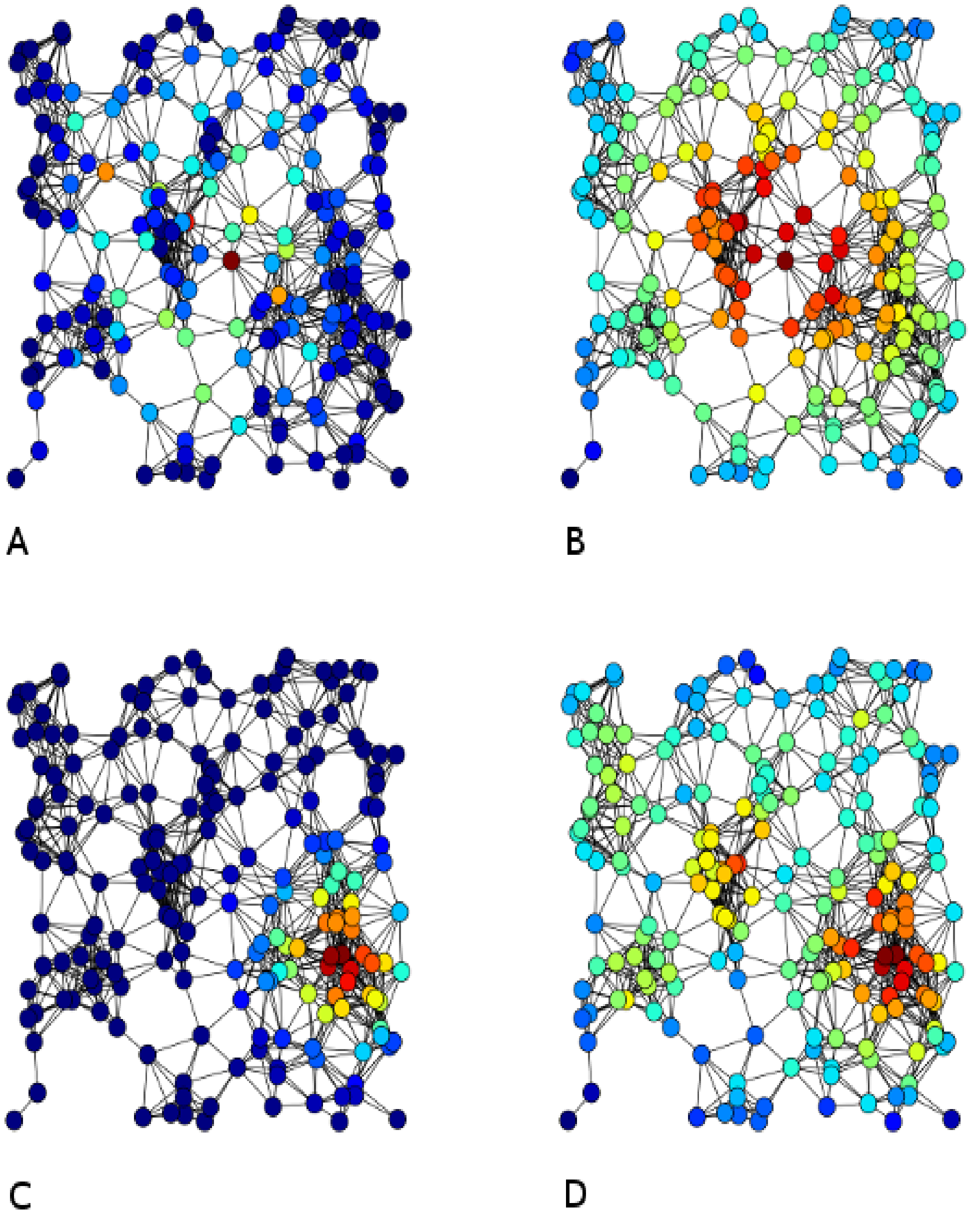


Figure 1.4: A. Betweenness B.Closeness C.Eigenvector D. Degree

1.3.2 Global Metrics

Size Measures The two most basic global measures are the number of nodes $|V|$ and the number of edges $|E|$. These measures give to the analyst an idea of the network size, both in term of actors (nodes) and relationships among them (edges). Other basic measures that describe the size of a network are the maximum and minimum number of edges, denoted by E_{max} and E_{min} respectively.

Average Degree A useful measure is the average degree, which allows understanding which is on average the degree of a node. Let C_d the degree of node v and $|V|$ the number of nodes. The average degree is computed as

$$AverageDegree = \frac{1}{|V|} \sum_{i=1}^{|V|} d_i \tag{1.15}$$

In a weighted graph, the average degree is computed as a weighted mean.

Average Shortest Path length The average shortest path length is calculated by finding the shortest path between all pairs of nodes, and computing the mean. This shows us, on average, the number of steps it takes to get from one member of the network to another.

Diameter As another means of measuring network graphs, we can define the diameter. It is a representation of the linear size of a network and gives an indication of how extended is a graph. It is calculated as the longest of all the shortest paths in a network.

Density This measure gives an overview of the density of a graph. The closer is to one, the more dense is the graph, while the opposite is true when it approaches to zero. It is computed dividing the actual number of edges $|E|$ by the possible number of them. Each vertex can be connected at most with $|V-1|$ other vertices, so the maximum number of edges that a network can achieve is $|V||V-1|/2$ or $|V||V-1|$ if undirected or directed respectively. So the equation in the former case is

$$D = \frac{2|E|}{|V||V-1|} \tag{1.16}$$

Cliques A clique is a subset of nodes such that all elements in the clique are fully connected. The clique number of a graph equals the largest clique available in the network. The first definition of cliques was given by Luce and Perry to study the society [13]. However it was widely used to handle problems in several other subjects, for instance in bioinformatic [14] and chemistry [15].

Cluster The clustering coefficient of a node is the ratio between the existing links connecting a node's neighbours and the maximum possible number of such edges. The clustering coefficient for the entire network is the average of the clustering coefficients of all the nodes. The equation for the clustering coefficient of a single node v is

$$Clus_v = \frac{2e_v}{k_v(k_v - 1)} \tag{1.17}$$

where k_v is the number of nodes connected with v and e_v is the number of connections between these neighbours. A high cluster coefficient for a network is an indication of small world. A network with small world property is one that shows dense local interconnections.

1.4 Application of Network Analysis

Network science has grown exponentially as a novel tool for the study of complex systems. Thanks to this approach, many unexpected phenomena of real world systems have been discovered, and increasingly sophisticated system structures are studied [16]. The fields of application vary among several subjects, for instance, in medicine, networks are used to assess the epidemic spreading [17] and the immunization strategy [18], in biology found wide application in the analysis of the operation and organization of the brain [19] and to decribe the protein networks [20] as well as in meteorology to study the factors that affect climate [21]. In communication science and information technology was used to model the structure of mobile communication network [22], of internet [23] and the opinion dynamics [24]. Many other fields exploited the strengths of the network analysis, varying across the most disparate subject, from the transportation [25], to the chemistry [26]. Network analysis ability to describe phenomena structure was widely exploited in the field of economics. An understanding of the growth, structure, dynamics, and functioning of the economic networks and their mutual interrelationships is essential in order to find precursors of changes, to make the systems resilient against failures, or protect them against external factors. In particular, after the financial crisis of 2008, the awareness about the importance of financial actors links increased the use of this tool, allowing the analysis of the contagion effect [27] [28]. The social network analysis was not used only in the financial industry, but over the last two decades has proliferated rapidly as a tool to assess the structure of several economic sectors. The strength of this analysis is that it allows to examine both supply and demand perspective, furthermore, for marketing purposes, it replaces the dominant view of the consumer as an individual with the more accurate model of the consumer as acting as part of a herd [29]. According to Scott et al. [30], the tourism sector is the network industry par excellence. Bjork and

Virtanen [31] defined tourism as a system where interdependence is essential and collaboration and cooperation create the tourism product [32]. In literature, the fragmented nature of the tourism industry has been widely recognized [33], [34]. Indeed the sector consists of small independent business with high staff turnover in a turbulent environment. The survivor of these actors depends mostly on collective action [35], which can be assessed only by means of a network approach. Network analysis can therefore deliver a number of useful outcomes for tourism studies. It provides a means of visualizing complex sets of relationships and simplifying them, and so can be useful in promoting effective collaboration within a destination group, supporting critical junctures in destination networks that cross functional, hierarchical, or geographic boundaries; and ensuring integration within groups following strategic destination restructuring initiatives [36].

Chapter 2

Structural Time Series

2.1 Introduction

In time series analysis two main approaches have been developed. The classical one was proposed by Box and Jenkins in 1970 [37] and it is based on ARIMA models. The second is called structural approach and the methodological base and main ideas beyond this method goes through the works of Harvey [38], West and Harrison [39], Kitagawa and Gersch [40]. The former deals with the non-stationary features removing them from the series by means of differentiation. Letting y_t denote the observed data, differenting involves the computation of a new variable y_t^* satisfying $y_t^* = y_t - y_{t-1}$ to remove the trend and $y_t^* = y_t - y_{t-s}$ to remove seasonal with periodicity s. Usually it is necessary a combined removal of trend and seasonality obtaining $y_t^* = y_t - y_{t-s} - (y_{t-1} - y_{t-s-1})$. In case y_t^* is still not stationary the differenting procedure can be continued by taking high integration order. Once the series achieves a sufficient stationarity an ARMA(p,q) model that best fit the observed data need to be identified. In contrast to this approach, Grether and Nerlove in their paper "Some properties of optimal seasonal adjustment" [41], treated the non-stationary elements in a different way. They intended the series as the summation of seasonal and non seasonal part. A turning point for this approach was achieved thanks to the works of Pagan [42] [43], who suggested the use of Kalman filter to estimate the components. This methodology was revised and improved by the Nobel prize Engle. In his paper "estimating structural models of seasonality" [44] sought to fit a model to each unobserved component, in particular the seasonal one. He found three main advantages in using this approach. First, with an explicit statistical model of the seasonal process, it is possible to calculate the properties of different methods and the variances of individual component estimates. Second, the method of seasonal adjustment will be tailored to the characteristics of the series, and third, it is possible to incor-

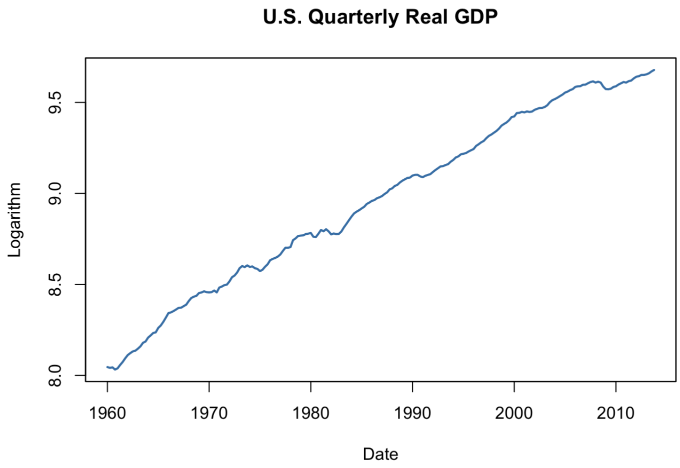


Figure 2.1: Quarterly Real GDP in United States for the period 1-Jan-1960 1-Jan-2012

porate additional information about economic trends and cycles, weather, strikes, and holidays which will help distinguish seasonal from nonseasonal behaviour and will provide means for automatically correcting for phenomena that are normally treated as outliers. Durbin and Koopman pointed out another pro of this method, focusing in the nature of the phenomena the econometrists analyze. In fact, in economic and social fields real series are never stationary, and the researcher who uses ARIMA model has to ask himself, how close to stationary is close enough? [45]. This hard question is not required if the researcher opts for the structural analysis. However, how outlined by Commandeur et al. [46], each ARIMA model can be put in state-space form and analyzed using structural method. Despite this commonality, the two approaches are actually the opposite. In Box-Jenkins, trend and seasonality are treated as nuisance parameters and removed from the series. In contrast the structural method emphasizes the unobserved components contained in the data. From a philosophical point of view, the main difference between them is on what the analyst is searching for, in the former case he is interested in the underlying generator process, while, in the latter, the focus is shifted on the stylized fact, that is the recursive empirical pattern that the series presents. The common salient characteristics of data behaviour identified by Harvey [38], are trend, cyclicity, seasonal movements, and daily effects. These facts can be found in thousands of series, and often, they have direct interpretation. As an example, is reported in figure 2.1 the logarithm evolution of the gross domestic product (GDP) in the United States for the period starting from 1960 up to 2012. It is evident a constant tendency to rise, in other words, an upward trend is

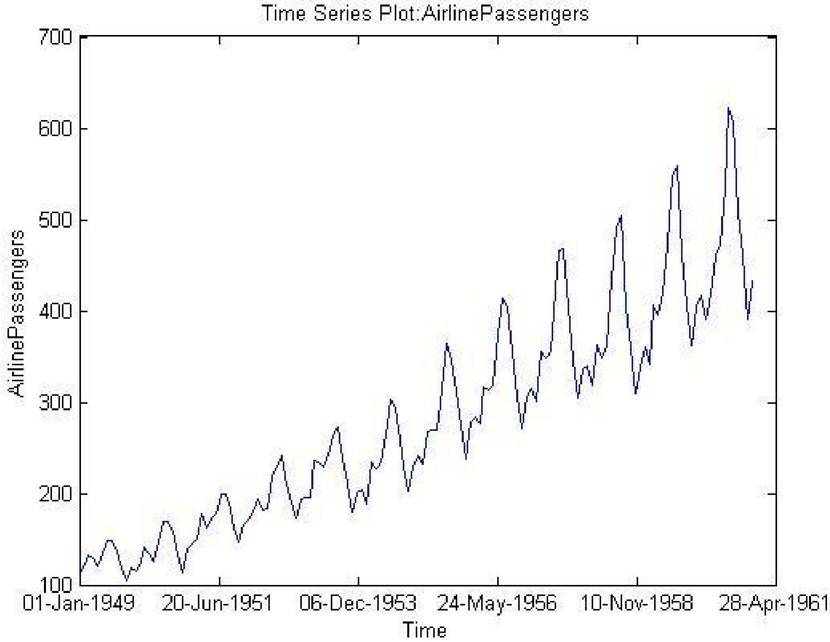


Figure 2.2: Monthly U.S. airline passenger for the period 1-Jan-1949 28-Apr-1961

present. There are other examples of data that easily allow us to appreciate other empirical recurrences. It is the case of the time series in figure 2.2, that represents the data of airline passengers from 1949 to 1961. Not only an upward trend can be found, but also an accentuated seasonality. Every year, there are months with high travelers and months with low level of them. The reason can be found in the differences between months that include some holiday periods, such as summer or Christmas time, and the other ones. How it is possible to appreciate looking at these two pictures, the data can be intended as the summation of the components previously stated plus an irregularity term, how showed in the equation below

$$Observation = trend + cycle + seasonal + daily + irregular \tag{2.1}$$

It is possible to implement a multiplicative form as well, which can be easily transformed in a summation taking the logarithm. One key aspect of the structural methodology is that the researcher can use these models not only to make prediction, but also to check if the components of the series are consistent with any prior knowledge about the phenomenon, sometimes only given by a purely theoretical explanation. This is a very useful tool to test the empirical evidences of economic models. An example is the work of Tawadros [47] who investigated the relevance of monetary models of exchange rate or the work of Beissinger [48] about the real wages and business cycle relationship. A thorough presentation of the main ideas and methodological aspects underlying structural time series models is contained in [38], [39] and [40]. In the next sections we are going to investigate the form that a time-series should take to asses its unobservable components, and we are going to talk about the state estimation system called Kalman filter.

2.2 State-Space Form

The Components in analysis are not directly observed, so we need to guess them. Before running such an estimate, data must be sort in a State-Space form. The observations are collected in a vector/matrix y, either of dimension $t \times 1$ if we are handling a univariate time-series or of dimension $t \times m$ in the case of multivariate time-series. Suppose we are in the former scenario, the equation that describes the behaviour of y is

$$y_t = Z_t \alpha_t + \epsilon_t \tag{2.2}$$

Where Z is a $t \times m$ matrix, α is an $m \times 1$ and epsilon is an n x 1 vector with the following characteristic

$$E(\epsilon_t) = 0 \quad Var(\epsilon_t) = H_t \tag{2.3}$$

Recalling the representation described in the previous section, ϵ_t can be intended as the uncertainty, while α, the so-called state vector, as the collection of the four unobservable components. To understand the state-vector representation, it is important to describe in depth the stylized facts, in order to find a form that can capture all the relevant features of them.

2.2.1 Trend

The trend represents the long-term movement, and here is denoted by μ. If we assume that the oberevd data y are given by a trend μ plus an idiosyncratic term, the representation will be

$$y_t = \mu_t + \epsilon_t \tag{2.4}$$

The trend can be expressed in various forms, the most common is the linear one

$$\mu_t = a + bt \tag{2.5}$$

that is

$$\mu_t = \mu_{t-1} + b, \quad t = 1, 2, ... \tag{2.6}$$

with $\mu_0 = a$. We allow the model to capture all the structural changes that the phenomenon may encounter, making the trend stochastic. This purpose can be reached adding an error term to the μ_t equation, and allowing the slope b to vary over time

$$\mu_t = \mu_{t-1} + b_t + \eta_t \tag{2.7}$$

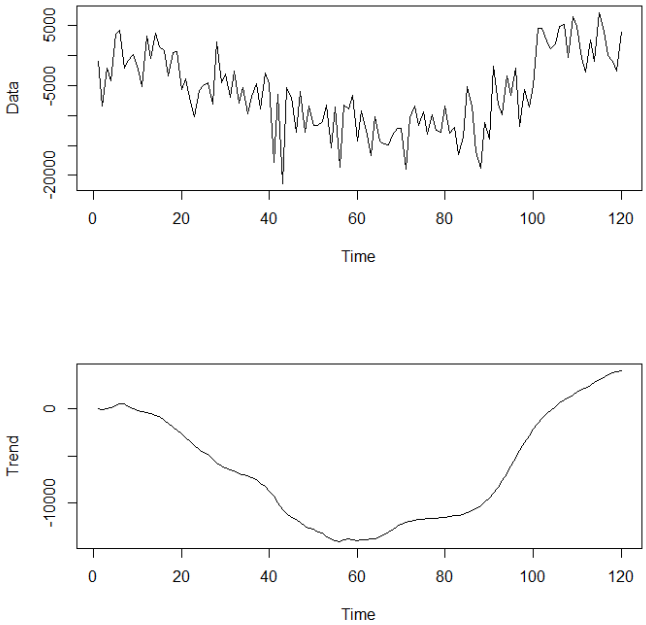


Figure 2.3: Time series simulation and trend decomposition

$$b_t = b_{t-1} + \zeta_t \tag{2.8}$$

Where η_t e ζ_t are two uncorrelated white-noise disturbances with $E(\eta_t) = 0$, $E(\zeta_t) = 0$ and $Var(\eta_t) = \sigma^2_\eta$, $Var(\zeta_t) = \sigma^2_\zeta$. The previous equations can be expressed in matrix form, in order to insert the relations concerning trend, in the state vector:

$$\begin{bmatrix} \mu_t \\ b_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix} \tag{2.9}$$

In figure 2.3 is run a simulation on R of a structural time series with trend. The first time series represents the observed data, and the second the underlying trend is reported. The parameters are: $\sigma_\epsilon = 4000,\ \sigma_\eta = 20,\ \sigma_\zeta = 100,\ \mu_0 = 0,\ b_0 = 3$

2.2.2 Cycle

The cycle is defined as the recurrent, though not exactly, deviation around the long term path of the series [49]. If cyclicaliy is the only component presents in the observations, they can be represented by

$$y_t = \psi_t + \epsilon_t \tag{2.10}$$

where ψ_t is a periodic function described by the following equation

$$\psi_t = a\cos(\lambda t) + b\sin(\lambda t) \tag{2.11}$$

with $\sqrt{(a^2+b^2)}$ amplitude, $tan^{-1}(\frac{b}{a})$ phase and λ the frequency. As we did for trend, we let the parameter a and b to be stochastic, and a recursion in the construction of ψ_t is made to ensure continuity.

$$\psi_t = \psi_{t-1}\cos(\lambda) + \psi^*_{t-1}\sin(\lambda) + \kappa_t \tag{2.12}$$

where ψ^* is described by

$$\psi^*_t = \psi_{t-1}(-\sin(\lambda)) + \psi^*_{t-1}\cos(\lambda) + \kappa^*_{t-1} \tag{2.13}$$

with $\psi^*_0 = b$, $\psi_0 = a$, κ and κ^* noise terms. In matrix form the previous relations becomes

$$\begin{bmatrix}\psi_t\\ \psi^*_{t-1}\end{bmatrix} = \begin{bmatrix}\cos\lambda & \sin\lambda\\ -\sin\lambda & \cos\lambda\end{bmatrix}\begin{bmatrix}\psi_{t-1}\\ \psi^*_{t-1}\end{bmatrix} + \begin{bmatrix}\kappa_t\\ \kappa^*_{t-1}\end{bmatrix} \tag{2.14}$$

Figure 2.4 is a simulation of a structural time series with trend and cycle, the parameters are $\sigma_\epsilon = 2000$, $\sigma_\eta = 10$, $\sigma_\zeta = 10$, $\mu_0 = 0$, $b_0 = 3$ $\lambda = 50$, $\psi_0 = 2000$, $\psi^*_0 = 2000$

2.2.3 Seasonality

The seasonality component can be intended as the systematic, regular, intra-year movement caused by the changes of the weather, the calendar and timing of decision [50]. The difference between cycle and seasonality is that the first is caused by the up and down movements related to a non fixed period, while the second consists in finding the repetitive behaviour due to something that will occur pretty much surely, because of natural phenomenon, such as hot temperatures in summer, or the snow in winter. If a series is influenced only by seasonal factors, each observation can be represented as the sum of a seasonal component γ plus an irregularity term

$$y_t = \gamma_t + \epsilon_t \tag{2.15}$$

The length of the season will be denoted with s, so the seasonal components will be $\gamma_1, \gamma_2 ... \gamma_s$. All the components must sum to zero in order to ensure that the seasonal movements does not lead to a permanent shift in the series

$$\sum_{j=1}^{s} \gamma_j = 0 \tag{2.16}$$

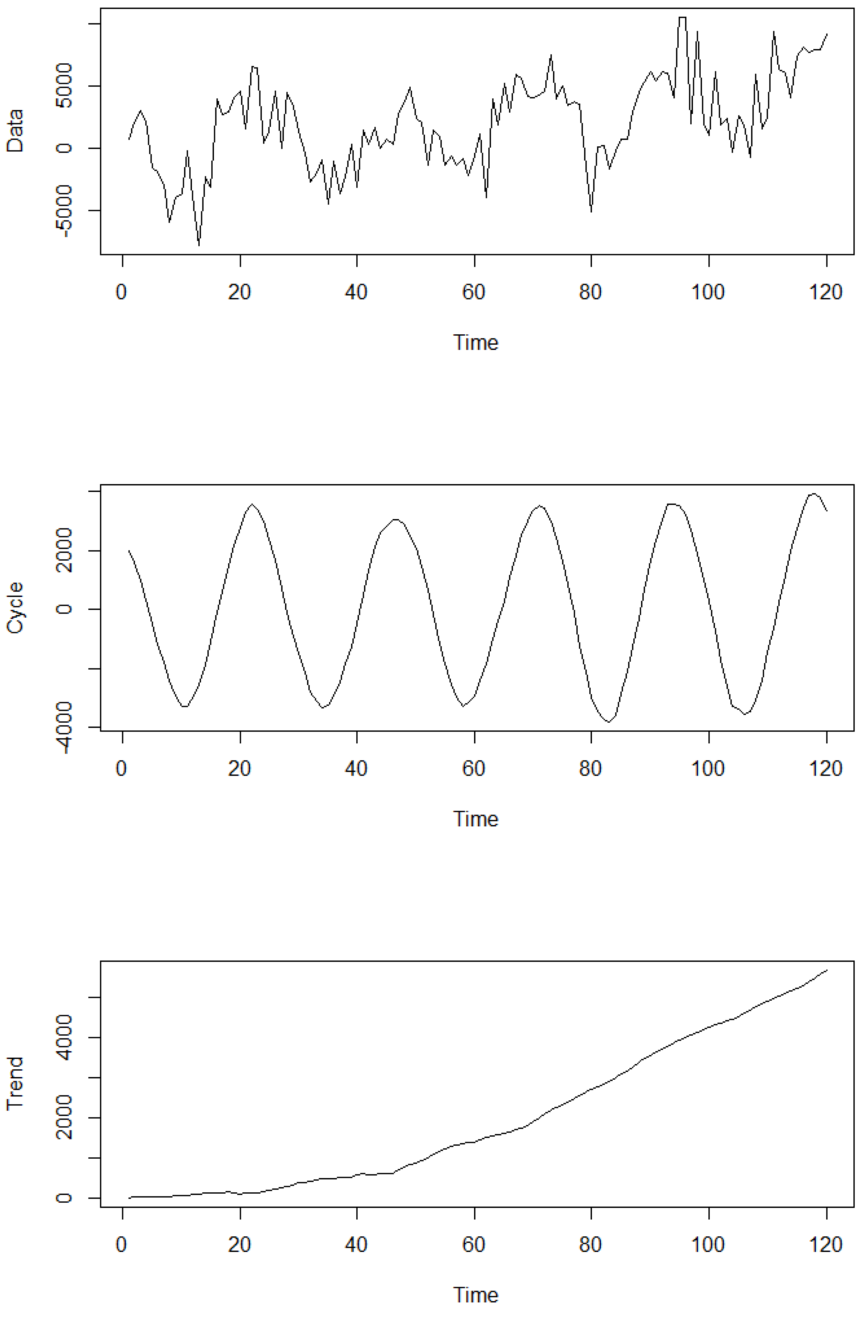


Figure 2.4: Simulated time-series and decomposition in trend and cycle components

Since $\gamma_1, \gamma_2, ..., \gamma_s$ sum to zero, the last seasonal component γ_s, must be equal to the negative sum of the previous $s-1$ terms. At a generic time t, γ_t, is equal to

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} \tag{2.17}$$

Adding an error term ω with white-noise characteristics, the seasonal effect is allowed to change over time

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \tag{2.18}$$

The last step is to re-arrange the relationship using matrices in order to obtain a suitable form for the state-space representation, obtaining

$$\begin{bmatrix} \gamma_t \\ \gamma_{t-1} \\ \vdots \\ \gamma_{t-s+1} \end{bmatrix} = \begin{bmatrix} -1 & -1 & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \gamma_{t-1} \\ \gamma_{t-2} \\ \vdots \\ \gamma_{t-s} \end{bmatrix} + \begin{bmatrix} \omega_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.19}$$

A thoruough presentation of the other methods used by pratictioners to handle seasonal time-series can be found in Franses [51]. The simulation in figure 2.5 is a series with trend, cycle and seasonality as components.The parameters for trend, cycle and observation uncertainty are the same as before. For the seasonal component we assume a season of length 12, so we consider each observation as a month. there is a peak on the sixth and seventh month of each year that is absorbed in the eigth and ninth ones. We can imagine a scenario where the hot temperatures in summer, in partiular in June and July, lead to an increase in the observed data.

2.2.4 Daily Effect

Daily effects are the systematic mouvement caused by the differnces among days, for instance between week ones and holydays. The Daily effect can be viewed as a generalization of seasonality and so it will be trated as it. For this reason in the state vector I will incorporate this component in the equation 2.19.

2.2.5 State Vector

The state vector assembles the four stylized facts presented in the precedent sections, that is, the trend, the cycle and all the seasonal behaviours, which include

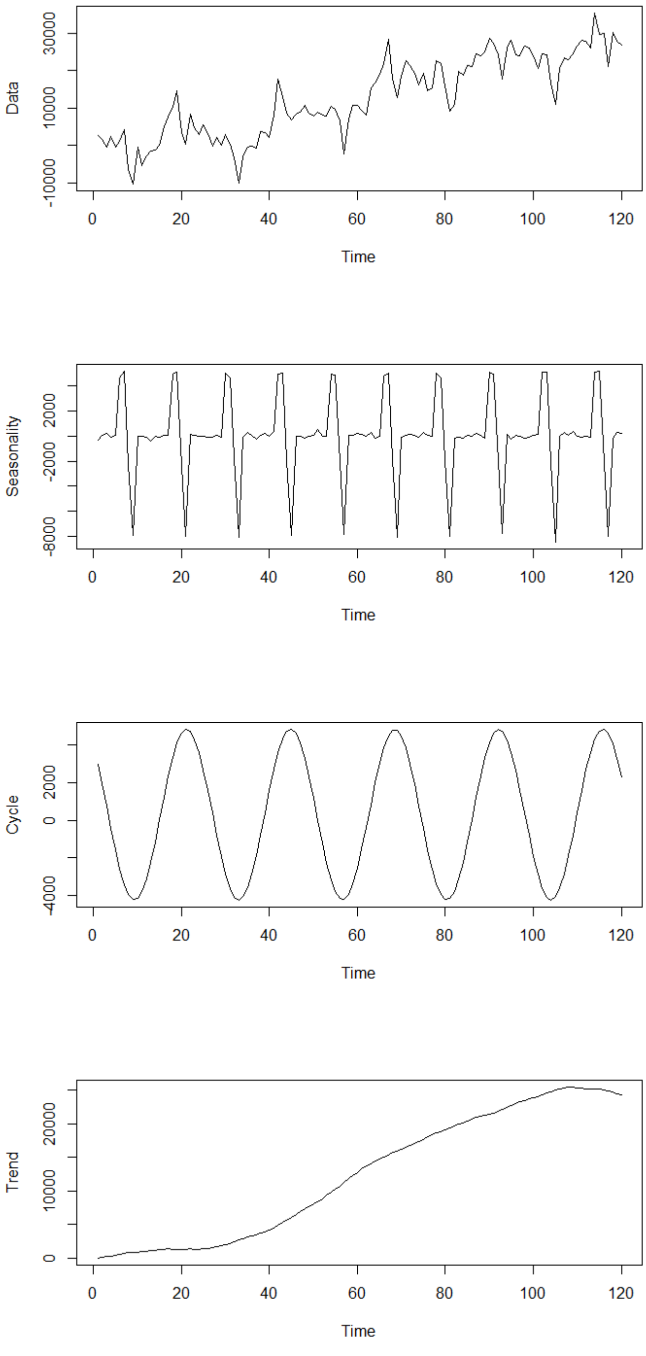


Figure 2.5: Simulated time-series and decomposition in trend cycle and seasonal components

seasonal and monthly recursive movements as well as daily ones (daily effect). It assumes the form

$$\alpha_t = \begin{bmatrix}\mu_t & \psi_t & \gamma_t\end{bmatrix}' \tag{2.20}$$

where μ_t is the trend, ψ_t the cycle and γ_t the seasonality plus the daily effect. These are the components that have direct interpretation, however, we have to add in this vector all the other variables that contribute to create them, that is b for the trend, ψ_t^* for the cycle and $\gamma_{t-1}, \gamma_{t-2}, ...\gamma_{t-s}$ for the seasonality, as shown in the equation below

$$\alpha_t = \begin{bmatrix}\mu_t & b_t & \psi_t & \psi_t^* & \gamma_t & \gamma_{t-1} & ... & \gamma_{t-s}\end{bmatrix}' \tag{2.21}$$

We can see from equations 2.9, 2.14, 2.19, that each component of the state at a given time t is equal to the value of itself at the previous time multiplied by a matrix of constant, plus an irregularity term. This relation can be weel resumed by the equation below

$$\alpha_t = T\alpha_{t-1} + \upsilon_t \tag{2.22}$$

where T is the matrix of constants, and υ_t the vector containing the uncertainty terms, with

$$E(\upsilon_t) = 0 \qquad Var(\upsilon_t) = Q_t \tag{2.23}$$

Determining the components as we did in the previous sections, the extended form of 2.22, will be

$$\begin{bmatrix}\mu_t\\ b_t\\ \psi_t\\ \psi_t^*\\ \gamma_t\\ \gamma_{t-1}\\ \vdots\\ \gamma_{t-s+1}\end{bmatrix} = \begin{bmatrix}1 & 1 & 0 & 0 & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0\\ 0 & 0 & \cos\lambda & \sin\lambda & 0 & 0 & \cdots & 0\\ 0 & 0 & \text{-}\sin\lambda & \cos\lambda & 0 & 0 & \cdots & 0\\ 0 & 0 & 0 & 0 & -1 & -1 & \cdots & -1\\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1\end{bmatrix} \begin{bmatrix}\mu_{t-1}\\ b_{t-1}\\ \psi_t\\ \psi_{t-1}^*\\ \gamma_{t-1}\\ \vdots\\ \vdots\\ \gamma_{t-s}\end{bmatrix} + \begin{bmatrix}\eta_t\\ \zeta_t\\ \kappa_t\\ \kappa_t^*\\ \omega_t\\ 0\\ \vdots\\ 0\end{bmatrix}$$

2.3 State estimation

The state vector is not directly visible, the only thing we can observe is an output related with the the state. The hidden variables of this type are known as latents, and they must be inferred through methodologies such as filtering and smoothing. The relating estimators are the filter and the smoother respectively. How reported by Einicke [52] The basics on this topic were developed by Norbert Wiener [53], Richard S. Bucy [54][55] and Rudolf E. Kalman [56] over half a century ago. In this thesis we will focus on the work of the latter one, using its recursive method known as Kalman filter. Before starting to explain how to implement the state estimation, it is useful to fix the differences between prediction, filtering and smoothing.

SMOOTHING:
Smoothing aims to find the distribution of a past hidden variable. In this techniques, we use informations on what happened before and after the moment of interest. Let suppose we have three time instant, $t < s < T$, and let denote with α the latent variable and with y the observations. The smoother attempt to find the distribution of α_s using the data from t to T, that is

$$Pr(\alpha_s|y_1, ...y_s, ...y_T) \tag{2.24}$$

FILTERING:
In contrast with smoother, a filter estimates the current value of a latent variable. The data available, are the observations from t up to the moment of interest. Using mathematical notation a filter can be intended as

$$Pr(\alpha_T|y_1, ...y_s, ...y_T) \tag{2.25}$$

In this case we are trying to estimate the present.

PREDICTION:
Prediction probably represents the most common practice in statistics, or at least, the most common way to think at this subject. It aims in using past data to forecast the future. The notation is

$$Pr(\alpha_T|y_1, ...y_s) \tag{2.26}$$

In this chapter we are interested in filter and smoother. Usually, an analytical solutions for these two estimators become impossible to find, for this reason a lot of numerical methods were developed. The algorithms for filtering and smoothing are different in nature, since the former proceed forward, whereas the latter backward. The next sections present one of the most known versions of the two estimators.

2.3.1 Kalman filter

The state estimation method known as Kalman filter was introduced in Kalman, 1960 [56]. It consists of a recursive procedure with the aim of computing an

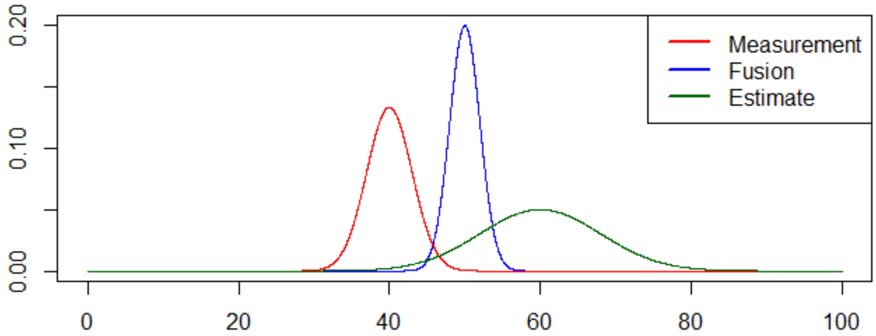


Figure 2.6: Kalman filter operation in graphic terms

optimal estimator of the state vector distribution at a given time t. Optimal estimator means that it is the one with the smallest mean squared error. The main idea behind the Kalman filter is to combine two unbiased estimates of a variable in order to obtain a better one as showed in Pei, 2019 [57]. So far, we have seen two equations from which we can infer something about the state vector, they correspond to 2.2, 2.22. They are reported below

$$y_t = Z_t \alpha_t + \epsilon_t$$

$$\alpha_t = T\alpha_{t-1} + \upsilon_t$$

The former equation relates an observable variable with a latent one. Acting as a detector, it gives an estimate of the state starting from a measurement of y. For this reason it is called measurement equation. The latter is a description of the state vector behaviour with respect to its past values. it gives us an estimation of how the state will move in the future, and it is known as estimation equation. In figure 2.6 is shown a graphical representation of how the Kalman filter combines two estimators in order to obtain a better one. Suppose the red and green represent the measurement and the estimation distributions, merging them, we obtain the blue one, which has a smaller variance, increasing the precision. The graphic procedure must be translate in equation, at each time, thanks to the estimation equation, we can obtain an a priori estimate, denoted by a^-, for the future value of α.

$$a_t^- = Ta_{t-1} \tag{2.27}$$

and P_t^+ for its variance-covariance

$$P_t^- = TP_{t-1}T' + Qt \tag{2.28}$$

Then we combine them with the successive measurement of y and its relative variance-covariance in order to obtain an a posteriori estimate, denoted by a_t^+ and P_t^+. This process is fixed through the following equations

$$a_t^+ = a_t^- + K_t(y_t - Z_t a_t^-) \tag{2.29}$$

$$P_t^+ = (I - K_t Z_t)P_t^- \tag{2.30}$$

The term K_t is the so called Kalman gain, it multiplies the difference between y_t and the a priori estimate multiplied by the matrix Z. The closer to one is the Kalman gain, the nearer will be the a posteriori estimates to y, while, in the opposite case, it will approach to the a priori one.

2.3.2 Kalman Gain

The proper value of the Kalman gain at time t is given by the one that minimizes the Mean Squared Error (MSE) defined as

$$P_t = E[(\alpha_t - a_t^+)(\alpha_t - a_t^+)^{'}] \tag{2.31}$$

Substituting 2.2 in 2.29, and the resulting equation in 2.31 we obtain

$$P_t = E[[(I - K_t Z_t)(\alpha_t - a_t^-) - K_t\epsilon_t][(I - K_t Z_t)(\alpha_t - a_t^-) - K_t\epsilon_t]^{'}] \tag{2.32}$$

The term $\alpha_t - a_t^-$ is the error of the a priori estimate, which is clerly uncorrelated with the measurement noise, therefore the multiplications between the expected values of them is equal to zero. Denoting the MSE of the a priori estimate P^- and using 2.3, we obtain

$$P_t = (I - K_t Z_t)P_t^-(I - K_t Z_t)^{'} + K_t H_t K_t^{'} \tag{2.33}$$

The minimization of the MSE can be obtained by the minimization of the trace of P, which can be reached computing its partial derivative with respect to K and setting it equal to zero

$$\frac{dT[P_t]}{dK_t} = -2(Z_t P_t^-)^{'} + 2K_t(Z_t P_t^- Z_t + H_t) = 0 \tag{2.34}$$

$$(Z_t P_t^-)^{'} = K_t(Z_t P^- Z_t + H_t) \tag{2.35}$$

Solving for K we obtain the equation that determine the value of the Kalman gain at time t

$$K_t = P_t^- Z_t^{'}(Z_t P^- Z_t + H_t)^{-1} \tag{2.36}$$

A detailed presentation of the Kalman gain demonstration is provided by Kim [58]. The logic behind the Kalman gain is straightforward to understand. Remembering that H_t represents the uncertainty in the observation while P_t^- the one in the a priori estimate, it is possible to observe that the greater is the error in the estimate the closer is K to one, and, as a consequence, the a posteriori estimates to y. In the opposite case, when the measurement error is large, the gain approaches to 0, giving more importance to the a priori estimate in the computation of the a posteriori one.

2.3.3 The Algorithm

In this paragraph, we are going to see the steps of the Kalman filter methodology. As sated before, it is a recursive procedure, so the operations must be repeated t times, where t represents the moment of interest. We can split the procedures in 4 steps:

STEP 1:
Compute the a priori estimate using equations 2.27 and 2.28. where α_{t-1}^+ is the latest estimate available and P_{t-1}^+ its relative variance-covariance matrix. Usually we obtain them from the a posteriori estimates computed in the past steps, but if we are at the first iteration, this data is not yet present, so we have to use an estimate of the vector that arises from our previous knowledge about the series. For this reason, before starting the Kalman filter procedure, we need to know the expected value and the variance of a_0.

STEP 2:
Compute the Kalman gain using equation 2.36 where P_{t-1}^- is the magnitude computed in the past step

STEP 3:
Once the measurement y is available, the a priori estimate can be updated through equation 2.29 and 2.30, obtaining the a posteriori one. It is easy to see that the error of the new estimate can not ever be greater of the previous one. The smaller is the error in measurement, the larger is the Kalman gain, which implies a steeply decrease in the updated matrix P. In this scenario, the state estimation quickly approaches to its actual value. In the reverse case, the opposite conclusions are true.

STEP 4: Repeat the process using the update estimates and error. This passage is iterated up to the time we are interested in knowing the values of the state vector.

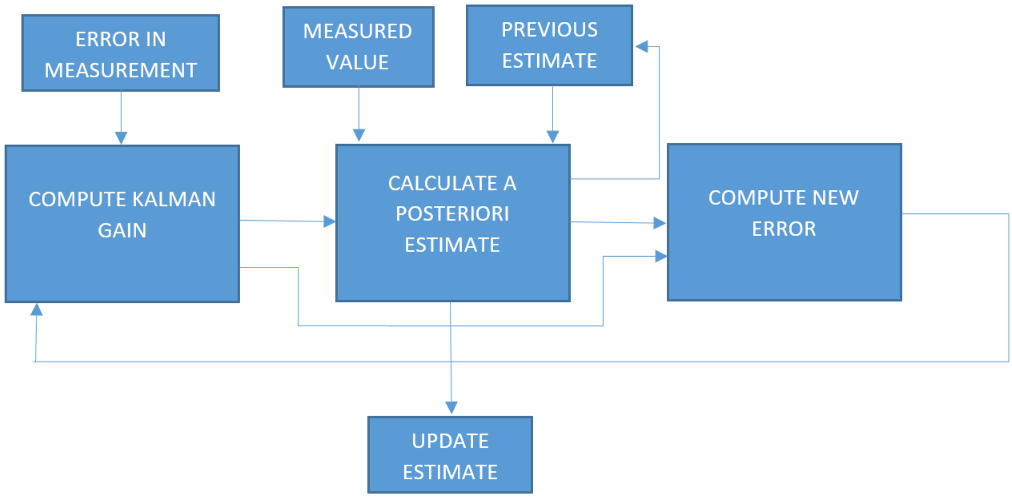


Figure 2.7: Flow chart of Kalman filter algorithm

In figure 2.7, is presented a flow chart which resumes how the Kalman filter works

2.3.4 Kalman Smoother

So far, we have seen how to estimate the present state distribution, now we assume that we are interested in the past one. The Kalman smoother can be computed using three categories of algorithm, fixed point, fixed lag or fixed interval algorithms. The first one is concerned in estimates the state vector at some fixed point in time, for instance, assuming a time interval that starts from 0 up to T, and with $0 < t < T$, an algorithm of this type estimates the latent variable value at time t only. The second typology computes an estimation for a fixed delay, that is the value of the state between T and $T - j$, for $j = 1, .., M$, where M is the maximum lag. With the last one category, we can figure out all the full set of smoothed estimates for a fixed span of data. We will use the latter one, in particular an algorithm developed in Rauch et al.[59], from now denoted by R-T-S algorithm.

2.3.5 R-T-S Algorithm

R-T-S algorithm is an optimal fixed-interval smoothing algorithm, which, in general, can reach higher accuracy with respect to the Kalman filter [60]. The R-T-S uses both a forward (the Kalman filter) and a backward algorithm[61]. The first step is to estimates the state vector $\alpha_{0:T}$ from time 0 up to time T (suppose $0 < t < T$) using the Kalman filter. Thanks to this passage we stored T a priori and a posteriori estimates, and T a priori and a posteriori variance-covariance matrices. Each estimate will be denoted from now with the superscript f to indicate that it is derived from the forward algorithm.

$$a_t^{f+},\ P_t^{f+} \tag{2.37}$$

$$a_t^{f-},\ P_t^{f-} \tag{2.38}$$

Instead the estimate derived from the backward algorithm will be denoted by superscript b.

$$a_t^{b},\ P_t^{b} \tag{2.39}$$

The first backward estimate a_T^b correspond with the last a posteriori estimate a_T^{f+}, the same thing goes for the variance matrix, $P_T^b = P_T^{f+}$. The backward value at the previous step is computed merging the forward a posteriori one at time $T-1$ and the difference between the backward and the forward a priori estimate at time T, by means of the following equation

$$a_t^b = a_t^{f+} + K_t^b(a_{t+1}^b - a_{t+1}^{f-}) \tag{2.40}$$

where K_t^b assumes the same meaning of the Kalman gain, and it is computed in the following way

$$K_t^b = P_t^{f+}T'(P_{t+1}^{f-})^{-1} \tag{2.41}$$

While the variance of the estimate is

$$P_t^b = P_t^{f+} + K_t^b(P_{t+1}^b - P_{t+1}^{f-})K_t^{b\,\prime} \tag{2.42}$$

As for the Kalman filter all these passages must be repeated T times, with the difference that in the previous scenario we start from 0 up to arrive at T, now we begin from T up to time 0.

Chapter 3

Review on Tourism Flows Forecasting Methods

3.1 Introduction

In the last few decades, the tourism industry showed a spectacular growth [62]. The United Nations World Tourism Organization (UNWTO) reported an increase of international tourism arrivals between 1950 and 2008 from 25 to 922 million that generated more than $944 billion in receipts and averaging a 4 percent growth per year. Greenwood [63] defined this sector as the largest scale movement of goods, services and people that humanity has perhaps ever seen. The literature gave several theoretical explanations to this phenomenon, that can be divided in demand and supply side. Theories in the first one category, find the justification for the expansion in an increasing desire during the post-World War II period, for touristic experience, fuelled by a quest either for authentic experience lacking in everyday life [64], an inversion of one's normal routine [65] or a chance to gaze upon exotic sights [66]. For other researchers, the reasons behind the phenomenon must be found in the supply side, in particular in the changes that the labour market faced after the World-War II. Honey [67] stated that paid vacation time, shorter hours of work, less physically taxing jobs, and better education pushed workers to spend their leisure travelling. Not only improvements in labour markets, but also the desire of industry operators to develop profitable enterprises can have pushed the demand for touristic goods, Fundamental in this respect was the development of the commercial airline industry in the 1950s [67]. Tourism dynamics is not something that only interests industry agents or travellers, indeed it has become a strategic sector or even vital for some less-developed countries. Munt [68] described tourism as a last-ditch attempt to break from the confines of underdevelopment and get the IMF to lay the golden egg of an upwardly-mobile GNP. Indeed, since the 1960s tourism has been embraced as a development strategy by a great number of development planners, including transnational institutions such as the United Nations and World Bank, international aid agencies such as USAID, and national governments worldwide [62]. Although for developed countries it is not an indispensable industry, tourism remains a strategic business for the economy. The World Tourism Travel Council (WTTC) estimated that it counts for the 11% of the global GDP. Furthermore, it creates noumerous employment opportunities, just think about the industries associated with tourism: transportation services such as cruise ships and taxis, accommodation such as hotels, restaurants, bars, and entertainment venues, and other hospitality industry services such as spas and resorts. An interesting work of Lee et al. [69] give more reason to focus on tourism industry expansion. They showed how there are long-run comovements and causal relationships between tourism development and economic growth for OECD and nonOECD countries using data for the 1990–2002 period. In the light of the aforementioned motivations, appears to be clear the importance of a good prediction regarding the future touristic flows. Furthermore, for the specific features of the main touristic goods, an accurate forecasting becomes even more crucial respect to the other industries. Tourism largely involves intangible experiences that customers cannot keep, except in photos, videos and memories. More importantly, unlike industries such as construction, manufacturing, or retail trade, tourism products and services cannot be stockpiled, for example unsold hotel rooms, airline seats or bungee jumps on one day cannot be treated as inventory for sale at a later date. Therefore, planning and forecasting are vital more for tourism that for other businesses [70]. Predicting human mobility can be useful even in other fields [71]. It plays a key role in urban planning [72], traffic management [73] and in the very topical issues of infectious disease [74] [75] and epidemic control [76]. In this chapter, I am going to do a little review of some key papers published between 1968 and 2018. According to Jiao et al. [77], I divided the works into three categories; econometric, time series and artificial intelligence (AI) models. the first two have already been widely used in the past, are more often used as benchmark models for forecasting performance evaluation and comparison with respect to new models. AI models are rapidly developed in the past decade and hybrid AI models are becoming a new trend.

3.2 Econometric Models

The econometric models concentrate their analysis on relationships among variables. The aim is to establish a causality structure between economic-social factors and the demand for services and goods regarding tourism. The econometric model starts from the links given by the economic theory, then it tests whether an empirical feedback is given and determines how significantly the explanatory variables affect future demand. These kind of models have been widely used for people flows prediction purposes and have played a distinctive role in tourism demand forecasting research and practice over the past five decades [78].

3.2.1 Static Regression Model

The static regression is used in the most basic econometric works. The adoption of this kind of model aims to find the influence of some factors in causing the present value of the forecasted demand. The goal is to find a linear relation as the one described in the equation below

$$Y_t = \beta_0 + \beta_1 X1_t + \beta_2 X2_t + ... + \beta_n Xn_t + \epsilon_t \tag{3.1}$$

where Y_t represents the tourism demand level, $X1_t, X2_t, ..., Xn_t$ are the variables that influence demand and ϵ is a random error. In this category usually fall the earlier studies, in particular in '60s and '70s, now they are used as benchmark for more sophisticated methods [79]. As an example I reported the works of Laber [80] and Martin et al. [81]. The former, implementing a linear regression, found that the the portion of population that visit Canada from USA, the average expenditure per day, the average length of stay and travel receipts per population are well explained by the per capita income, the number of people that were born in the arrival location, the average distance between the greatest urban centre and a dummy variable that assumes value 1 if the state borders with Canada and value 0 otherwise. The results were very good, the R^2 values ranged from 0.67 up to 0.95 for the various regressions, showing a pretty much satisfactory result. The second paper estimates the demand using the cost of travel to the destination and the cost of living for tourists in the destination.

3.2.2 DL and ADL Model

One drawback of static regression is that it does not take into account the intertemporal relationships between demand for tourism and the factor that influence it. For this reason several researchers prefer the use of distributed lag or autoregressive distributed lag models. The application of the first is actually very limited due to the more general and advanced counterpart (ADL). The DL is usually used as benchmark to measure the performance of other techniques. The ADL model consists in writing the demand as a function of its past values and the present and past values of some explanatory variables. Calling Y_t the tourism demand at time t, and assuming the presence of only one explanatory variable X, the ADL(1,1) model takes the form of:

$$Y_t = \mu + \alpha Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \epsilon_t \tag{3.2}$$

The model allows us to find a steady-state solution which should be the value that the varibles assume in the long term. It is calculated through the following formula

$$Y = \frac{\mu}{1-\alpha} + \frac{\beta_1 + \beta_2}{1-\alpha} X \tag{3.3}$$

Taking the expected values it is possible to obtain a relation that describes the influence on Y_t of a permanent shift on X_t, which is equal to

$$\frac{\beta_1 + \beta_2}{1-\alpha} \tag{3.4}$$

The possibility to analyze in-depth the relation among variables has made the ADL model very used in tourism demand forecasting, as an example I report three works. Ongan et al. [82], undertook an innovative research which relates tourism and the economic policy uncertainty index (EPU). They found that one standard deviation increase in EPU leads to 4.7% decrease in Japanese tourist arrival. Lin et al. [83] used the ADL to analyze the factors that push the Chinese outbound flows finding that the income level and the cost of a stay at a tourism destination compared with that of staying at a Chinese tourism destination are two important factors that affect Chinese residents' traveling abroad. Other noteworthy work is the one of Smeral [84] that investigated the existence of asymmetric income and price effects on tourism demand across business cycles.

3.2.3 ECM

Using the ADL model we have included lags of variables to capture the short-term relationship between the demand and magnitudes of interest. The error-correction model is useful to understand the long-run link. It includes a term for the deviation from the long-term relationship that estimates how much of the disequilibrium will dissipate in the next forecasting period. Before implementing such a model it is necessary to investigate whether the variables in analysis are cointegrated. This happens when two time series have a long-run common movements. This can be carried out regressing the explanatory variables on the interest one and performing an augmented Dickey-Fuller test to check the stationarity of residuals. In affirmative case an EC Model can fit well the data. Subtracting on both sides y_{t-1} and adding and subtracting β_1 on right side of 3.2 we obtain

$$\Delta Y_t = \mu + (\alpha - 1)Y_{t-1} + \beta_1 \Delta X_t + (\beta_1 + \beta_2)X_{t-1} + \epsilon_t \tag{3.5}$$

or alternatively

$$\Delta Y_t = \beta_1 \Delta X_t - (1-\alpha)\Big(Y_{t-1} - \frac{\mu}{1-\alpha} - \frac{\beta_1 + \beta_2}{1-\alpha} X_{t-1}\Big) + \epsilon_t \tag{3.6}$$

Calling $\frac{\mu}{1-\alpha}$ a and $\frac{\beta_1+\beta_2}{1-\alpha}$ b we can observe that the attractor Y_t^* is equal to $a + bX_t$ and ΔY_t reacts to deviation on long run solution

$$Y_t - Y_t^* = Y_t - a - bX_t \tag{3.7}$$

In tourism demand forecasting the error correction model is widely used, of particular interest is the work of Goh [85] who found a relation between climate variable and tourist flows, suggesting to use such a magnitude as explanatory variable in order to improve the prediction model performance. In general the ECM is used to analyse the variables that has an impact on the inbound-outbound tourist flows in the long term such as the economic and financial crisis [86] [87] and difference in income elasticities [88].

3.2.4 Panel Data Regression Models

Another type of analysis used in tourism demand studies is the panel data regression. This regression incorporates information regarding both the intertemporal movements and the cross-sectional heterogeneity of the data. So far the use of PDR for prediction purposes is pretty much rare, the only noteworthy work is the one of Wen et al. [89]

3.2.5 Other on Econometric Models

In general the autoregressive distributed lag and the error correction models work very well. The paper of Song et al. [78] pointed out that among the 26 works that compared the ADL with other methods, 16 find it to be the best performing, while in 24 that tests ECM, 17 indicate it as the most recommended. Due to their flexibility forms, these two models can be used with other features that represent parameters assumptions or data utilisation. For instance time-varying parameters were used in the works of Li et al. [90] and Song et al. [91]. Such an integration allows to capture the gradual structural changes. To deal with the problem of mixed frequency data Bangwayo [92] integrated in a reduced form of ADL model the mixed data sampling (MIDAS) in order to estimates tourist arrivals in Caribbean. Among all this plethora of econometric models the most important and significative determinants of tourist demand have found to be tourists income level, exchange rate, prices of tourist goods and services in the destination relative to those in the origine and the price of tourist products in competing destinations. Others factors such as climate change [93], political stability [94], one-off event [95], terrorist attacks [96] and financial crises [97] are also found to have important effects on tourist demands. Although without direct causal relationships with the demand variables, search engine data from sources such as the Google

[98] and Baidu [99] indices have proved to be good indicators of tourism demand fluctuations.

3.3 Time-Series Models

The time-series approach estimates the future value of tourism demand based on its past value. Unlike econometric methods time series analysis uses successive observations taken at regularly time interval (e.g. daily, monthly or yearly) of the same magnitude. The strengths of this method are the capability to capture trend, cycle and seasonal patterns as well as the simplicity of data collection, indeed, contrary to the econometric approach, the measurement of only a variable is needed. According to Peng et al. [100] I divide this model class into basic and advance techniques. The former classification includes Naïve, auto-regressive (AR), exponential smoothing (ES), moving average (MA) and historical average (HA) models, while in the latter falls the ones that includes additional time-series features such as trend seasonality, like ARIMA-based, advanced ES and structural models.

3.3.1 Basic Techniques

Naïve 1 The Naïve 1 method consists in setting the future value of the variable we are studying equal to the last available observation. Letting T the last observation, the h-step ahead forecast value is equal to

$$Y_{T+h|T} = Y_T \tag{3.8}$$

This methodology works well only if data follows a random-walk. There are some forecasting fields (e.g. stock price) where this method reach quite good result, but a cause of the marked seasonality of tourist demand this is not the case. It is usually used as benchmark to test the performances of other techniques [101] [102].

Naïve2 This model attempts to improve the drawbacks of Naïve 1, taking into account seasonality and trend. The equation 3.8 is modified as follow

$$Y_{T+h|T} = Y_T + (Y_T - Y_{T-1}) \tag{3.9}$$

for trend, and

$$Y_{T+h|T} = Y_T + (Y_T - Y_{T-s}) \tag{3.10}$$

for seasonality, where s is the seasonal lag. As for the previous method, it is used for comparing purposes only [103] [104]. Despite their simplicity the naïve models gave reasonably results in particular for short-term horizon forecast. Among the basic techniques they are the most used thanks to their easy implementation.

AR The auto-regressive of order p model consists in regressing the demand level on its p past observations. Doing so we obtain the prediction as a function of the past values. The one-step-ahead forecasting is computed through the following equation

$$Y_{T+1|T} = \mu + \phi_1 Y_{T-1} + \phi_2 Y_{T-2} + ... + \phi_p Y_{T-p} \tag{3.11}$$

Between the works that uses this model I report the one of Fildes et al. [105] who uses AR together with other estimation techniques to forecast the air passenger traffic flows.

MA The moving average method of order q computes the predicted value with a linear combination of the q past values of a stochastic and perfectly unpredictable term. The equation for one step ahead forecasting is

$$Y_{T+1|T} = \mu + \theta_1 \epsilon_{T-1} + \theta_2 \epsilon_{T-2} + ... + \theta_q \epsilon_{T-q} \tag{3.12}$$

Kim [106] analyse the inbound tourism flow in south Korea using a moving average approach together with other models. The MA is usually combined with the previous model in order to create a more sophisticated technique. Indeed, it is very rare the use of AR and MA alone, often they only serve as a benchmark.

ES The exponential smoothing technique is strictly related to the moving average model. The difference is that the former assign exponentially decreasing weights to the past error terms. In particular the simple ES equation is

$$Y_{T+1|T} = \alpha Y_T + (1 - \alpha)\hat{Y_T} \tag{3.13}$$

where $\hat{Y_T}$ indicates the previous one-ahead forecast. Geurts et al. [107] used this model to predict the Hawaian tourism flow, finding that it works equally well as the Box-Jenkins approach. The ES is often used in its advanced forms that remedy to the flaws of simple version integrating both trend and seasonal terms.

Due for their simplicity the basic techniques are adopted to check the performances of advanced model. The use for this task is very frequent, Song et al. [78] found that in the 101 papers that they analyzed regarding time-series approach, 55 use the basic techniques as benchmark.

3.3.2 Advanced Techniques

ARMA based The auto-regressive moving average model combines the AR and MA models. The most common is the ARIMA model, which deals with trend differentiating the series until stationarity is reached. Once stationarity is achieved an ARMA model is fitted. Considering an ARIMA(1,1,1) model the one-step-ahead forecast is made by means of the following equation

$$Y_{T+1|T} = Y_{T-1} + \phi(Y_T - Y_{T-1}) + \theta\epsilon_{T-1} + \epsilon_T \tag{3.14}$$

In tourism literature this kind of model was very used, for instance Chan [108] used it for the tourism arrivals in Singapore and Dharmaratne [109] for the long-stay journey in Barbados. The ARIMA model is widely used in some modified form. For instance the SARIMA model, which differentiate also for the seasonal lag, it was used by Lim [110] to study the arrivals from Hong Kong and Singapore to Australia, and by Millan [111] for the analysis of the oleotourism in the south of Spain. ARFIMA models are also used for the series which exhibiting long memory, such a series can not be considered neither stationary nor integrated of order one. The ARFIMA model allows for non integer order of integration in order to capture both the mean-reversion and the covariance non-stationarity. It was used by Chu [112], who showed that to predict the tourist inbound in Singapore it is the best among the time-series model, in that it exhibits the lowest mean absolute percentage error (MAPE). Other method to deal with long memory series is to fit an ARARMA model as did by Chu [113]. To take into account external events, such as policy changes, strikes, advertising promotions, natural disasters, epidemic etc. Goh and Law [114] used an ARIMA-In, which allows the interventions of some independent variable on a dependent variable of interest. To model the volatility as well, Chan et al. [115] used for the first time in tourism prediction literature an ARIMA-GARCH model to predict the conditional volatility of the tourism arrival in Australia from Japan, New Zealand, UK and USA, finding the presence of interdependent effects in the conditional variances of these four countries, and asymmetric effects of shocks in two of them. A detailed collection of the ARMA based models used in tourism forecast can be found in Song et al. [78].

Advanced ES Advanced exponential smoothing methods search to remedy to the flaws of simple ES models. It integrates trend adding an additional smoothing factor to control the decay of the influence of the change in trend. The method supports trends that change in different ways: an additive and a multiplicative, depending on whether the trend is linear or exponential respectively. To deal with seasonality, a third factor is added. As for the trend, the seasonality may be modeled as either an additive or multiplicative process for a linear or exponential change respectively. This method is sometimes called Holt-Winters Exponential

Smoothing, named for two contributors to the method: Charles Holt [116] and Peter Winters [117]. It was used in touristic field, with very satisfactory performances by Weatherford and Kimes [118] to forecast the hotel revenues as well as by Lim to predict quarterly tourist arrivals to Australia from Hong Kong, Malaysia, and Singapore. Many other works used Holt-Winter models type, a list of it can be found in Song et al. [78].

Strctural Time Series A detailed explanation of the structural time series methodology is presented in the previous chapter. In touristic field I reported the work of Turner et al. [119], who analyze inbound tourism to New Zealand from Australia, Japan, UK and USA, the work of Guizzardi and Mazzocchi [120] who create an innovative insight on modelling the effects of the business cycle on domestic and inbound tourism demand and the work of Clewer et al. [121] who introduced Intervention variables to take account of sudden shocks to tourism demand, and demonstrates that Structural model generally provided more accurate forecasts than Box-Jenkins models.

In the tourism forecasting methods review written by Song et al. [78], they found that among the 101 papers that opts for time-series methods, 74 use ARMA based models, and in 56 of them, such models are identified as the best predictors. Thanks to its ability in dealing with seasonality, the SARIMA is the most used (33 papers), and was showed to give outstanding forecast (12 papers). Advanced exponential smoothing was used in 29 papers while structural time series in 15 papers.

3.4 Mixed Models

In this section, I analyse models that are at half-way between Econometrics and Time-series methods. Whereas the ADLM and the ECM extend the static single equation model by introducing time dynamics, the already dynamic time series models discussed in the previous section can be given similar extensions by including exogenous variables. We can regroup these models in the category ARMAX-based models, where X represents the exogenous variables. Both the ADLM and the ECM emphasise measurement of the cause and effect relationships between influencing factors and tourism demand. On the other hand, the ARMAX-based models focus heavily on discerning the dynamics of tourism demand. An ARX model was used by Li et al. [122] to assess the effects of relative climate variability on seasonal tourism demand. The ARMAX model was used in work of Pan and Tang [123] where outperforms the ARMA model in forecasting hotel occupancies as well as ARIMAX generates better long-run forecasts of the numbers of airport

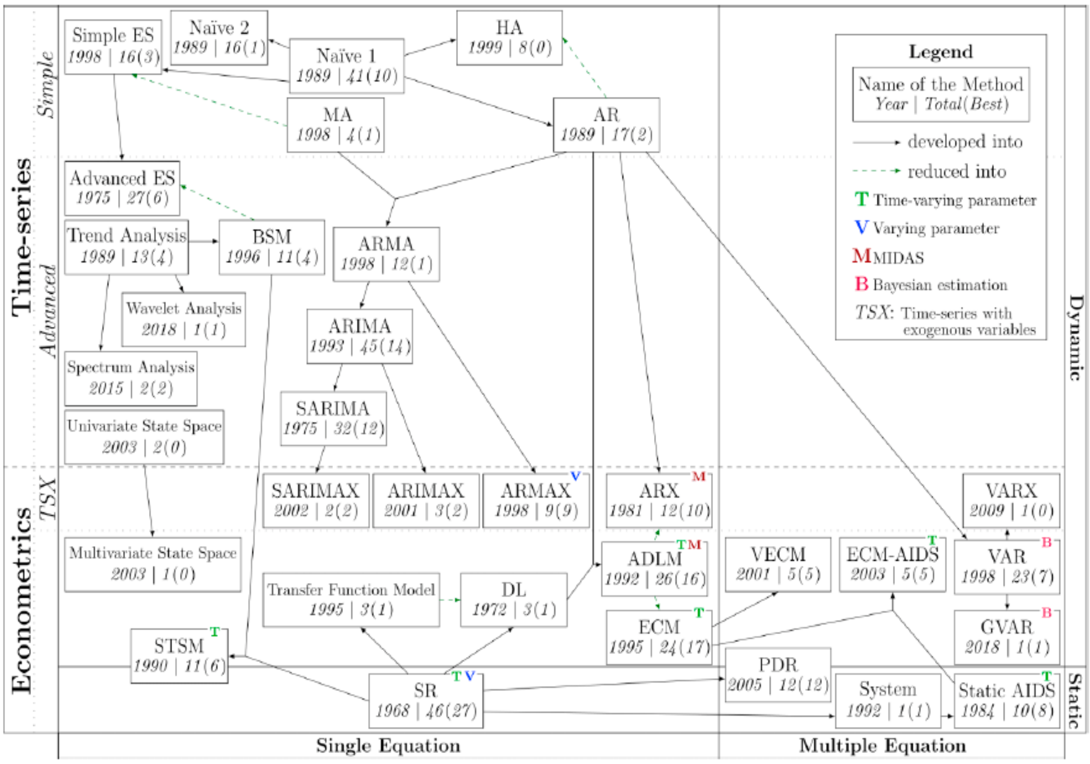


Figure 3.1

passengers in Hong Kong than the SARIMA model [124]. SARIMAX gives better forecasting respect to classical time series models in the work of Park et al. [125]. Figure 3.1, taken by by Song et al. [78], represents a resume of the time-series and econometrics models used in tourism demand forecasting. There are included techniques that I have not treated. By means of arrows Song outlined the relationships that these models present. It is also reported the year of the first introduction in touristic field, the number of researches about it and in brackets how many works found it to be the best predictor.

3.5 Artificial Intelligence Models

The first paper regarding tourism demand forecasting using artificial intelligence was written in 1999 [126] and since then, the adoption of AI-based techniques has seen a substantial development in particular in the last decade. The strength of these kinds of models is that they are able to explain complex non-linear data, without a priori knowledge about the relationships between input and output. AI models are data-driven and model-free techniques. There are some researchers that questioned the explanatory value of these model. For instance in artificial neural network the input enter in a ‘black box’ where it is transformed to produce an output, but what happen inside the box is very often too much hard to interpret. Furthermore the absence of an apriori model leads to a lack of theoretical background. Despite these drawbacks AI-based methods are widely used to predict phenomena in scientific fields with very successful results, for these reason touristic researchers have started to use them more and more frequently in forecasting tourism demand. Following Song et al. [78], I found five AI techniques used for touristic purposes, the artificial neural network (ANN), support vector regression (SVR), fuzzy time series, the rough sets approach and grey theory.

3.5.1 ANN

The artificial neural network is a computation system inspired to the biological neural network contained in the brain. This system is able to learn without any previous knowledge about what it is analyzing. It is composed by nodes (neurons) which are connected to other (synapses). Edges transmit a signal to the output neuron that transforms with a non linear function all the values contained in the input neurons. It is an excellent method of prediction and it is by far the AI method most used in tourism flow forecasting, but it suffers from the problems mentioned in the previous section. The first work in touristic field using AI-based method, implemented an ANN [126], but it was also used in more recent works. For instance in 2016 Constantino et al. [127] used it to forecast the number of overnight stays in hotels and similar places in Mozambique using harmonized index of consumer prices, gross domestic product per capita, and exchange rate (ER) as explanatory variables. In ANN various training algorithms can be used. For touristic purposes, but in general for time-series analysis, Chen et al. [128] suggested to use the backpropagation, which thanks to its high learning accuracy and quick retrospect speeds can achieve higher performances. Thus, BPN has been applied in tourism demand forecasting as well, for instance in Lin et al. [83] to forecast touristic demand in Taiwan.

3.5.2 SVR

The support vector regression is a method that, in contrast to linear regression, is concerned about contain error in a certain range. The analyst must specifies a certain margin, called the maximum error, which represents the tolerance to deviation from prediction. To obtain the regression the machine learning algorithm called support vector machine is used. SVR has frequently appeared in touristic papers, as an example I report the work of Cang and Yu [129]. Actually this method is often combined with other algorithms and models, I will talk about these hybrid techniques in section 2.5.6.

3.5.3 Fuzzy Time Series

Because the existing statistical time series methods could not effectively analyze time series with small amounts of data, fuzzy time series methods were developed. The beginning of this theory is the work of Zadeh [130] which described by means of mathematical models the fuzzy information. The fuzzy is a logic in which one can attribute to each proposition a degree of truth different from 0 and 1 and between them. It was integrated in time-series analysis in 1993 by Song and Chissom [131]. An application in touristic field is the work of Wang and Hsu [132] who analyzed the tourism arrivals in Taiwan. As for SVR, this method is often combined with other algorithm (see section 2.5.6).

3.5.4 Rough Set Approach

Rough set approach is by far the less used, indeed only one study used it in tourism demand forecasting in the past 10 years [77]. Rough set approach is an informative technique that is good at handling vague data and identifies relationship and patterns in hybrid data, with both quantitative and qualitative information. In the opinion of Goh et al. [133] numerous tourism data have a high degree of vagueness and roughness; for this reason, the use of rough set based model is ideal in this field. They found that qualitative noneconomic factors, such as leisure time index and climate index, have stronger impacts on tourist arrivals to Hong Kong from USA and UK than economic factors.

3.5.5 Grey Theory

Grey theory, was developed by Deng in 1982 [134]. It focuses on model uncertainty and information insufficiency. In the field of information research, deep or light colours represent information that is clear or ambiguous, respectively. Meanwhile, black indicates that the researchers have absolutely no knowledge of system structure, parameters, and characteristics; while white represents that the information is completely clear. Colours between black and white indicate systems that are not clear, such as social, economic, or weather systems. The grey forecasting model adopts the essential part of the grey system theory and it has been successfully used in various fields. In tourism forecast was used by Sun et al. [135] to predict annual foreign tourist arrivals to China

3.5.6 Hybrid Models

All the AI models described above have advantages and disadvantages, so in the literature they were combined in order to improve forecasting accuracy minimizing their limits. SVR is very often combined with genetic algorithm in order to select hyperparameters [136] as well as with a recently developed new algorithm called fruit fly optimization [137]. Genetic algorithm was also combined with fuzzy time series to forecast monthly tourist arrivals to Japan [138]. AI models were also merged with classical methods, Chen [139] proposed an hybrid model composed of a linear autoregression structure modeled using ES, ARIMA and Naive and ANN and SVR to model nonlinear residual.

Despite their pitfalls, AI-based are for sure the forecasting models that will experience greater development. Indeed these are the techniques that well deals with big data. However, questions remain about the interpretation of the analytical results when such data-driven techniques are applied.

3.6 Model Comparison

The touristic demand forecasting literature is in continuous development, models that were widely used in the past have practically disappeared while new techniques were born. Usually the model adopted in earlier works are used as benchmark in most recent papers, it is the case of linear regression, that from the '60s to '70s was one of the main methods and now is primarily used to test the performances of more sophisticated techniques. Following the work of Song [78], we can draw a chronological development of the last six decades. in '60s and '70s the single regression was in absolute the dominant method until, in '80s, it was surpassed by time series techniques. In '90s the attention was shared between Box-Jenkins approach and dynamic regression methods such as the ADL, but it is the decade when also structural time series approach began to take hold. In the first ten years of 21° century there was an expansion of econometric techniques and the artificial-intelligence based method started to be implemented. The last decade saw a flourishing of hybrid methods that combined various models in order to improve accuracy of forecasting. Many works compared the performances of techniques through the use of mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE). The first simply consists in computing the mean of the difference between the forecasted and the actual value in absolute value.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{Y}_i - Y_i| \tag{3.15}$$

One problem with the MAE is that the relative size of the error is not always obvious. Sometimes it is hard to tell a big error from a small error. For this reason the MAPE is used. The difference is that inside the absolute value the difference among actual and forecasted is divided by the actual value.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \tag{3.16}$$

Since both of these methods are based on the mean error, they may understate the impact of big, but infrequent, errors. To adjust for large rare errors, the RMSE is computed. By squaring the errors and then taking the square root of the mean, we arrive at a measure that gives more weight to the large but infrequent errors.

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(\hat{Y}_i - Y_i)^2}{n}} \tag{3.17}$$

Although the several works and developements, the conclusion of Li et al. [140] that no single forecasting method has been found to outperform another method in all circumstances still seems to be valid. However, there are some models that seems to be the best predictor more times than others. For instance, among ARMA based approach analyzed by Jiao and Chen [77] the ARFIMA models is the one that more times outperforms the others, followed by SARIMA and ARARMA.Structural time series models are considered superior techniques together with the SARIMA-GARCH model. The class of model that in absolutely seems to have the best performance more frequently are the hybrid ones, for this reason, many researchers think that this will be the direction that future research will take. Figure 3.2 is taken by Song et al. [78] and it is based on the sample they built. It summarizes the application of various models, including some that I have not treated, indicating in the left panel how many times they were used and in the right panel how many times they are selected as best predictor. A thorough presentation of the main works in tourism demand forecasting is contained in Song et al. [78] and, with a focus on the works written in the last decade, in Jiao and Chen [77].

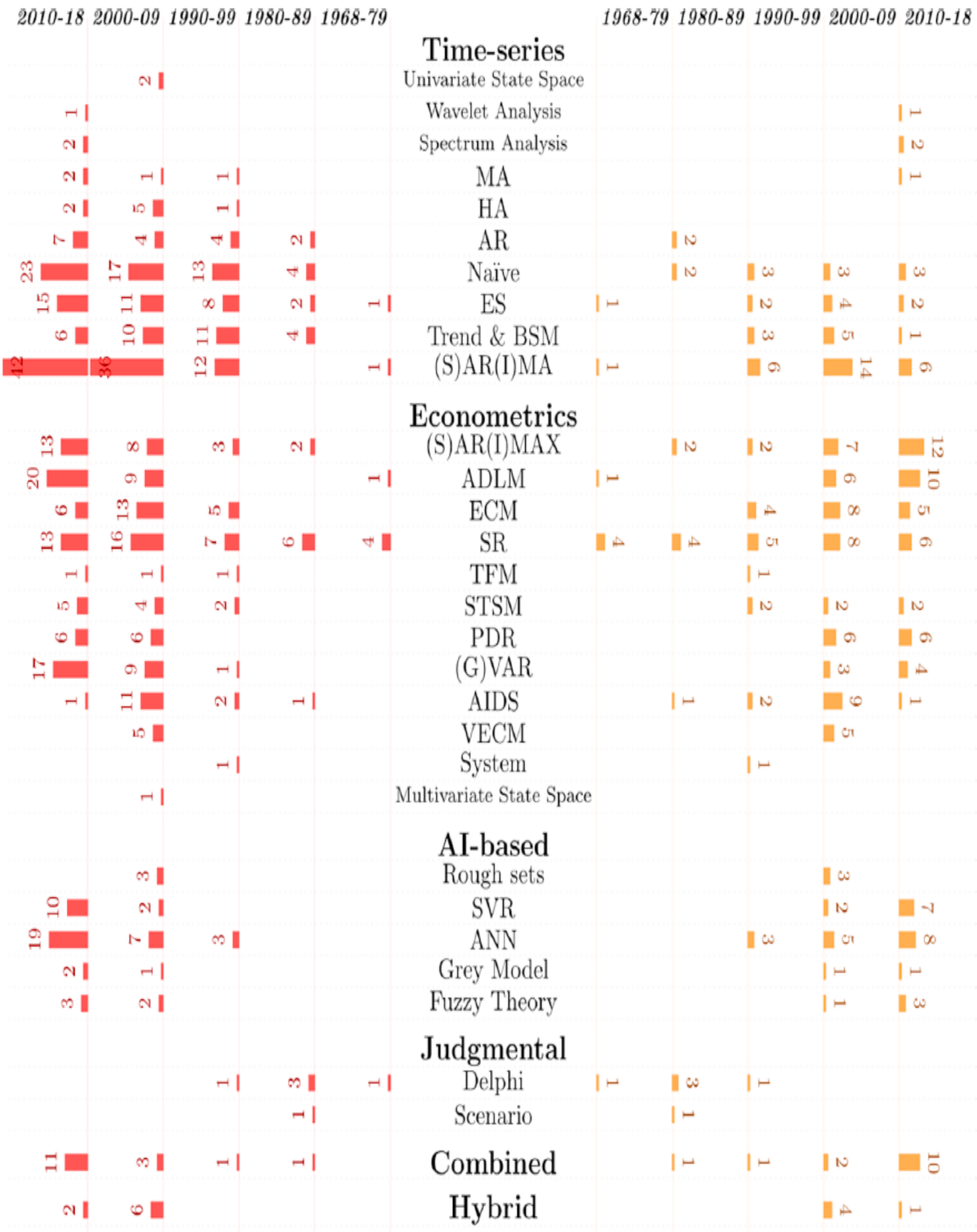


Figure 3.2: Models performances

Chapter 4

Data and Methodology

This work is based on the touristic relations among 24 countries. Unlike the vast majority of quantitative touristic research, this work has a global point of view, since each continent is represented by at least one nation. Indeed, reading the review on tourist forecasting of Song et al. [78], no one paper uses a global approach. The countries taken into account are Argentina, Brazil, Mexico, United States of America, Canada, South Africa, Australia, New Zealand, Thailand, Indonesia, South Korea, Japan, Israel, Turkey, Poland, Greece, Italy, Germany, France, Belgium, Netherland, UK, Spain and Portugal. Although these countries represent only the 10% of the overall nations, the majority of tourist flows pass through them. Following the annual UNWTO relations about international tourism [141], in 2017 the selected countries represented the 69,4% 41,2% 81,2% and 22.7% of European, Asia-Pacific, American and African tourist receipts and the 59.3% of the world tourist earnings. Furthermore, nine of them happear among the first ten countries by number of tourist arrivals. I am aware that I did not consider China, which is one of the most important countries in terms of tourist arrivals (fourth in 2017 [142]). I exclude it because of the difficulties encountered in the collection of data, since I found only series regarding the overall tourism arrivals and not the ones regarding the travellers inbound flows by country of origin. The same reason can explain the absence of other relevant countries such as India and Russia. In this chapter, I will analyze the data collecting and the methodology used to examine them.

4.1 Data

Several entities provide data about tourist flows, firstly, the national institute of statistics, then the ministries of interior and tourism, the government tourism agency and other private entities. At an international level, the most important institution is the world tourism organization (UNWTO), which is a specialized agency of united nations responsible for the promotion of sustainable and universally accessible tourism. Since the sources of information are so heterogeneous, it is practically impossible to find data collected with the same criteria. A high degree of harmonization about data collecting methodology can be found only among the states that belong to the European Union. Furthermore, almost all countries do not collect exactly the tourism inbound, but some proxies such as the number of visitors that stay for at least one night in a hotel, or the foreign citizens that pass the national borders. There are differences also among the meaning of foreign tourists, for instance, some countries use citizenship, while others the nation of residence. For all these reasons, the data used are not precise and must be considered as proxies of touristic flows. Although these pitfalls, all these proxies remark very well the non-stationary components such as seasonality and trend, which are the objects in analysis in this work. In this section, I am going to describe all the data used, specifying the source of information and the methodologies of collection.

4.1.1 Argentine

The data used for Argentine are taken by both the "Istituto Nacional de Estatistica y Censos" (Indec) [143] and UNWTO. The former provides monthly data of all the international airport movements presents in the territory of Argentine. The data are contained in the "Encuesta de Turismo Internacional" which measures the number of tourists and the average expense of non-resident and resident travellers. The data are collected from January 2013 up to December 2019. Since the datatabase provides only aggregate data about Europe, Asia and Oceania without specifying the country of origin, I integrate it with the data provided by the UNWTO "Country-Specific: Outbound Tourism 1995-2018". These database contains yearly observations of outbound tourism, specifying the country of origin and arrival. To make it monthly, I computed how much a state weights in percentage terms on the total amount of its area each year, and I assumed that the reported weight remains constant over the corresponding annual period. I make this assumption because the factors that affect seasonality in touristic data such as weather conditions and feast days are similar between countries belonging to the same geographical and cultural area, such as Europe or Asia.

4.1.2 Australia

The data used for Australia were produced by the "Australian Bureau of Statistics" and collected on the web site Knoema [144]. The observations are monthly and start from January 2010 up to March 2020. The database contains statistics on the travel movements of persons arriving in, and returning to, Australia divided by country of visitors residence.

4.1.3 Belgium

The data for Belgium are provided by the national institute of statistics "StatBel", in the table "Touriste selon pays de provenance" [145], the data available are yearly and start from 2015 up to 2018. To construct monthly data I took the ones of the neighboring Netherlands, for each country I checked how a month weight in the total year in percentage term. Then I multiplied the resulting weight for the Belgian date obtaining an estimate of the monthly time-series. I am aware that the seasonal pattern might be severely biased, but the trend is correctly preserved. The data reports the number of tourist arrivals intended as each time a person comes up at a hotel or any other accommodation and stay for at least one night. The visitors are split by country of residence, intended as the country the traveller stayed the most, in the last 12 months. As for Argentina, I integrated these data with the ones provided by UNWTO, in order to consider all those countries that are specified in aggregate terms only.

4.1.4 Brazil

The Brazilian database is provided by the "Instituto Brasileiro de Geografia e Estatìstica" (IBGE) and the minister of tourism of Brazil [146], which reports tourism flows data in the yearly publication "Anuario estatistico de turismo". The table "Chegadas de touristas" reports tourist arrivals by country of permanent residence. The data are monthly and starts from January 2012 up to December 2019.

4.1.5 Canada

The data for Canada are provided by the Canadian government in its website [147] and is divided in two databases, in the former are collected data of United States only, while in the latter information about the others countries are supplied. With travellers is intended a resident of a country other than Canada who is travelling to Canada for a period fewer than 12 months. The data are sorted by country of residence and recorded monthly for a period that starts from November 1993 up to March 2020.

4.1.6 France

French data are provided by the “Institut national de la statistique et des études économiques” (INSEE) [148]. The data collected are available only quarterly, but they are sufficient to appreciate the seasonal patterns. The tourist number is intended as the number of visitors that stays at least one night in hotels or other accommodations, and the visitor residence is intended as the country the traveller stayed the most, in the last 12 months. The data are available only for few countries of residence (Italy, Germany, UK, Belgium, USA, Japan, Spain and Netherlands), so I inferred the ones of other countries, integrating it with the UNWTO database as I did before. Unfortunately not even the UNWTO has data for Brazil, South Africa, Thailand and Indonesia, so I did not conidered in-flows from that countries. The observations start to the first quarter of 2011, up to the last of 2019.

4.1.7 Germany

German data are provided by the national statistic authority “Statistisches Bundesamt” (DESTATIS) [149]. In the monthly bulletin of tourism statistic “Monatserhebung im Tourismus” the institut reports the number of tourists who have stayed in a hotel or any other type of accommodation in the German territory sorted by country of origin, intended as the country of permanent residence of travellers. I used monthly data from January 2018 up to March 2020. Unfortunately neither DESTATIS nor UNWTO provide data for the flows from Thailand, Indonesia and Argentina.

4.1.8 Greece

Greek data are elaborated by the “Hellenic Statistical Authority” (ELSTAT) [150] based on the border survey made by the bank of Greece. The tourists arrivals are intended as the number of people that cross the greek boards. Data are sorted by country of visitors residence. The data are monthly and starts from January 2008, up to December 2016. To contruct the series of the following years, I took the yearly data of UNWTO, and I supposed that the seasonal pattern of 2016 has remained unchanged. For each country of origin, I computed how the total flows of 2016 are distributed among the months in percentage terms, and I applied the same proportions to the succesive years. Neither ELSTAT nor UNWTO provide data for Thailand, Indonesia and New Zealand.

4.1.9 Indonesia

Indonesian data are provided by the national institute of statistic “Badan Pusat Statistik” (BPS) [151]. The tourists arrivals are intended as the number of non-Indonesians that cross the national borders both by air and sea. The data are monthly and for a period that starts from January 2018 up to April 2020.

4.1.10 Israel

Israelian data are provided by the “Central Bureau of Statistics” of Israel [152]. Touristic information can be found in section E (Migration and Tourism) of the monthly bulletin of statistics. Table 7 provides the monthly number of tourists, intended as who enters Israel under a tourist visa and leaves it on a date other than the entry date (one day visitors not included), sorted by country of citizenship. I downloaded the bulletins from January 2012, up to April 2020.

4.1.11 Italy

Data about Italian tourist flows are provided by the Italian institute of statistics (ISTAT) [153]. The number of tourists is proxied by the number of guests that spent at least one night in a hotel or any other accommodation. The data are sorted by country of residence. The time series is monthly and starts from January 2008 up to December 2018.

4.1.12 Japan

Data about Japanese tourist flows are provided by the “Japan National Tourist Organization” (JNTO) [154] which is an independent administrative institution of the government of Japan responsible for promoting travel to and in the country. Visitor arrivals are calculated on the number of travellers passing the Japanese national borders. Those figures exclude permanent residents having Japan as their place of residence and include travellers entering Japan for the purpose of transit as well as international students. The data are sorted by citizenship. Time-series are monthly and start from January 1990 up to February 2020.

4.1.13 Mexico

Mexican data are provided by the autonomous agency of the Mexican Government “Instituto Nacional de Estadística y Geografía” (INEGI) [155]. The data reports the arrivals of foreign tourists in the 56 international Mexican airport. For each nationality, I summed up the arrival in each airports in order to estimate the total visitors in the Mexican territory. The data are monthly and starts from January 2012 up to April 2020.

4.1.14 Netherlands

Data about tourist flows in the Netherlands are provided by the "Centraal Bureau voor de Statistiek" (CBS) [156]. The number of tourists is proxied by the number of guests that spent at least one night in a hotel or in any other tourist accommodation. The data have monthly frequency and starts from January 2012 up to February 2019. There is a lack of data for the visitors coming from Thailand, Argentina, Brazil, Mexico and South Africa.

4.1.15 New Zealand

Data about New Zealand touristic flows are provided by the site "Figure.nz" [157], which is a public data provider that collect information about New Zealand. Tourist arrivals are intended as overseas residents arriving in New Zealand for a stay of less than 12 months. They are sorted by port of entry and by place of residence, intended as the country where the person last lived or will next live for 12 months or more. Since the database supplies data for geographic area, I disaggregate them using UNWTO database as did for Argentina. The time series are monthly and start from April 1978 up to March 2020.

4.1.16 Poland

Data about polish tourist flows are provided by the national institute of statistics "Główny Urzad Statystyczny" (GUS) [158]. The number of tourists is proxied by the number of guests that spent at least one night in a hotel or in any other tourist accommodation establishments that possessing ten or more bed places. The data are sorted by country of residence. The time series is monthly and starts from January 2017 up to February 2020.

4.1.17 Portugal

Data for Portugal are provided by the "Instituto Nacional de Estadística" (INE) in two different reports [159]. The former reports the number of guests that spent at least one night in holiday camps, intended as a holiday complex with appropriate facilities for providing free or low-cost holidays, usually as a social service by public or private entities. The latter indicates the number of guests that spent at least one night in a youth-hostel, intended as a non-profit-making establishment providing accommodation for young people or limited groups of young people. The data are sorted by country of residence and collected monthly for a period that starts from June 2008 up to December 2018. Since holiday camps and youth-hostel are only a little part of the accomodations, I summed up the data contained in the two files, I extrapolated from the resulting series the seasonal pattern and I applyied it to the yearly observations of the UNWTO. Unfortunately neither INE nor UNWTO provide data for the flows from, Thailand, Indonesia, Argentina, Mexico, South Korea and New Zealand

4.1.18 South Africa

Data of South Africa are provided by the department of statistics of South Africa (STAT SA) in its yearly report [160]. The bulletin includes information on South African residents and foreign travellers who passed through all South African ports of entry/exit (air, land and seaports). It summarises data published in the monthly statistical releases on Tourism and Migration (Statistical release P0351), highlighting annual and monthly numbers of travellers and tourists providing Details on the country of residence of tourists. Time series are monthly and starts from 2017 up to 2019.

4.1.19 South Korea

Data about South Korea tourists are provided by the “Korea Tourism Organization” (KTO) [161], which is an organization of the Republic of Korea (South Korea) under the Ministry of Culture and Tourism aimed at promoting the country’s tourism industry. The data are sorted by nationality, are monthly based and strat from January 2014 up to March 2020.

4.1.20 Spain

Spanish data are provided by the “Instituto Nacional de Estadística” (INE) [162]. The number of tourists is intended as the number of the non-resident travellers that enter in Spain by all the access routes (airport, road, harbour and train). The data are monthly and starts from October 2015 up to March 2020. The travellers are sorted by country of residence. As did before, because of aggregation of states, I integrated them with the tourist outbound data provided by the UNWTO. Unfortunately neither UNWTO nor INE provide data for Thailand, Indonesia and New Zealand.

4.1.21 Thailand

Data on Thailand's tourism are provided by the "National Statistics Office" (NSO) [163] of Thailand. In the database are reported the number of international arrival in Thailand's territory sorted by nationality. The data are available only on a yearly base, for this reason, I extrapolate the seasonal pattern of Indonesian data and I integrate it in the Thailand's series as did for Belgium. I choose Indonesia because it is a country similar to Thailand in terms of factors that can influence seasonality, such as weather conditions and culture (feast days). The data starts from 2009 up to 2018.

4.1.22 Turkey

Data on Tourism in Turkey are provided by the "Turkish Statistical Institute" (TURKSTAT) [164]. The visitors are intended as foreign or citizen visitors resident abroad above the age of 14 who visit Turkey for the purpose of staying at least one night or same-day (except the military personnel and diplomats). The travellers are sorted by country of citizenship. TURKSTAT provides quarterly data only between 2006 and 2012. For the successive period, exclusively yearly data are furnished. For this reason, I used the seasonal pattern extrapolated in 2012 to estimate the one of the following years. Doing this I assume that the seasonal characteristics remained inhaltered during time.

4.1.23 United Kingdom

Data about United Kingdom tourist flows are provided by the website "visitbritain.org" [165], which reports visitors who spend at least one night in the UK during their trip, sorted by nationality. Data are available only quarterly, but it is sufficient to assess the seasonal pattern. The time-series starts from the first quarter of 2002, up to the last of 2018.

4.1.24 United States of America

Data for United States are provided by the U.S. Department of Commerce together with ITA, I&A and the National Travel and Tourism Office (NTTO) [166]. Data are reported in a monthly publication "Summary of International Travel to the United States" highlighting overseas visitor arrivals by country of residence. The travellers are intended as the people who cross the national borders with a valid touristic visa. Time series start from January 2011 up to April 2020.

4.1.25 UNWTO

Because of the lack of data, I often integrated the information previously described, with data provided by the world tourism organization (UNWTO) in its e-library [167]. In particular, a lot of databases report arrivals series by aggregate geographic region or by continent. I use UNWTO data to extrapolate the country specific data, as described in paragraph 3.1.1. For each country, the UNWTO provides the number of tourism outbound, specifying the country of arrivals. The pitfall is that data are available only yearly, thus not allowing extrapolation of seasonal features. There is no uniform methodology of data collecting since The information is obtained on the basis of data supplied by each of the destination countries, therefore correspond to the definitions of tourist arrivals used in these countries.

Differences in data sources can leads to inaccurate observations, for this reason I spent an entire section to outline the different methodologies. Probably the countries that estimate tourist arrivals through the nights spent in accommodations underestimate the true number of travellers since they do not take into account the daily tourists, while the countries that use the foreigners that pass the borders tend to overestimate the true number of travellers since they also consider who enters for other purposes. Unfortunately, the use of different sources is the only method to collect global data about tourist flows, since there is not an international organism that collects on its own such a data [1]. Although this problem, the main features of touristic time series such as seasonality and trend are preserved, allowing us to implement an effective structural time series analysis.

4.2 Methodology

The aim is to capture the tourisms connections among countries together with their dynamics. For this task, I used a social network approach as widely did in touristic literature [168][169]. I intended the network as a set of elements and a set of ties between them, where the elements are the 24 countries previously listed, and the ties are the tourist flows between pairs of country. I constructed a network for each quarter of 2018 and 2019 and I analyzed how they change in terms of centrality measures. In particular, I computed the weighted outdegree and indegree of each vertex, analyzing how the countries importance in terms inbound and outbound flows changes among the years and quarters. The weighted indegree of country A corresponds to the total touritic inflow in that country. In

[1] The UNWTO doe not collect data on its own, it uses data provided by national authorities

contrast, the outdegree corresponds to the total touritic outflow of A. A graphical representation of networks is provided using the software GePhi, which allows to modify the sizes of vertices and edges on the basis of their degree and wheight respectively. Although network analysis was widely used in tourisms researches, until now, quantitative research has been lacking to understand how the dynamics of global tourism networks have changed over time and how these networks affect and are affected by, tourism supply and demand. The first work on this topic was written in 2020 by Chung et al. [170] that conducting the first international tourism study to adopt a social network analysis approach that quantifies the complex structure of global tourism networks. Differently to Chung et al., to fill this research gap, I analyzed the dynamics of edges by means of a structural time series model, as described in the second chapter. I split the analysis into two parts, in the former one, I considered the time series of each connection from the first observation available to the last observation before January 2020. I do not consider successive realizations in order to do not take into account the structural break due to the coronavirus case, which is assessed in the second part. The model fitted is the local linear trend with seasonality, which is a powerful tool for time series displaying upward or downward trends together with seasonal patterns. The equation for the trend is stated below:

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t \beta_{t+1} = \beta_t + \zeta_t \tag{4.1}$$

the trend slope behaves as a random walk. In some cases this model is not suitable, since it allows trend to continue its direction causing the series to assume unreasonable values (e.g. negative number of tourists). For this reason, in some cases, I used a more conservative model for the trend specification, stated in the equation below.

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t \beta_{t+1} = D + \phi(\beta_t - D) + \zeta_t \tag{4.2}$$

Here the slope follows an AR1 model with $-1 < \phi < 1$, centered in a potential non-zero value D, instead of a random walk. Each series is analyzed using the R packages 'bsts' and 'KFKSDS', the former is used to fit a structural model together with the functions 'AddSeasonal' and 'AddLocalLinearTrend' or 'AddGeneralized-LocalLinearTrend' depending on which model follows the trend. Then the future values of edges for the period 2020-2025 are estimated through the 'predict' function together with the 'KFKSDS' function in order to specify the use of Kalman filter as prediction method. By means of the forecasted values, future networks structures are estimated and a qualitative analysis is reported. The next chapter finishes analyzing the trend resulting from the structural analysis, thus estimating which will be the countries that will gain or lose importance in the global tourism network both in inbuond and outbound tourist flows. Coronavirus outbreak is considered in the last chapter, analyzing the policy on human mobility introduced in the sample of 24 countries used in this thesis. The information about tourism restriction is provided by the UNWTO and by the national government websites. Looking at the gouvernment behaviour, I forecasted the possible measures that can be adopted by the various policy-maker, and the duration of them. To take into account the bans, I deleted the edges that were or are expected to be interrupted because of the restrictions and the compulsory quarantine. Then, I computed the nodes weighted out and in degree of the resulting networks, analyzing how they are expected to change. To capture the differences with the network of a normal year, I computed the density metric, that give a measure of how the general global touristic level will diminuish.

Chapter 5

Results

In this chapter are presented the tourist flows analysis results. How stated before, here, the massive decline of touristic flows due to the Covid-19 pandemic is not taken into account, indeed, this chapter aims to analyze the tourism trend of each country, searching to predict how the importance in terms of both inbound and outbound flows, will modify in the next five years. In the next section are presented the networks of the quarters of the two past years. In section two are shown the time series of the edges and the predictions for the period 2020-2025. In section 3 are reported the results of a becktesting implemnted on the time-series of Italian edges. In the last section, the networks derived by the forecasting of series are displayed and commented. I remind that the results regard the flows among the 24 countries listed in the previous chapter, so the past and forecasted number of tourists does not correspond to the total of these countries, as well as the resulting network structure does not correspond to the actual global one.

5.1 Past Networks

5.1.1 Inbound Analysis

In figures 5.1-5.4 are presented the networks of the four quarters of 2018, the size of nodes and edges are based on their in-degree values. In the first quarter USA, Spain, France and Italy are the locations with the most significant number of visitors. Looking at the image is possible to see that the USA inbound tourists are pushed up by the Mexican and Canadian travellers, while the tourists in Spain come mainly from the UK and France. In contrast to these countries, Italy and France have more heterogeneous visitors, showing edges of uniforms sizes. In the second quarter, France becomes the most visited country, surpassing the USA as well as Italy surpassed Spain. In this period, their weights started to be less diversified, indeed, it is possible to notice a marked link between Germany and Italy and between UK and France. Mexico and Canada remain the most important source of tourists for the USA. In the third quarter, coastal locations become more central in the network. Spain reached the first position in terms of weighted-indegree value, Italy the third places while Greece gained five positions. In this period the inequality in edges size continues to accentuate, showing a significant dependence of Spanish tourism to the German, British and American visitors as well Italy is strongly dependent to the German and Dutch visitors. In the fourth quarter, the edges seem to be more homogeneous again, except for Spain, which presents a strong dependence on the flows come from UK France and Germany.

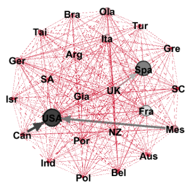


Figure 5.1: I-quarter 2018

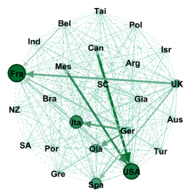


Figure 5.2: II-quarter 2018

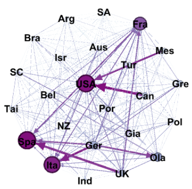


Figure 5.3: III-quarter 2018

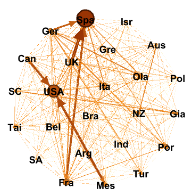


Figure 5.4: IV-quarter 2018

In figure 5.5-5.8 are reported the networks for 2019. The nodes size is given according to the in-degree value and the thickness of the edge according to their weights. There are no evident changes among the two years, except for USA and

Italy that surpassed Spain in the summer quarter. Spanish tourism suffered a decrease in visitors from the UK and the USA. In appendix A are shown the table with the numeric results

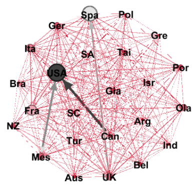


Figure 5.5: I-quarter 2019

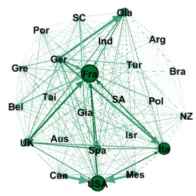


Figure 5.6: II-quarter 2019

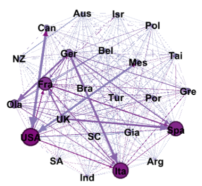


Figure 5.7: III-quarter 2019

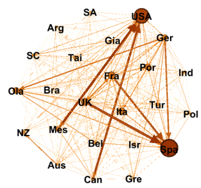


Figure 5.8: IV-quarter 2019

5.1.2 Outbound Analysis

In pictures 5.9-5.12 are reported the networks of the four quarters of 2018. Unlike the figures in the previous section, the sizes of nodes are set based on the out-degree values. The corresponding numeric measures are reported in the table in appendix A. Unlike in-bound ranking, out-bound one seems to be more stable throughout the year. Germany, UK, USA and France are ever in the first five places, while Mediterranean countries, such as Spain and Italy that occupy the highest positions in the in-bound list, fall to the half of the ranking. Germany is the leader in terms of visitors in the second and third quarters, showing a strong preference for for The Netherlands and the Mediterranean destinations such as Spain and Italy, while British seems to prefer only Spain among coastal places.

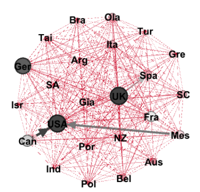


Figure 5.9: I-quarter 2018

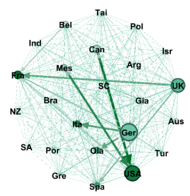


Figure 5.10: II-quarter 2018

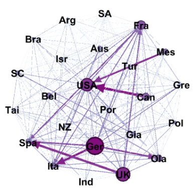


Figure 5.11: III-quarter 2018

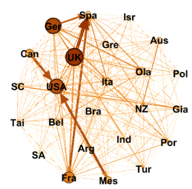


Figure 5.12: IV-quarter 2018

In figures 5.13-5.16 are reported the networks for the quarters of 2019. No evident changes are reported, except for a little increase of the US importance in the first and fourth quarters of the year.

5.2 Time Series Analysis

For each country, I computed a time series of the past and future total tourist in-flows, obtained as the summation of the past and forecasted series of each edge arriving in a node (inbound visitors in a country). Each country has 24 time series, plus the aggregate one. The overall time series are 600, to not break readability I

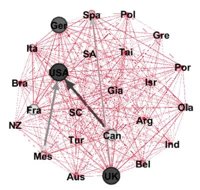


Figure 5.13: I-quarter 2019

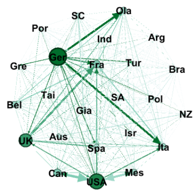


Figure 5.14: II-quarter 2019

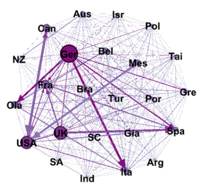


Figure 5.15: III-quarter 2019

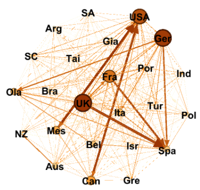
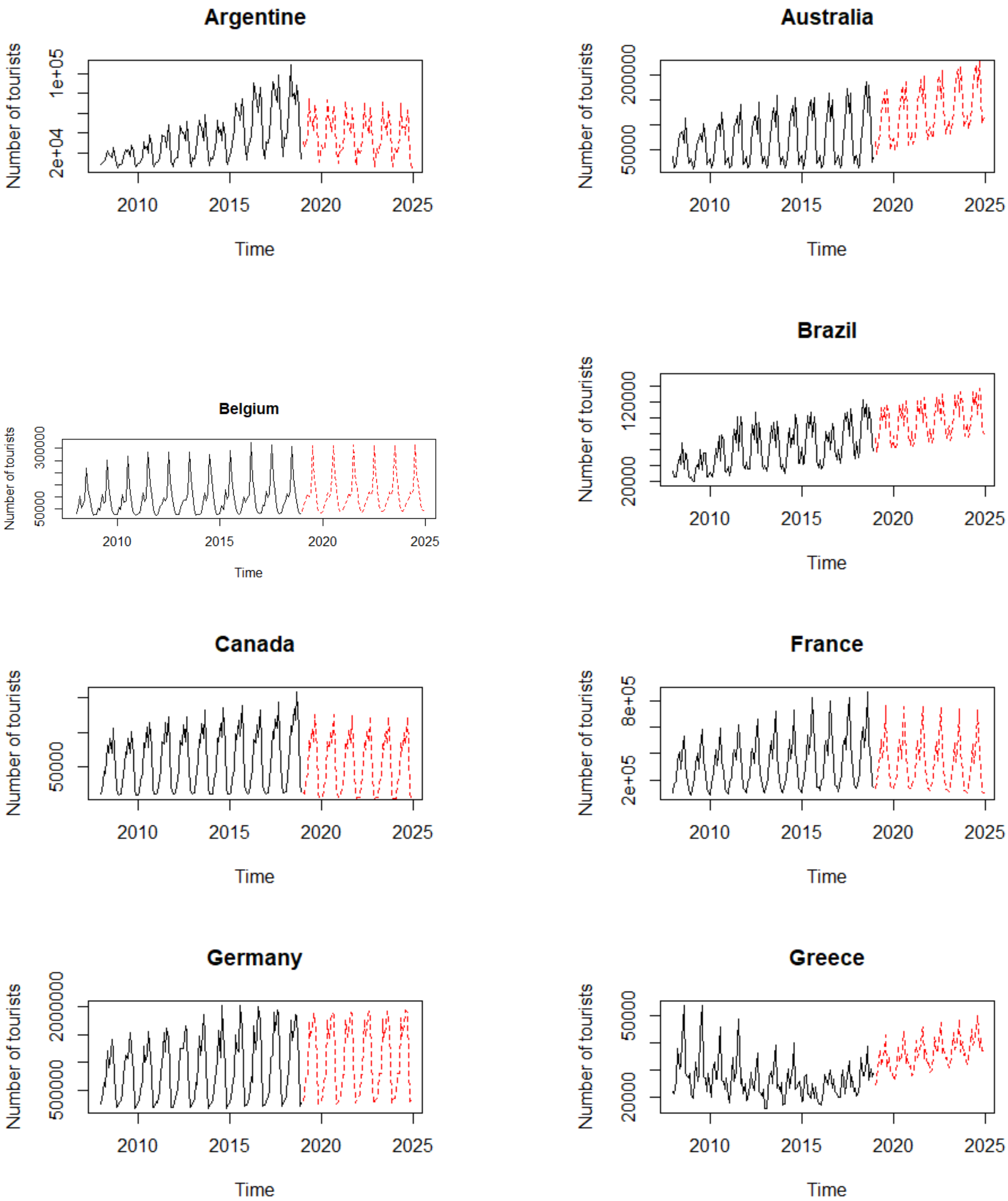
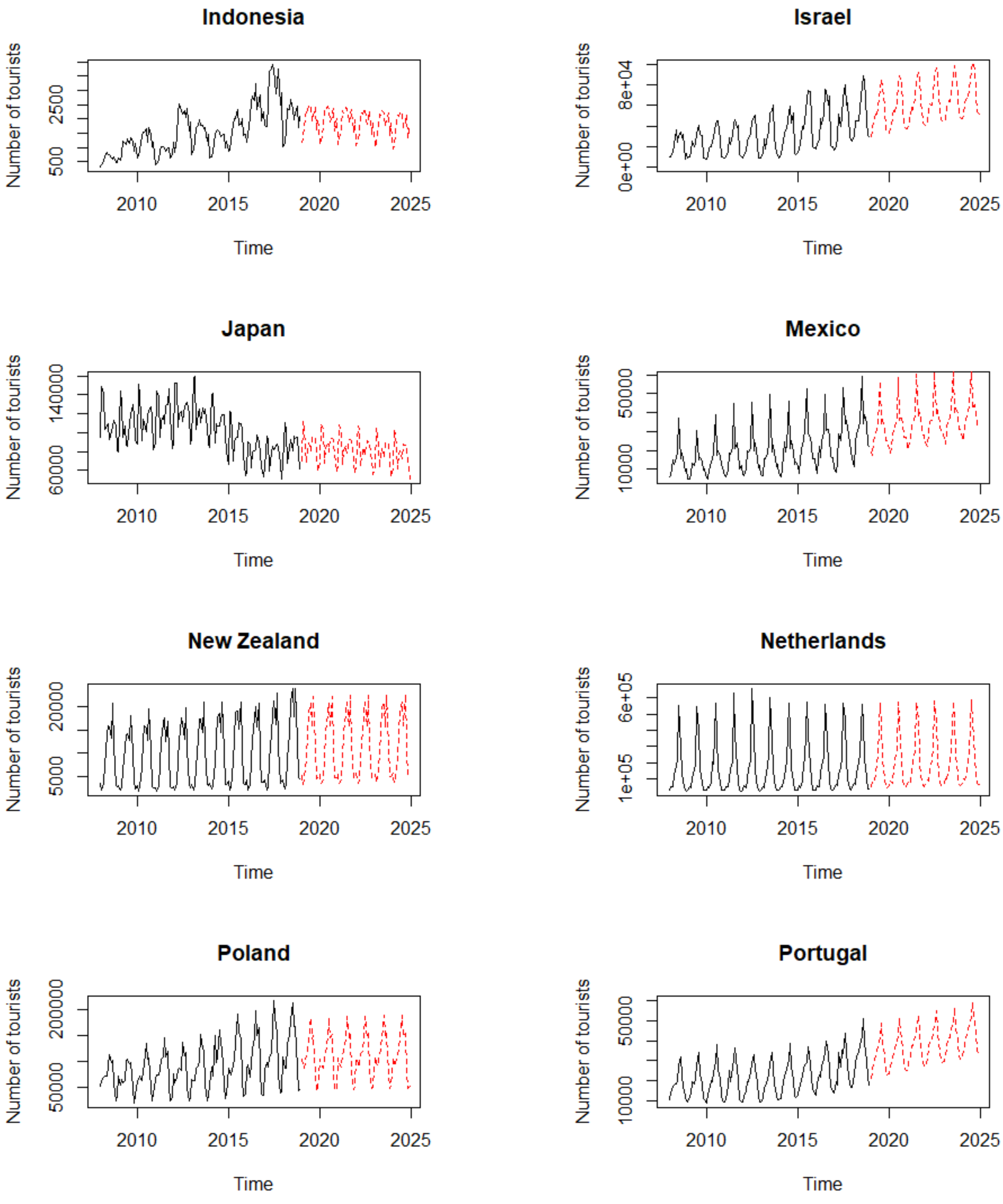


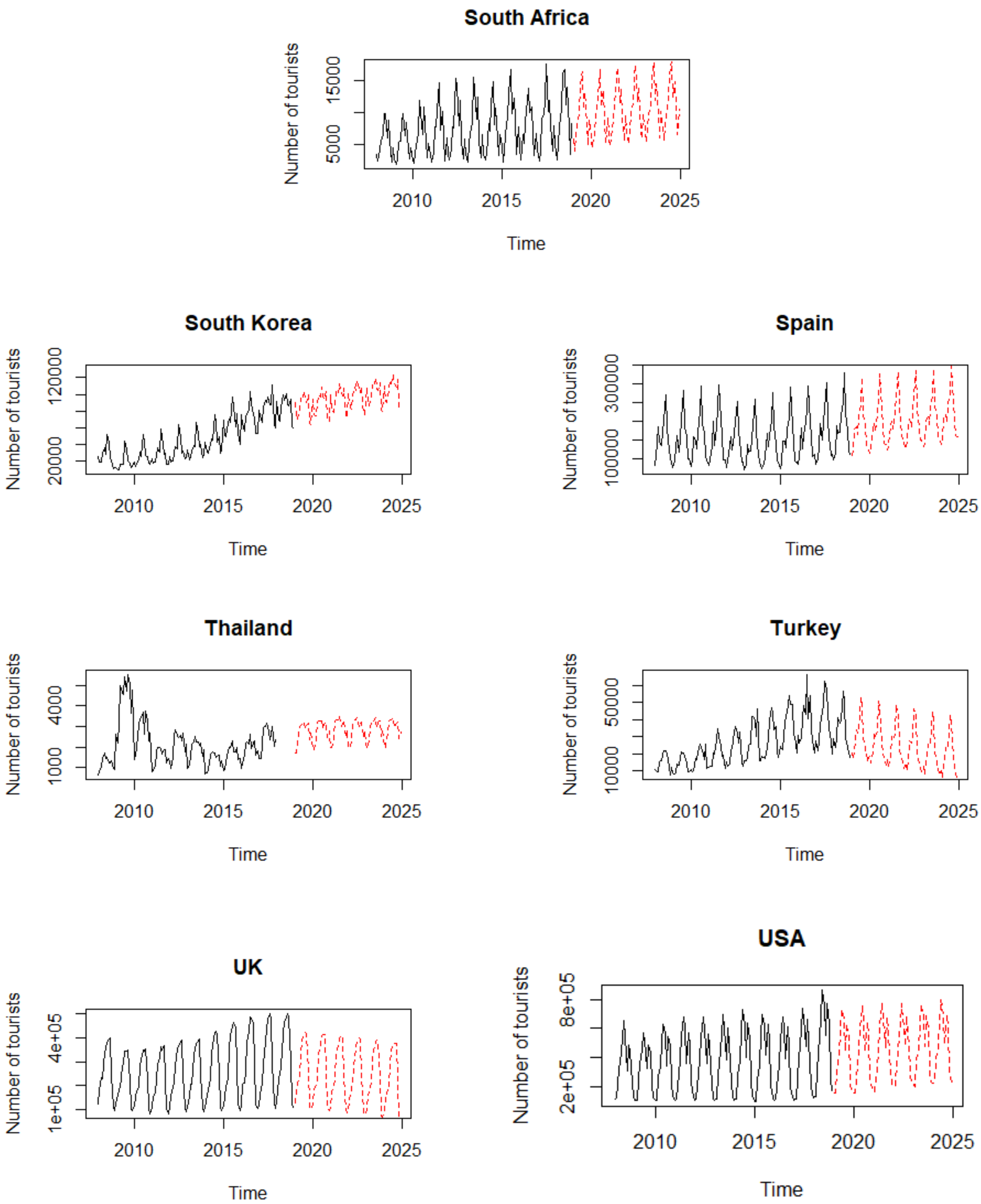
Figure 5.16: IV-quarter 2019

reported the series of each connection only for Italy, while for the rest of countries I reported only the aggregate one. In each figures, the black lines represent the past observations, while the red dotted line represents the forecasted values. In some series there are discontinuous observations, this is due to the lack of data, which are replaced with NaN values.

Italy In the figures below, are displayed the edges time series for Italy. It is evident in each plot a strong seasonal component, and a less accentuated trend component. Looking at the prediction, we can see that is expected an increase in the tourists from Australia, Brazil, Greece, Israel, Mexico, Portugal, South Africa and South Korea, while it is expected a decrease in the flows from Turkey and the UK. The prediction for the other countries is pretty much stable.


Argentine
Number of tourists
2e+04
1e+05
2010
2015
2020
2025
Time
Australia
Number of tourists
50000
200000
2010
2015
2020
2025
Time
Belgium
Number of tourists
50000
300000
2010
2015
2020
2025
Time
Brazil
Number of tourists
20000
120000
2010
2015
2020
2025
Time
Canada
Number of tourists
50000
2010
2015
2020
2025
Time
France
Number of tourists
2e+05
8e+05
2010
2015
2020
2025
Time
Germany
Number of tourists
500000
2000000
2010
2015
2020
2025
Time
Greece
Number of tourists
20000
50000
2010
2015
2020
2025
Time


Indonesia
Number of tourists
500
2500
2010
2015
2020
2025
Time
Israel
Number of tourists
0e+00
8e+04
2010
2015
2020
2025
Time
Japan
Number of tourists
60000
140000
2010
2015
2020
2025
Time
Mexico
Number of tourists
10000
50000
2010
2015
2020
2025
Time
New Zealand
Number of tourists
5000
20000
2010
2015
2020
2025
Time
Netherlands
Number of tourists
1e+05
6e+05
2010
2015
2020
2025
Time
Poland
Number of tourists
50000
200000
2010
2015
2020
2025
Time
Portugal
Number of tourists
10000
50000
2010
2015
2020
2025
Time


South Africa
Number of tourists
5000
15000
2010
2015
2020
2025
Time
South Korea
Number of tourists
20000
120000
2010
2015
2020
2025
Time
Spain
Number of tourists
100000
300000
2010
2015
2020
2025
Time
Thailand
Number of tourists
1000
4000
2010
2015
2020
2025
Time
Turkey
Number of tourists
10000
50000
2010
2015
2020
2025
Time
UK
Number of tourists
1e+05
4e+05
2010
2015
2020
2025
Time
USA
Number of tourists
2e+05
8e+05
2010
2015
2020
2025
Time

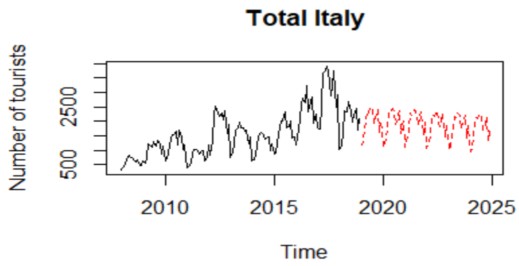


Figure 5.17: Total flow in Italy 2008-2025

In figure 5.17 is presented the series of the total flows in Italy, how we can see the prediction forecasts a stable number of visitors in Italy in the next five years. It is expected to maintain a strong seasonal component, with positive peaks in the summer months and negative in the winter months.

Argentine Since now, only the aggregate series are shown. In figure 5.18 we can see that neither an upward nor a downward trend is expected for the total tourist flows in Argentine. In particular, is expected a decrease of tourist from South-East Asian countries such as Thailand and Indonesia and of visitors from the richest countries in Europe such as Germany, Italy and France. This decrease is compensated by an increase in visitors from developing countries such as Turkey. How expected, unlike Italy, positive peaks are presented in the firsts and lasts months of the year.

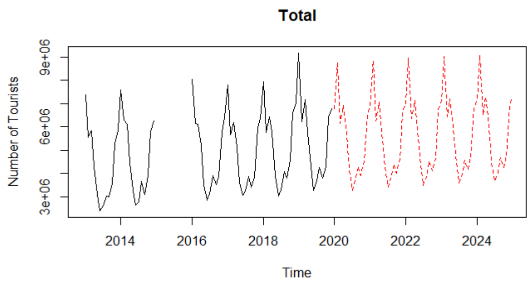


Figure 5.18: Total flow in Argentine 2013-2025

Australia Looking at the prediction in figure 5.19, Australia seems to be a destination that will gain even more importance in touristic terms. An increase in the number of visitors is expected for the period January 2020-January 2025. Looking at the series of each edge is possible to notice that this upward trend is led by an increase in visitors from Europe, in partcular from Spain, Portugal, Belgium, Germany and France as well as an increase in tourists from North America (Canada and USA). Visitors from neighbouring New Zealand are expected to remain constant. At the same time, it is predicted a slight decrease in the travellers from South Africa.

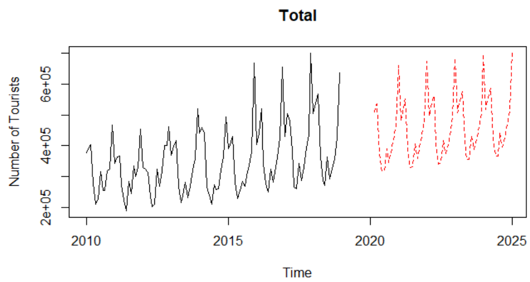


Figure 5.19: Total flow in Australia 2010-2025

Belgium Belgium shows a predicted increase in the number of tourists. An increase in inbound is showed in almost all countries except for United Kingdom, Italy, Turkey, South Korea and Japan. The number of total tourists is expected to increase from 7337168 of 2019 to 7746719 in 2024, with an exprected growth percentage of 5.58%

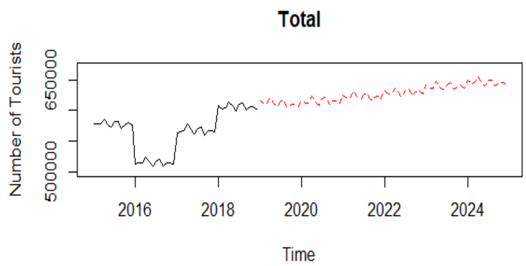


Figure 5.20: Total flow in Belgium 2018-2025

Brazil In figure 5.21 is showed the prediction for Brazil, the lack of data in the first part of the graphics is due to values not available. The flows are expected to suffer a slightly decrease, by varying from 4217870 in 2019, to 4024091 in 2024, but, looking at the image, no down-ward trend seems to be present. The single edge series reveal a constant values of connection over years, except for Portugal and Indonesia, which are expected to experience a mild increase in the number of tourists that will choose Brazil.

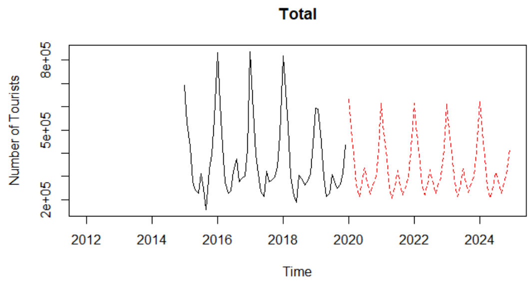


Figure 5.21: Total flow in Brazil 2012-2025

Canada The flows in Canada are expected to decrease from 13914776 in 2018 up to 8770482 in 2024, loosing the 36% of tourists in four years. This strong downward trend is mainly due to the decline in tourist arrivals from Europe and the United States, which is the most important source of inbound visitors in the country. These decrease completely offsets the expected increase in inbound flows of Mexican visitors.

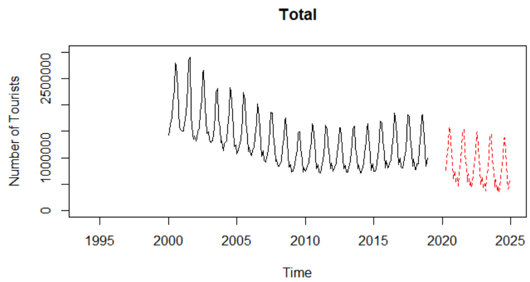


Figure 5.22: Total flow in Canada 1994-2025

France The flows in France are expected to fall in the next five years from 47405823 of 2018 up to 43514102 of 2025, loosing the 8% of visitors. The decrease is caused by a drop in the arrivals from the developed countries, such as Germany, Japan and the United States for touristic purposes. If the scenario were to come true, it would be detrimental to the France importance in tourism network, since they are the countries that weight the most in the touristic economy of France. Although this general decline, developing countries such as Mexico and Turkey shows an upward trend in terms of tourist arrivals in France.

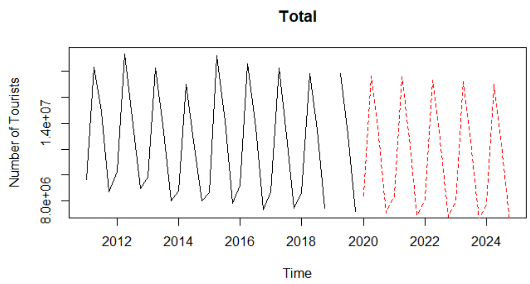


Figure 5.23: Total flow in France 2011-2025

Germany The volume of tourist inbound in Germany is expected to decrease by 35% in five years, by varying from 21378167 to 13753306 of 2019 and 2024 respectively. The decrease is general, without countries in contrast with this behaviour.

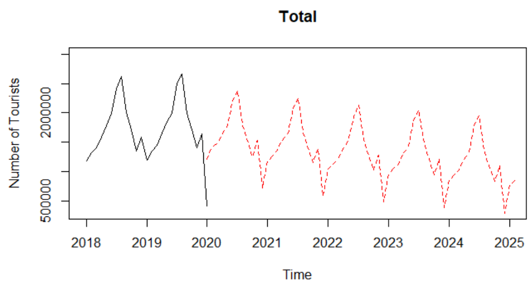


Figure 5.24: Total flow in Germany 2018-2025

Greece The number of tourists that choose Greece for their trips is expected to encounter a slight increase, by varying from 8086237 of 2019 to 9111449 of 2024. The countries that contribute the most are the European ones, which show a continuous increase in the travellers that plan to spend their holidays in Greece. A positive trend can also be found in the countries of South America, while for the rest of states is expected a constant volume of travellers.

Indonesia In figure 5.26 is presented the time series of tourist arrivals in Indonesia with its forecast. It is strongly evident an upward trend in the series, with an expected increase of 33% by varying from 4559293 tourists of 2019 to 6071756 of 2024. This behaviour is shared by the majority of the single edge time-series, except for the relatively close countries such as Australia and Japan.

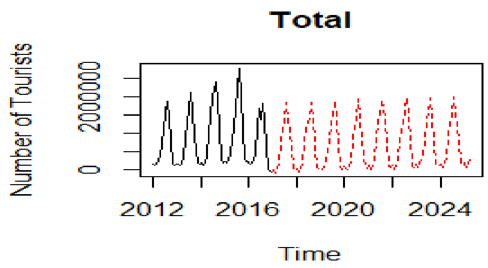


Figure 5.25: Total flow in Greece 2012-2025

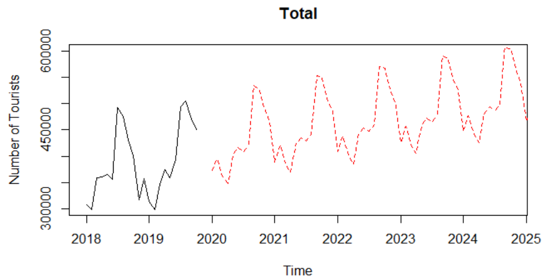


Figure 5.26: Total flow in Indonesia 2018-2025

Israel The time series forecasting predicts that Israel will gain more importance in the global touristic scenario, with an expected growth of visitors of 77%, by varying from 2311247 in 2018 to 5300547 in 2024. The only country that presents a downward trend is Turkey, while the rest of nations move in the same way of the aggregate time series.

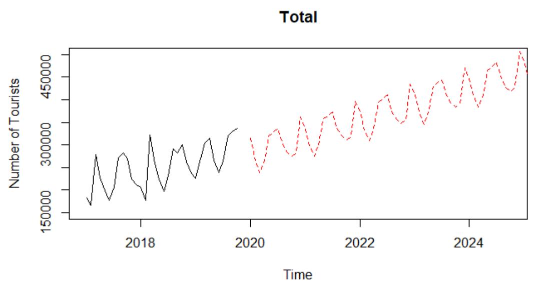


Figure 5.27: Total flow in Israel 2017-2025

Japan The constant growth that Japanese tourism experienced in the past decade is expected to continue in the next five years, increasing the number of tourists from 11015468 of 2018 to 15892128 of 2025. Looking at the dis-aggregate series, it is possible to notice that there is no one nation that follows the opposite trend.

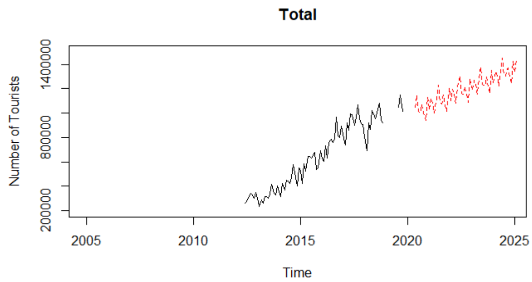


Figure 5.28: Total flow in Japan 2005-2025

Mexico As for Japan, Mexico has experienced constant growth in the past years and is expected to continue for the next five years. The model predicts that from the 17042531 visitors of 2018, the number of travellers will grow up to 20636113 of 2024, increasing the number of tourists of more than 20%. All countries contribute to delineate the total upward trend except for Italy and Argentine that show an expected constant path with no tendency neither to increase nor to decrease.

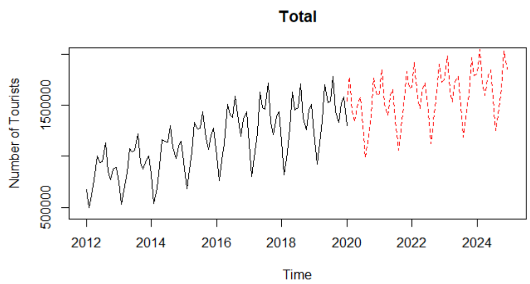


Figure 5.29: Total flow in Mexico 2012-2025

New Zealand The flows to New Zealand coming from the close Asian countries (Japan and South Korea) seems to be entered in a negative trend. Nevertheless, the positive ones of the other countries completely offset them and lead to a positive prediction for the number of visitors in the next five years. It is forecasted a percentage increase of 30% for the period 2017-2025, by varying from 9406192 of 2017 to 12197122 of 2025.

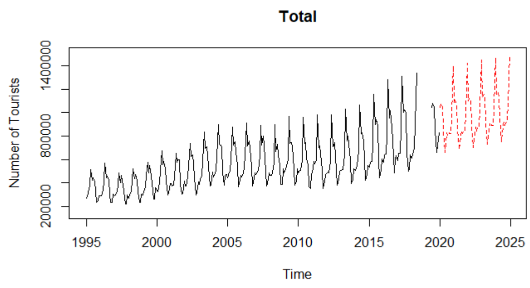


Figure 5.30: Total flow in New Zealand 1995-2025

The Netherlands Tourist flows in the Netherlands are expected to increase, by varying from 26717329 to 36286311 of 2017 and 2024 respectively. The European countries mainly lead this growth, that, with the exception of Italy, show all an upward trend. In contrast, American and Asian countries show a flat tendency.

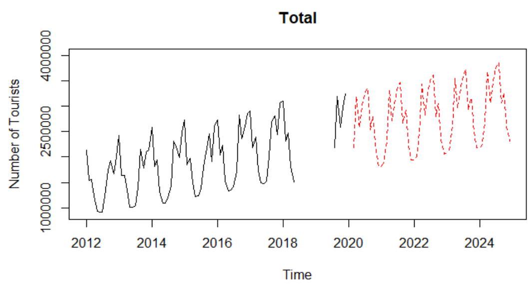


Figure 5.31: Total flow in The Netherlands 2012-2025

Poland Tourist flows in Poland are expected to increase of 30% from 2017 to 2024, by varying from 4241666 to 5472355. There is no geographic area that is expected to have a strong positive trend. Japan and Korea time series show the opposite behaviour of the aggregate one, as well as Spain and Portugal. The countries that are forecasted to increase the most the number of residents that will visit Poland are the UK and Germany.

Portugal Looking at figure 5.33 it is evident a predicted downward trend for the number of tourists that will visit Portugal. The model forecasts a decrease of 40% of tourists from 2019 to 2024, by varying from 11825657 to 6853835. Although this general decline, there are some countries that are expected to increase the number of residents that will visit Portugal in the next years. Spanish, South African,

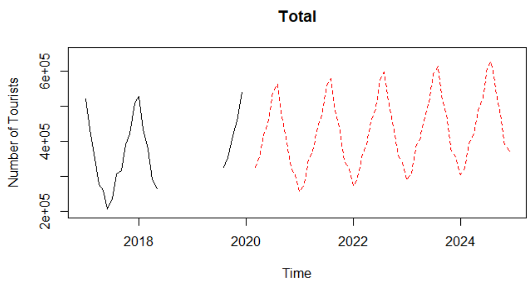


Figure 5.32: Total flow in Poland 2017-2025

British and US visitors are the ones that will increase the more, while all the other countries are expected to decrease or to follow a flat trend.

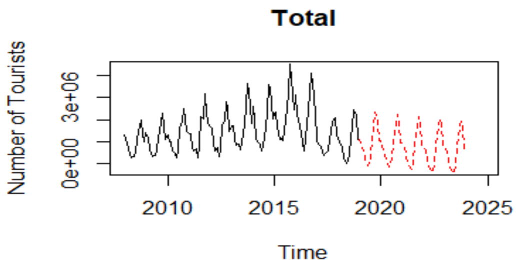


Figure 5.33: Total flow in Portugal 2009-2025

South Africa The model predicts a decrease of tourists in South Africa, by varying from 2046769 of 2019 to 1880261 in 2024. The loss of tourists is mainly due to a decrease in the volume of visitors coming from Europe, indeed there is no one country in this area that shows a contrary tendency. A different scenario is forecasted for Southern and Central American country as well as Australia, where an upward trend is expected.

South Korea South Korean tourist inbound are expected to increase in the next five years, from 5724955 of 2018, up to 6980435 of 2024. The growth is driven mainly by developed countries such as USA, Australia, Canada and European countries. The only nations belonging to this area that show a flat trend are the one that most suffered the economic crisis of the last years, Italy and Greece.

Spain Spain is one of the most important countries in terms of touristic flows, nevertheless, in the last few years, a negative trend has started. The model predicts

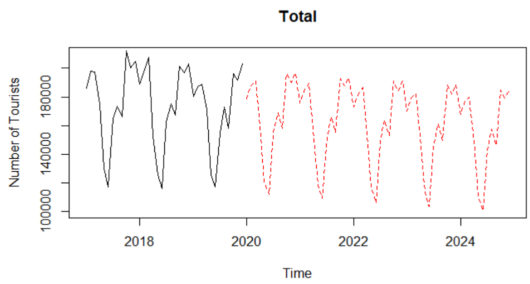


Figure 5.34: Total flow in South Africa 2017-2025

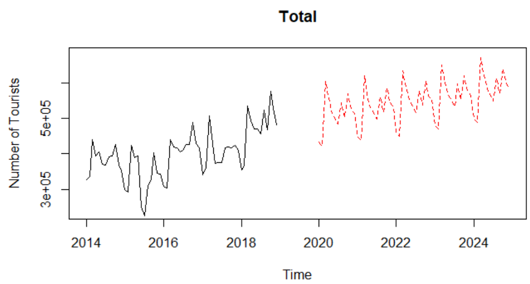


Figure 5.35: Total flow in South Korea 2014-2025

that this decline will continue thorough the next five years. The only countries that show constant growth are South Korea, Mexico and Greece. Although this general decline, the absolute value of arrivals in Spain remains one of the highest in the world. The expected number of tourist in 2024 is 37157697.

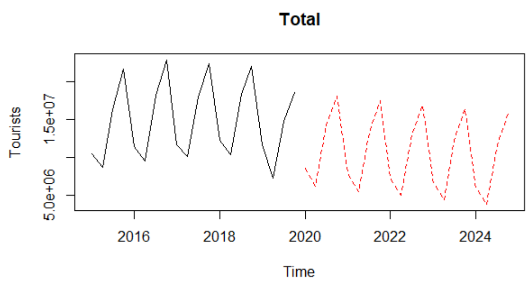


Figure 5.36: Total flow in Spain 2015-2025

Thailand The tourism in the Southern-East Asia countries faced a great growth in the last few years, Thailand is the leader among them. This positive trend is expected to continue in the next five years, with a forecasted tourism volume in 2024 of 12958388. In 2018 the visitors in Thailand was 10948119. The growth is driven mainly by European and North American countries, while a downward trend is present in developing countries of South America.

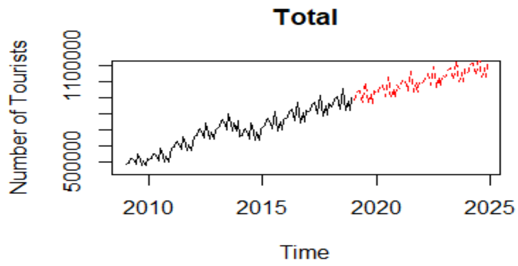


Figure 5.37: Total flow in Thailand 2009-2025

Turkey Looking at image 5.38 it is possible to see a flat tendency of the tourist volume in Turkey. Almost all places of origin share such behaviour. The exception are Indonesia, Thailand Argentina and Mexico, which showed an expected slight increase in the number of tourists visiting Turkey

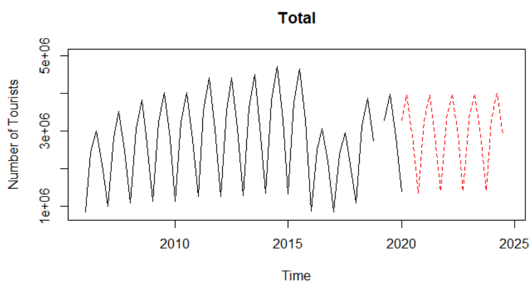


Figure 5.38: Total flow in Turkey 2006-2025

UK European touristic flows to the United Kingdom are still consolidated and present flat trends. The increase that it is possible to appreciate in figure 5.39 is mainly driven by developing countries. The touristic flows are expected to increase from 25571997 in 2019 to 27541742 in 2024.

USA The USA has ever been one of the leading countries in touristic terms, both in terms of inbound and outbound flows. The positive trend of the past

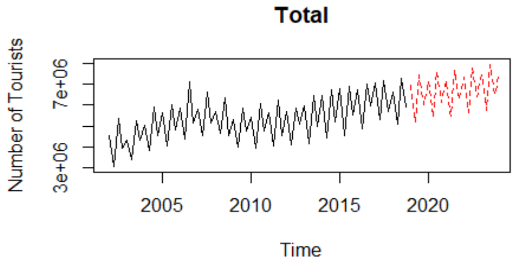


Figure 5.39: Total flow in UK 2002-2025

years seems to does not change direction in the prediction period. It is expected a growth of 13% from 2019 to 2024, with a number of tourists that varies from 60094446 to 68470439. Looking at the series, a flat trend is presented in the vast majority of series. A decrease is expected in the countries of South America, while an increase is forecasted in the extra American english-speaking countries, such as New Zealand and South Africa, and in Israel.

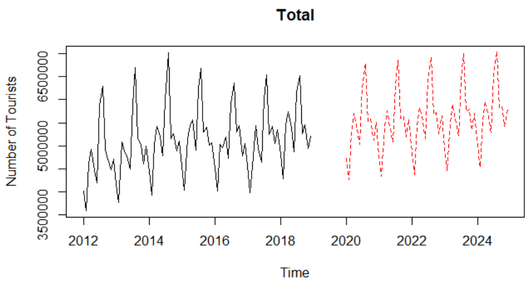


Figure 5.40: Total flow in USA 2012-2025

5.3 Backtesting

To investigate the effectiveness of the structural time series approach, I implemented a backtesting. I split the past data of the 23 time series referred to Italian touristic flows into two parts, the former composed by the first 75% observations and the latter by the remaining portion. I fitted the model taking into account only the first set, and I tested its reliability using the second one as virtual future. The measures used to check the goodness are the mean absolute percentage error (MAPE) and the root mean squared percentage error (RMSPE), defined by means

Country	MAPE	RMSPE	Interval	Country	MAPE	RMSPE	Interval
Arg	0.273682	0.3660809	2	Mex	0.1516197	0.2052339	1
Aus	0.166191	0.1953154	2	Net	0.1081521	0.1269	3
Bel	0.1574319	0.2303538	2	NZ	0.1482046	0.1743532	1
Bra	0.172607	0.192511	0	Pol	0.1590476	0.178894	2
Can	0.123498	0.1499399	0	Por	0.0913038	0.1097499	1
Fra	0.1793278	0.1935617	1	SA	0.2023648	0.2878304	1
Ger	0.1597241	0.2052984	2	SK	0.1052154	0.1346167	0
Gre	0.164795	0.1932559	0	Spa	0.09950459	0.1323683	2
Ind	0.235670	0.3209278	0	Tur	0.2627044	0.361667	1
Isr	0.162323	0.1862186	2	UK	0.123021	0.1314255	4
Jap	0.418187	0.4509111	5	USA	0.1258529	0.1559726	0

of the following equation

$$MAPE = \frac{1}{n}\sum\left(\frac{y_i - \hat{y}_i}{y_i}\right) \tag{5.1}$$

$$RMSPE = \sqrt{\frac{1}{n}\sum\left(\frac{y_i - \hat{y}_i}{y_i}\right)^2} \tag{5.2}$$

where n is the number of observations in the test set, y_i is the actual value at time i and $\hat{y}_i$ is the forecasted value at time i. Both measures are expressed in percentage terms, where a value of zero represents a perfect prediction. The former measure treats all the error in the same manner while the latter gives a greater weight to the extreme ones. I also computed the number of observations that fall outside the upper and lower bounds set with a confidence level of 95%. The results show a good model performance, with an average MAPE of 0.17 and an average RMSPE of 0.21. Furthermore, the average number of observations that fall outside the confidence interval is 1.45, very close to 1.65, which correspond to the 5% of 33 (number of observation in the test set). In the table are reported the results for all the series. In the figure below is reported in black the actual values of the total Italian inbound flows series, computed as the sum of all the country-specific series, in red is displayed the forecast, calculated as the sum of all the forecasted value of the country-specific series, and in green, the upper and lower bound, computed at the same manner. In appendix B are displayed the backtesting for the country-specific series.

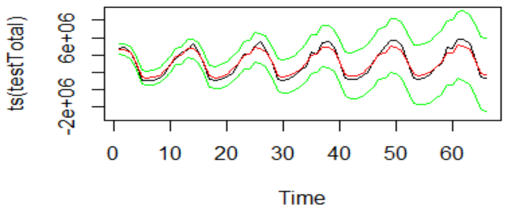


Figure 5.41: Total flow in Italy backtesting

5.4 Future Networks

5.4.1 Inbound Analysis

In figure 5.41-5.60 are reported the networks of the four quarters for the period 2020-2025. The size of Vertices is set on the basis of the in-degree values. The numeric results can be found in the tables in the appendix. It is possible to notice how the model predicts a general decline in the importance of Spain, in partiular in the fourth quarter of 2022, when it will lose the leadership in terms of visitors in favour of the USA. USA will maintain the first position in the first and third quarters as well as France in the second one. Italy affirms a constant centrality in the network, except for the first quarter when the Netherlands surpasses it. South American countries occupy the lasts positions in all the networks together with South Africa. Mexico is expected to increase its tourists flows, surpassing Portugal in the fourth quarter, Canada and Turkey in the third and Germany and Spain in the second quarter, becoming one of the leading countries in this period of the year. Japan is expected to experience a boost in its flows in the first quarter, overcoming the number of tourists of nine other countries, becoming the eighth most visited country. The model forecasts an increase of Australia importance in the last three months of the year. It is expected to become the fourteenth most visited destination, gaining eight positions. As regards European countries, the ones that are expected to increase their centrality are only The Netherlands and France but only in the first quarter, all other nations will remain stable or will lose some positions. The UK is a particular case because in the fourth and first quarters is expected to have a decline in its relative touristic strength, while in the second the opposite scenario is forecasted. The numeric results are reported in appendix A.

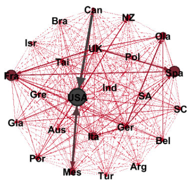


Figure 5.42: I-quarter 2020

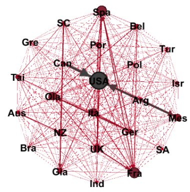


Figure 5.43: I-quarter 2021

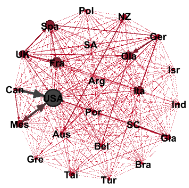


Figure 5.44: I-quarter 2022

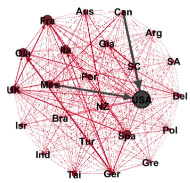


Figure 5.45: I-quarter 2023

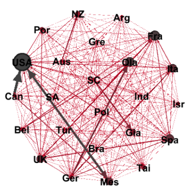


Figure 5.46: I-quarter 2024

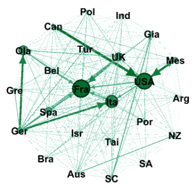


Figure 5.47: II-quarter 2020

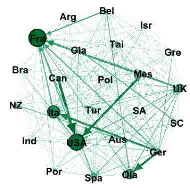


Figure 5.48: II-quarter 2021

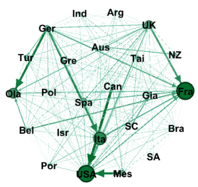


Figure 5.49: II-quarter 2022

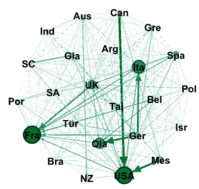


Figure 5.50: II-quarter 2023

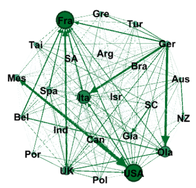


Figure 5.51: II-quarter 2024

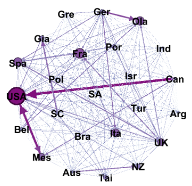


Figure 5.52: III-quarter 2020

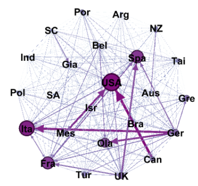


Figure 5.53: III-quarter 2021

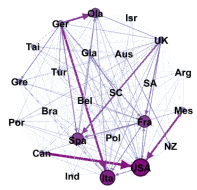


Figure 5.54: III-quarter 2022

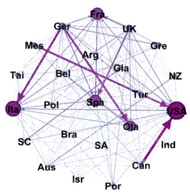


Figure 5.55: III-quarter 2023

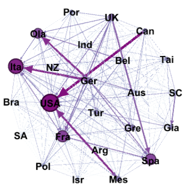


Figure 5.56: III-quarter 2024

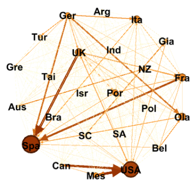


Figure 5.57: IV-quarter 2020

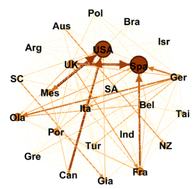


Figure 5.58: IV-quarter 2021

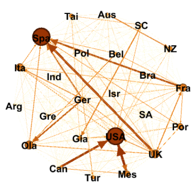


Figure 5.59: IV-quarter 2022

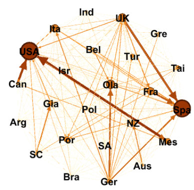


Figure 5.60: IV-quarter 2023

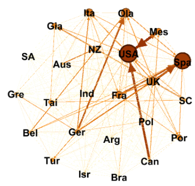


Figure 5.61: IV-quarter 2024

5.4.2 Outbound Analysis

In figure 5.61-5.80 are reported the networks of the four quarters of the period 2020-2025. The size of Vertices is set based on the outdegree values. Analyzing of the results shown in the tables in appendix A, it is possible to notice that the USA will maintain the first position as “exporter” of tourist in the first quarters of the overall period as well as Germany in the second and third ones. A change in the leadership can be found in the fourth qarters, when in 2022 Germany is expected to surpasse the UK. Italy is expected to will lose importance in the outbound tourist network in all the quarters. The model forecasts increasing importance of Asian countries, in particular Thailand in the second and third quarters and South Korea in the other two. Oceanian countries are expected to become more central in the network in the fourth quarter while a stable position is forecasted in the other three. South American countries are predicted to maintain their position at the end of the ranking, while Mexico will gain importance in the first three quarters of the year. As regards the other European countries, a slight general decline is expected, in particular in the first and last quarters.

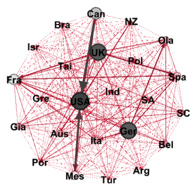


Figure 5.62: I-quarter 2020

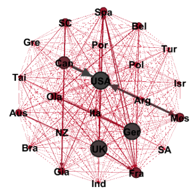


Figure 5.63: I-quarter 2021

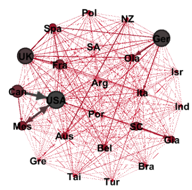


Figure 5.64: I-quarter 2022

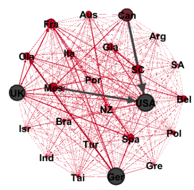


Figure 5.65: I-quarter 2023

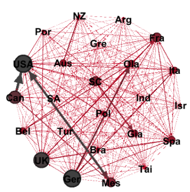


Figure 5.66: I-quarter 2024

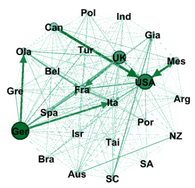


Figure 5.67: II-quarter 2020

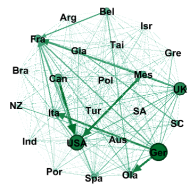


Figure 5.68: II-quarter 2021

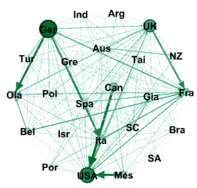


Figure 5.69: II-quarter 2022

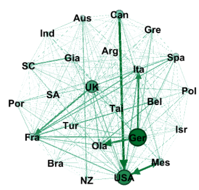


Figure 5.70: II-quarter 2023

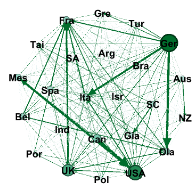


Figure 5.71: II-quarter 2024

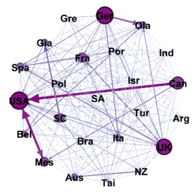


Figure 5.72: III-quarter 2020

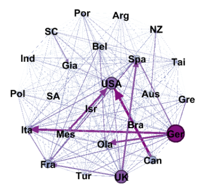


Figure 5.73: III-quarter 2021

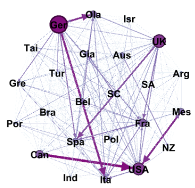


Figure 5.74: III-quarter 2022

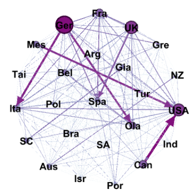


Figure 5.75: III-quarter 2023

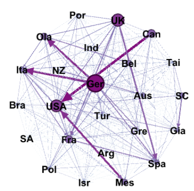


Figure 5.76: III-quarter 2024

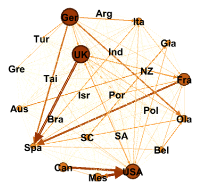


Figure 5.77: IV-quarter 2020

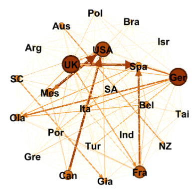


Figure 5.78: IV-quarter 2021

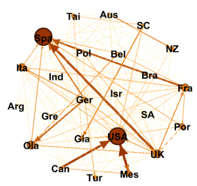


Figure 5.79: IV-quarter 2022

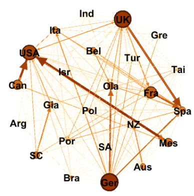


Figure 5.80: IV-quarter 2023

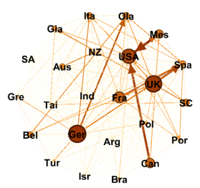


Figure 5.81: IV-quarter 2024

Chapter 6

Covid-19 Pandemic and its Implication on Tourist Flows

Covid-19 has become a global issue. The pandemic is expected to cause a decline in economic growth leading to one of the worst recession. Experts forecasts that the domestic product in developed countries will decline by 6.1 percent in average, while economic growth in emerging markets and developing economies will decrease by 1 percent. Furthermore, world trade is predicted to fall between 13% and 32% in 2020 because the Covid-19 pandemic disrupts normal economic activity and life throughout the world [171]. The tourism industry is one of the most affected, because the gouvernment responses to the disease imply human mobility restriction. To reduce the spread of this pandemic, all countries have imposed lockdowns, widely restricted international travel, bans on all foreign visitors, travel restrictions from various places with confirmed cases. In April 2020 the 93% of destinations in Europe have fully closed the border for international tourism. In America, this proportion reaches 82%, in Asia and the Pacific 77%, in the Middle East 70% and Africa 60%. The UNWTO (2020b) has reported that the COVID-19 pandemic has caused a 22% reduction in international tourist arrivals during the first quarter of 2020 and may decline by 60% -80% throughout 2020. In the past global tourism has been exposed to a wide range of crises Between 2000 and 2015, major disruptive events include the September 11 terrorist attacks (2001), the severe acute respiratory syndrome (SARS) outbreak (2003), the global economic crisis unfolding in 2008/2009, and the 2015 Middle East Respiratory Syndrome (MERS) outbreak. None of them led to a longer-term decline in the global development of tourism, suggesting that tourism as a system has been resilient to external shocks in the long period and a catastrophic one in the short [172]. For this reason the Covid-19 analysis implemented in this chapter consider only the current year. The chapter aims at forecasting the evolution of the tourist network taking into account the travel restriction and the forced quarantine imposed by national authorities. Obviously, these estimates need to be treated with extensive caution, as it remains fundamentally unclear how the pandemic will develop, and how the gouvernments will react in the future months. Furthermore, it is very difficult to estimate how the desease fear will cange the travellers behaviour. In the next two sections I am going to describe how I tryied to take into account these factors, in the third section I displayed the resulting networks for each quarters of 2020 and the last section I will draw some qualitative conclusions.

6.1 Border Restrictions and Expected Reopening

Tourist flows among countries are subordinated to the decision taken by the gouvernments about the measures to adopt at border. Not only the closed border leads to an annulment of the travellers number, but also the compulsory two weeks of quarantine has the same effect. Forecasting the period of reopening is a very difficult task, indeed neither the evolution of the disease nor the decision rule adopted by policymaker are known. I base the prediction primarily following the public information issued by the authorithies. The first countries that start to open borders are the european ones. On June Italy, Germany, Belgium, Portugal, Spain[1], France, Netherlands, Poland and Greece allows european and UK citizens to cross their borders without quarantine. No official programme has yet been issued with regard the reopening to non-Schengen countries. The Governor of the Tourism Authority of Thailand said that tourism could return in the fourth quarter of this year, so I considered the closure continues up to the first October 2020. Turkey has reopened on June for touristc purposes to 40 countries, among these nations, the ones included in the network are Greece, Germany, Japan, Poland, Netherlands, South Korea and Belgium. The United Kingdom reopened to international tourists in June, but all visitors must self-isolate for 14 days. This measure will strongly influence the choice of UK as a tourist destination, disincentivating the influx of tourists. US has not ever closed their territory except for Italy, China, Iran and Brazil. However it suspended tourism visa from some countries. The nationality included in the studied network that are allowed to enter in US territory are Australia, New Zealand, South Korea, Japan and Canada[2]. Mexico, as well as Turkey from June 12, does not have any entry restrictions, but travelers arriving from countries affected by COVID-19 will be screened and quarantined if necessary. Information on other countries are not yet available, so to forecast the border reopening I used a simple rule of thumb. I looked at the number of active

[1]Spain is expected to reopens the first of July
[2]By air only

cases the various country had at the moment of the border closing and I assumed that the reopening will occour in the first day of the month the number cases is expected to reach the same value. The reason behind this choice can be found in the fact that border restriction is a combination of various factors. The number of cases is only one of them, it must be combined with economic and diplomatic relationships and with the political line of governments. Checking when the restriction have started, I tried to capture all these factors. The assumtpion is that the more rapid was the border closing among countries, the less important are their relations and the more prudent the government policy are. On the contrary, the slower the frontier cross ban was apllied, the more important are the dyplomacy and the more reluctant is the government to containing measures. Since the curve of infected is bell-shaped, the greater is the number of infected that have been sarted the ban, the earlier will be reached the same number in the future, and the contraty is true. The projection of the number of cases is taken from the web-site `https://covid19-projections.com` of Youyang Gu, data scientist at the MIT. The prediciton is based on a SEIR model. This model divide the population at time t $N(t)$ in four subclasses, the susceptible individuals $S(t)$ (everyone who had not yet contracted the disease), the exposed population $E(t)$ (individuals who have contracted the virus and may infect others, but are still asymptomatic), the infected population $I(t)$ (exhibiting signs and symptoms of the illness) and the recovered one (in an oversimplified view, the number of individuals who can no longer infect others). To quickly summarize how an SEIR model works, at each time period, an individual in a population is in one of four states, and the model aim is to forecast the movement of individual among these classes [173]. The classic SEIR dynamics is described by means of the following system

$$\begin{cases} \frac{dS}{dt} = -\beta S(I + qE)/N \\ \frac{dS}{dt} = \beta S(I + qE)/N - \delta E \\ \frac{dI}{dt} = \delta E - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases} \tag{6.1}$$

The constant β is the product of the per capita contact rate and the probability of infection after contact with an infected individual. $1/\delta$ is the average time for a latent individual to become infectious, $1/\gamma$ is the average time it takes a person to die or recover once in the infectious stage, and q is a wheight applyed to the latent individual because of their less infectivity. Unlike the traditional approach, The parameters/inputs of this simulator are then learned using machine learning techniques that attempts to minimize the error between the projected outputs and the actual results. The predictions are made using daily data provided by Johns Hopkins CSSE, what is considered by experts to be the “gold standard” reference data. For further information about the model it is possible to consult the website

[174] or see the analysis published by Carl T. Bergstrom, Professor of Biology at the University of Washington [175]. looking at the prediciton, all European countries and UK, seems to have surpassed the worst period of the health emergency. The only exception is Poland, which is expected to face the peak in the end of August. The same scenario is expected for Oceanian countries and Turkey. South East Asian coutries, Mexico, South Africa, Brazil and Argentina are expected to face the peak in the next months. Japan, South Korea and Israel predictions, show the presence of a slight second wave of contagions in the third quarters, while a more evident and fat second waves is displayed in the predictions of the USA and Canadian outbreak.

6.2 Internal Restrictions

So far, I investigated the entry bans, but, for a more accurate analysis, an assesement of the exit ones is also needed. The pandemic has caused worldwide curfews and similar restrictions (stay-at-home orders, shelter-in-place orders, shutdowns/lockdowns) established to prevent further spread of COVID-19. The pandemic has resulted in the largest amount of shutdowns/lockdowns worldwide at the same time in history. By 26 March, 1.7 billion people worldwide were under some form of lockdown, which increased to 3.9 billion people by the first week of April, more than half of the world's population. The impossibility of leaving home, obviously, also affects the possibility of leaving the country for tourist purposes, and for this reason I took into account the quarantine measures. In Argentine the quarantine started on the 19th of March and ended the 28th of June while in Brazil, the president Jair Bolsonaro has never apllyed such kind of measure. In Mexico the quarantine measure started from 23rd of March to the 1st of June, while Canadian measure at national level were not applied. Japanese prime minister Shinzo Abe has been criticized for his mild response to the outbreak. Indeed in Japan a stay-at-home order was never applied, but only a closure of elementary school and some non-essential activity. A regional lockdown was applied in South Korea, but no-measures on a national level. In Thailand the curfews period sarted in 25th of March and ended on May the 31st, while in Indonesia there was a lack of such kind of measures. In South Africa, President Cyril Ramaphosa declared that would undergo a national lockdown, for a period of 21 days, from 26 March to 16 April 2020. This drastic measure was extended until 30th of April. European countries were the most affected by the pandemic, nevertheless, the gratest part of them surpassed the worst period. For this reason, almost all of them ended the stay-at-home order in the past months. Belgium was under quarantine for the period 18th of March-4th of May, France between 17th of March and the 11th of May, Germany from 23rd March to the 10th of May, Greece 23/3-4/5, Italy

9/3-18/5, Netherlands 15/3-6/5 [3], Poland 13/3-11/4, Portugal 19/3-2/4, Spain 14/3-9/5. Turkish prime minister ordered only five days of lockdown, from the 23rd to the 27th of April. New Zealand started the internal restrictions on the 26th of March and finished it in May the fifteenth, while Australian government, as well as British and Israelian ones, have not yet ceased lockdown. To forecast the end of it, I used the same method to predict the end of external bans. I checked the number of cases at the starting date of lockdown, and, by means of forecasted number of cases, I supposed to end when the same number will be reached. Merging the information collected in this and previous section, I constructed a matrix with the periods of absence of tourism connections. The matrix is reported in the table in appendix C.

6.3 Networks

In the figures below are reported the networks constructed taking into account the mobility limitations. From figure 6.1 to figure 6.4 the focus is on the inbound flows, the size of the nodes is set based on their indegree values, instead, in pictures 6.5-6.8 the focus is on the outflows, the size of the nodes is set based on their outdegree values. Looking at the images, it is possible to see evident changes among tourist flows networks with and without bans. Since the pandemic spread and the related bans in the countries in analysis started at the end of the first quarter, the networks related to the first part of the year are very similar. Italy is the country that suffered the most in this period, since it is the first one where both the inflows and outflows of visitors were forbidden. In the second quarter almost all countries closed their borders, and the greatest part kept the ban for the whole quarter. An exception are the countries of the European Union that decided to re-open their borders for the European citizens at the end of June. Italy and France are the ones that enjoied the most the end of the bans, since their flows are mostly dependent on Schengen countries such as UK and Germany. USA does not benefit of similar treaties, indeed its importance in the network drastically decrease. It is possible to observe the same scenario as regard the outflows. The USA almost disappeared from the network while Germany and UK became the leader source of visitor. The network density falls in the first two quarters from 0.833 to 0.065 as well as the average degree that falls of 93%. The third quarter is expected to report a slight recovery in international tourism movement. The density is expected to varies from 0.065 to 0.457 thanks to the reopening of the less affected destinations. Italy is expected to obtain the leadership surpassing Spain and USA, which number of cases are decreasing more slowly. Germany remains the

[3]Netherlands has never applied a true stay-at-home order, but all the commercial activity and the school remained closed for this period

most important "exporter" of travellers, while UK, which is expected to remain isolated from many countries due to its high level of contagion. The normality is quite reached in the fourth quarter, when the density is expected to achieve a value of 0.639. The USA restart to be one of the most important destinations as well as Spain, but countinues to be below normal levels as regard the outflows levels.

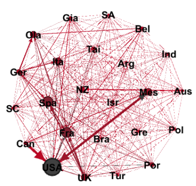


Figure 6.1: I-quarter 2020

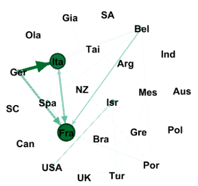


Figure 6.2: II-quarter 2020

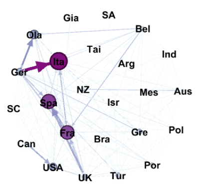


Figure 6.3: III-quarter 2020

6.4 Conclusion

As I outlined before, the spread of the pandemic and the connected policy decision are complicated to forecast, so the actual future architecture of the network can severely differ from the prediction presented in the previous part. Furthermore, in this chapter, I assumed that the only factor that affects tourist flows during an outbreak are the barriers to human mobility that the governments impose in order to contain the spread of the virus. I am aware that others variables

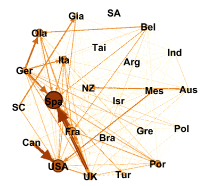


Figure 6.4: IV-quarter 2020

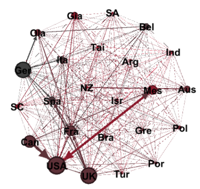


Figure 6.5: I-quarter 2020

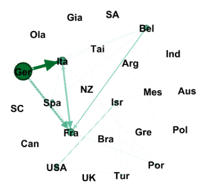


Figure 6.6: II-quarter 2020

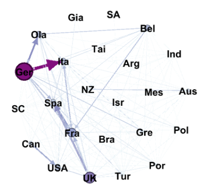


Figure 6.7: III-quarter

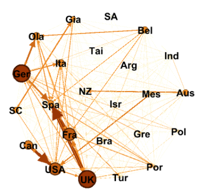


Figure 6.8: IV-quarter 2020

can influence the level of tourist flows since this industry is particularly prone to external changes [176]. The most obvious factor is the fear of contagion, that leads tourist to modify their holiday's plans, opting for domestic or closer destinations [177]. Also the typology of visited area is expected to vary, preferring remote places such as mountain and hill and avoiding cities or other visiting crowded tourist attractions [178]. The communication methods of both mass media and governments can affect tourist behaviour too, modifying the travellers' perception of the riskiness of a country [179]. An outbreak can also have an indirect effect on the tourism industry, for instance, the numerous lost of job places and the fall on the participation of labour market [180] can leads household to be less wealthy and, as a consequence, to spend less for travel purposes. Because of all this plethora of factors, a numerical forecast of the inbound and outbound flows can be severely biased. For this reason I opted only for an analysis of node centrality and network density. Indeed, while the aforementioned variables affect the whole global tourism industry, the entry bans and the number of cases are the only factors that can lead to a country-specific decrease in tourism level. For this reason, the accuracy of network forecasting and nowcasting should remain pretty much unaltereted without taking into account the others source of travellers behaviour modifications. Looking at the results, two things stand out. The first one is that countries involved in international agreements are the ones that will suffer the less the effects of the pandemic on tourism. Indeed, the countries that are expected to obtain the leadership both as "exporter" and "importer" of visitors, belong to the Schengen treaty. The second is that the countries that introduced a heavier internal measures (e.g. Italy), are the ones that are expected to recover more rapidly, while governments that opted for a softer line are the ones that are expected to return to pre-pandemic levels more slowly (e.g. USA). Furthermore, the countries that face the outbreak later, or even that are expected to reach the peak in the future are the ones that will suffer the most the effects on tourism. This inequality occurs because during the first period of the pandemic, all the countries imposed entry restriction, even to nations with few cases, while when the spread started to be under control, the border reopened only among states with a strong downward trend of cases. As regard the global network, the complete recovery is expected to be reached only in 2021, but at the end of 2020 the 77% of international tourist links will be restored.

Conclusions

The methodology proposed in this thesis can be a handy tool to predict how the importance of countries in term of touristic flows, will change, both during the year and in the long term. Understand how the network will evolve, can be used by firms to know where to invest and which will be the countries with the most significant touristic business opportunity. Furthermore, policy-makers can use the results to better manage the human flows, for instance knowing how to organize and adapt the routes and means of transport. Further research should concentrate on the determinants of touristic flows, introducing in the time series analysis some regression term, such as the per-capita GDP of departure country, the number of the touristic sites, the average temperature of the destination. Such an analysis allows understanding which are the factors that attract the more the visitors of a specific country, knowing where to act in order to modify the visitor behaviour, for instance for marketing purposes or to redistribute the number of tourist among a geographic area. Furthermore, future research should concentrate on using as many data as possible, searching to include the majority of world nations.

Appendix A

Tables of Results Chapter 4

Number of visitors in 2018 by destination (inbound tourists)

Quarter	Arg Gre Por	Aus Ind SA	Bel Isr SK	Bra Ita Spa	Can Mex Tahi	Fra NZ Tur	Ger Net UK	Jap Pol USA
I	640065 691290 2822307	1612049 965383 593981	1814539 654600 1257173	226191 5585970 12185458	2430549 4663029 2586748	8631687 3024056 1088485	3896082 5010368 5320618	2821978 745724 14549802
II	490079 2123224 1417971	923038 1081579 394214	1821439 775800 1434167	269714 12838760 10344056	3534703 3933193 2605839	17845035 1669423 3120624	5357355 7933315 7165061	2858850 1123361 16000380
III	544463 4668170 3154431	981810 1397235 505082	1823271 595400 1446435	296039 16910666 18256545	4962989 3406457 2657020	13504511 1892254 3855766	7047372 8483069 5679629	2342522 1467514 17943270
IV	668076 1283011 5330169	1420523 1076390 600994	1816409 804100 1587180	419785 6193012 22019085	2940219 3944819 2599815	7424590 2942065 2726909	4641379 5792248 6620199	2753213 930753 15600994

Number of visitors in 2019 by destination (inbound tourists)

Quarter	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
I	707423	1572493	1839496	616426	2187897	7753000	3978439	2903003
	811180	958238	720200	5698426	4640159	2903461	4639000	721327
	3309456	556563	1412037	11650652	2718095	1129193	5107103	13534986
II	566474	949082	1838731	494251	3358821	17811994	5487728	2859781
	2243435	1125352	844300	11945725	4084943	1734646	7814000	1127567
	1627919	413262	1694598	7196994	2726415	3270180	7293748	15420376
III	620041	1000063	1836147	377118	4507617	13129966	7153927	846643
	4814904	1470419	732800	15876113	3660250	1900499	8114000	1450104
	3406733	485463	1587934	14710850	2776335	3970814	5889224	17779574
IV	721938	1483402	1822794	268404	2724898	7196873	4758073	1198830
	1340611	1218283	883200	6175222	4175791	2876090	5810000	934013
	5753776	591481	1558719	18584855	2727273	2852364	6938919	15628551

Number of tourists in 2018 by country of origin (outbound tourists)

Quarter	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
I	872918	3028294	3231631	1486934	7021075	6210570	10146941	3540176
	483507	466466	559733	3926615	4814937	723908	3922902	1617896
	1115552	190754	3996810	3501505	572562	455686	11045400	10885358
II	1025373	3535960	5496052	1842616	7085999	7285571	16902410	3539294
	568452	615516	839708	5345727	5224226	622498	5489526	2130297
	1453766	305151	3781246	4489553	710172	583929	13351044	14837093
III	1039373	3978956	5755576	2043063	8179388	10084409	21242067	3956641
	661231	536734	1043360	6098946	5314488	660164	7303427	2429509
	1519818	281122	3722385	4741024	495496	648276	16773051	15313417
IV	924012	3656700	4217084	1931891	6677366	9562856	13617601	3558564
	604365	568507	663630	5089912	5585445	593587	4879863	1762563
	1553903	282673	3589385	3834901	708648	417260	15177658	12677566

Number of tourists in 2019 by country of origin (outbound tourists)

Quarter	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
I	1099137	2860641	3169710	1492559	6693584	6490201	9999048	3653970
	399932	397792	501615	3893059	4259403	673984	3788469	1240444
	800279	165883	3927567	3738329	550696	393424	10834415	11044111
II	1045618	3538055	5354988	1785280	6880631	6726936	16853634	3825664
	503234	474851	712470	5093207	4997145	611947	5283208	1864613
	1334192	296359	3533848	4711574	609735	562519	12591651	14738962
III	991741	3887018	5622860	2114684	8045919	9581079	20966467	3964994
	572988	397094	1024322	5514100	5628351	655945	7031329	2365415
	1403171	282170	2245078	4849046	394900	655964	15381060	14521839
IV	781396	3670956	4205416	1845585	6601744	9712132	13437966	3436411
	559285	484769	748715	4890891	5824064	656859	4818628	1790514
	1458201	283992	2105910	3910890	727389	458931	13946485	11867229

Number of visitors in the first quarters for the period 2020-2024 by destination (inbound tourists)

Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	767730	740287	1839430	213823	1899493	8386101	3014366	911774
	879773	1114230	821989	5807726	4847677	3182019	4990151	770675
	3151376	557800	1457576	8482536	2822127	1380173	5198008	14297674
2021	828234	1553969	1867649	229784	1623675	8295270	3151087	3162132
	921464	1179724	932346	5946105	5058832	3262958	6025544	872977
	3067292	550544	1502085	7864903	2929596	1350620	5401332	14510680
2022	884827	1595675	1886188	621566	1454192	8076118	2752499	3389048
	1011859	1229903	1037893	6106809	5258192	3318546	6404141	925120
	2894772	540595	1544839	7368445	3028582	1392575	5468279	14675775
2023	940792	1630406	1914258	206346	1308221	7939767	2481019	3601210
	1088304	1284615	1146804	6205404	5460318	3405622	6780141	975311
	2884044	531762	1601017	6821600	3128467	1392665	5609657	14887399
2024	996663	1664494	1940422	226079	1146848	7750513	2186440	3818990
	1170436	1348498	1259164	6315751	5655182	3467340	7136408	1019369
	2890339	524723	1654534	6290534	3225689	1412402	5715805	15119228

Number of visitors in the second quarters for the period 2020-2024 by destination (inbound tourists)

Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	628332	1013748	1845230	268954	2935539	17620324	4915206	3087441
	2273812	1246139	967695	12129409	4289898	2232863	8769914	1224623
	1737631	393480	1571169	6208703	2832874	3279666	7429359	16080025
2021	687717	1049890	1872529	272268	2796280	17580589	4536585	3317579
	2337942	1304702	1075315	12265546	4503008	2291637	9150614	1281509
	1830266	384944	1614252	5648150	2940526	3274930	7586143	16271724
2022	743400	1087357	1886237	493401	2645105	17336459	4109459	3536330
	2413135	1361179	1186602	12365320	4703215	2359025	9507262	1333528
	1900715	377789	1669126	5170606	3039432	3299905	7680342	16511481
2023	798526	1126413	1919365	258082	2503941	17209540	3854051	3749445
	2487078	1416824	1294833	12437970	4899232	2445424	9883800	1382012
	1955725	368551	1718937	4669084	3136038	3308205	7792830	16700869
2024	854333	1158085	1941199	270572	2348905	17015794	3565669	3978803
	2497583	1477821	1405530	12566443	5098566	2504568	10264650	1429153
	2060293	360260	1771433	4182970	3232734	3310246	7918398	16904170

Number of visitors in the third quarters for the period 2020-2024 by destination (inbound tourists)

Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	679431	1114158	1848767	334732	4227820	13015535	6386467	2935969
	4832271	1553532	838743	15988682	3864780	2451889	9141324	1567416
	3195691	481668	1529278	14082719	2881503	3963120	6005711	18597330
2021	738768	1156648	1873325	298063	4065896	12912102	6007303	3158025
	4950720	1611034	947571	16187493	4078333	2516590	9514076	1621314
	3063523	473357	1576468	13342797	2988739	3962036	6114502	18798357
2022	794905	1194418	1890773	378676	3920276	12701017	5616789	3387092
	4950103	1665337	1057946	16273117	4272934	2583947	9899166	1674097
	2938743	465491	1632295	12828574	3085422	4001211	6250833	19022728
2023	849086	1240871	1920757	320316	3778727	12526621	5395651	3599381
	4963546	1726337	1166649	16326659	4466150	2673081	10266299	1720868
	2720038	456942	1682014	12204585	3181243	4001195	6382569	19231753
2024	904043	1264593	1938543	306384	3620192	12362944	5095352	3830683
	5052364	1778853	1273609	16421872	4657032	2729375	10657010	1767683
	2742694	446490	1732634	11667084	3276942	3999435	6485566	19402489

Number of visitors in the fourth quarters for the period 2020-2024 by destination (inbound tourists)

Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	781173	1549751	1844266	276372	2217842	7053274	4310516	3211954
	1356512	1276659	997416	6314180	4380635	3208531	6712182	1033885
	5521401	583465	1608399	18074279	2833236	2835237	7092858	16203322
2021	838766	1583874	1865466	413108	2036113	6880467	3946589	3431067
	1462115	1332772	1105306	6493580	4584416	3268176	7100247	1090438
	5116594	574160	1658599	17465116	2934784	2838905	7177959	16400517
2022	894523	1623613	1890330	262868	1900380	6734528	3575043	3658546
	1514994	1386497	1216851	6607252	4781251	3350925	7480290	1141793
	4853722	566901	1716946	16924500	3033983	2861374	7319784	16641563
2023	948997	1659137	1916943	264900	1745550	6573517	3350061	3866061
	1576500	1450464	1326469	6671822	4979655	3420459	7823892	1186046
	4634036	558329	1763542	16369392	3131181	2879669	7421973	16833045
2024	1003503	1685882	1931468	414901	1609031	6384851	3016206	4101749
	1663457	1501998	1435377	6679310	5165323	3495839	8228244	1234993
	4408086	548788	1821834	15835166	3223023	2887134	7523844	17044552

Number of tourists in the first quarters for the period 2020-2024 by country of origin (outbound tourists)

Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	712412	3066499	3125079	1432387	6862570	5945758	9674740	3393083
	519547	477658	594932	3332410	4777325	205701	3517082	1722705
	1057485	185993	2095119	3655494	657516	429714	9744230	10349072
2021	736488	3224727	3217654	1454517	7073355	5813816	9852405	3582520
	557824	590046	667977	3316954	4893130	647634	3793152	1871591
	1146985	219507	4394353	3745252	736542	466852	9738245	10347278
2022	1104984	3271400	3260924	1445703	7184197	5582756	9979749	3534324
	570776	625388	680944	3259616	5006662	662405	3744175	1928220
	1190144	229088	4648392	3773879	788644	462361	9525638	10406071
2023	729053	3328405	3314496	1453220	7308740	5409182	10155127	3517395
	585261	661717	683551	3204823	5137440	675334	3735531	1995961
	1229092	240143	4896493	3838080	840337	461847	9371707	10452214
2024	759310	3386682	3357766	1466643	7438713	5228876	10331196	3489710
	598768	697829	684906	3148185	5268988	686592	3752793	2059958
	1271235	251603	5144964	3890487	893506	457398	9204380	10465361

Number of tourists in the second quarters for the period 2020-2024 by country of origin (outbound tourists)

Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	837575	3741648	5392163	1774226	7220112	6531430	16700976	3753970
	640642	670400	855268	4756002	5249264	628253	5367096	2173518
	1387417	328173	4094793	4696937	814991	599962	12195470	14571748
2021	848579	3794603	5445088	1782822	7367124	6402173	16914815	3733126
	655471	706371	861323	4711606	5366959	640362	5390816	2246678
	1442742	339832	4351373	4770833	868406	602483	12021416	14609643
2022	1054475	3846002	5482441	1781310	7491578	6280873	17095491	3700062
	668549	742868	870843	4628210	5504512	654142	5377016	2307772
	1482394	349485	4606434	4805196	920834	601984	11842670	14621268
2023	852315	3903989	5533178	1770251	7619037	6141829	17279039	3682866
	682513	778020	870562	4586444	5626669	668961	5419633	2371259
	1524604	360800	4850747	4856500	971709	599478	11687814	14678559
2024	875384	3971431	5582785	1788440	7750538	6025698	17479492	3649180
	695420	813571	871409	4531549	5751244	678311	5462625	2429139
	1565825	371642	5108076	4911100	1026104	589172	11500365	14689678

Number of tourists in the third quarters for the period 2020-2024 by country of origin (outbound tourists)

Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	973714	4052543	5648052	2096698	8469614	9266652	20933738	4040667
	707000	605280	1101229	5363101	5639653	684226	7116346	2459908
	1482970	302623	4362389	4804532	651845	690605	15576594	14488558
2021	928837	4105639	5673476	2106268	8606126	9049466	21019423	4008533
	721002	642052	1106713	5308634	5765246	700306	7091548	2531504
	1529060	314359	4612545	4869073	705038	696542	15396707	14468947
2022	1014220	4158194	5715734	2110294	8736538	8847996	21107697	3975993
	735206	678527	1111016	5243218	5898928	714453	7079690	2587987
	1573796	324827	4874175	4913280	758780	689112	15225321	14410906
2023	985241	4219355	5765727	2089767	8844404	8640778	21128741	3954809
	749550	714226	1107786	5187026	6026603	734142	7123867	2652514
	1610874	335840	5121221	4955612	809984	682586	15024059	14336627
2024	958464	4266669	5814893	2108557	8962980	8510848	21252777	3919881
	761394	748812	1103510	5145556	6140902	737523	7169258	2710654
	1652876	346351	5374579	5014078	864236	677666	14859879	14311522

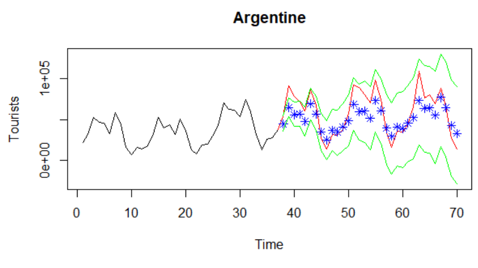
Number of tourists in the fourth quarters for the period 2020-2024 by country of origin (outbound tourists)

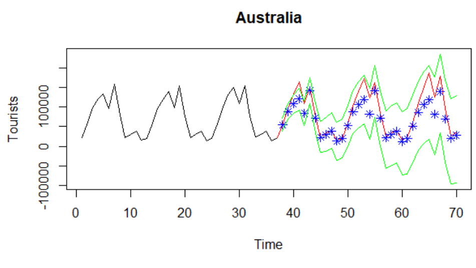
Year	Arg	Aus	Bel	Bra	Can	Fra	Ger	Jap
	Gre	Ind	Isr	Ita	Mex	NZ	Net	Pol
	Por	SA	SK	Spa	Tahi	Tur	UK	USA
2020	780744	3787605	4240133	1852750	6870793	9482209	13566268	3508971
	657181	650887	807535	4690868	5959744	679657	5008823	1910966
	1542812	302379	4177267	3915266	863590	481578	13852097	11687224
2021	871165	3818749	4270208	1844947	6998962	9252753	13648714	3459887
	664928	686942	819797	4547049	6082411	689806	4976513	1971547
	1597173	313663	4425481	3956342	915504	479543	13648922	11658130
2022	810702	3868609	4294170	1842371	7127462	9064992	13720894	3430832
	674613	723632	826257	4411914	6219887	705565	4935840	2038493
	1632591	323828	4687218	4009962	969343	478173	13510827	11630283
2023	797755	3926123	4324026	1849953	7248926	8883565	13807497	3408838
	683892	758445	827776	4287510	6340861	717886	4964606	2105665
	1675057	334576	4929024	4069350	1019546	475140	13351767	11563856
2024	901932	3975965	4355476	1860377	7377537	8682102	13908854	3372879
	691241	793900	827519	4176261	6466474	724076	4963920	2159629
	1724463	345991	5191003	4120671	1075261	469502	13194440	11485087

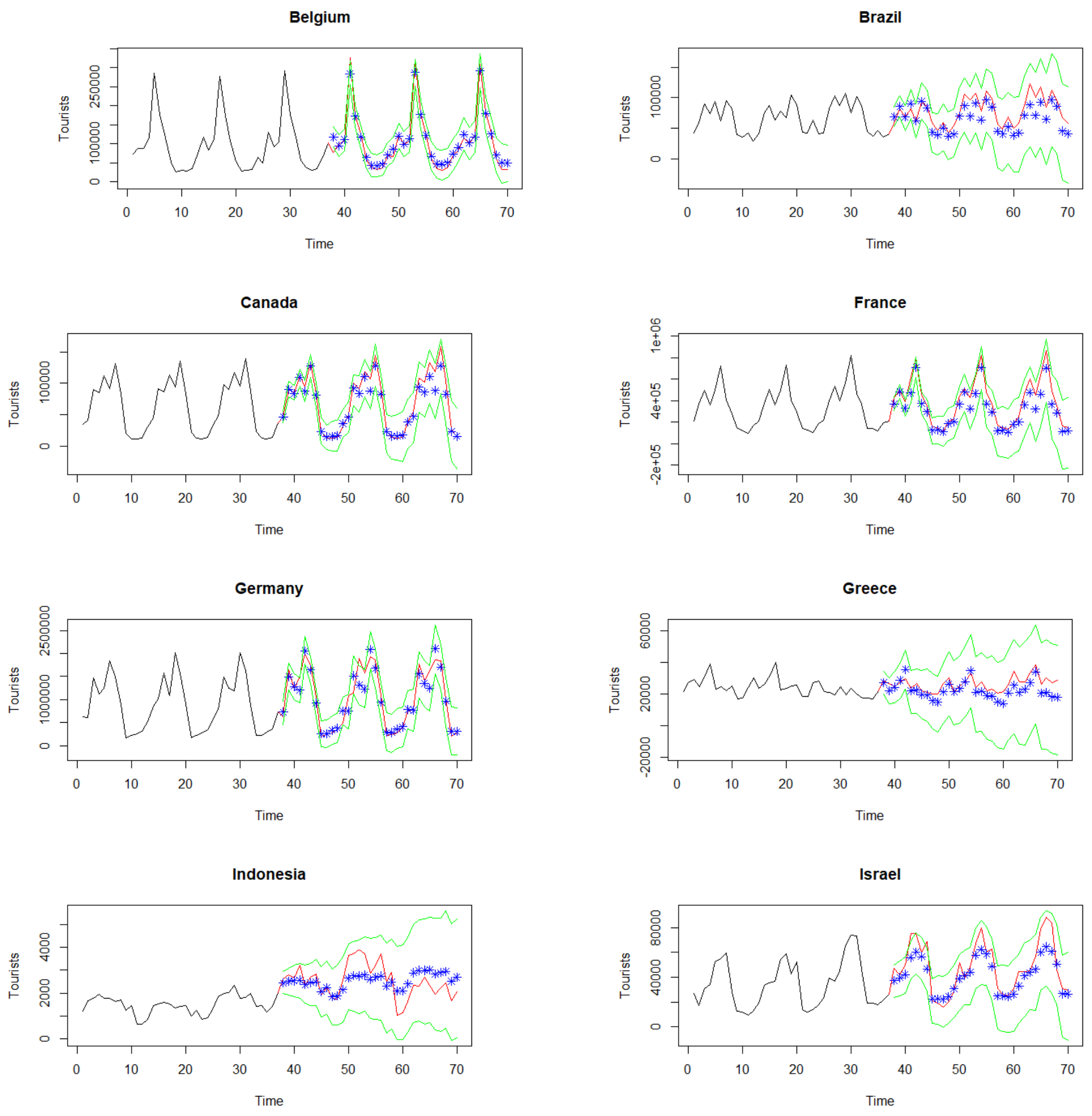
Appendix B

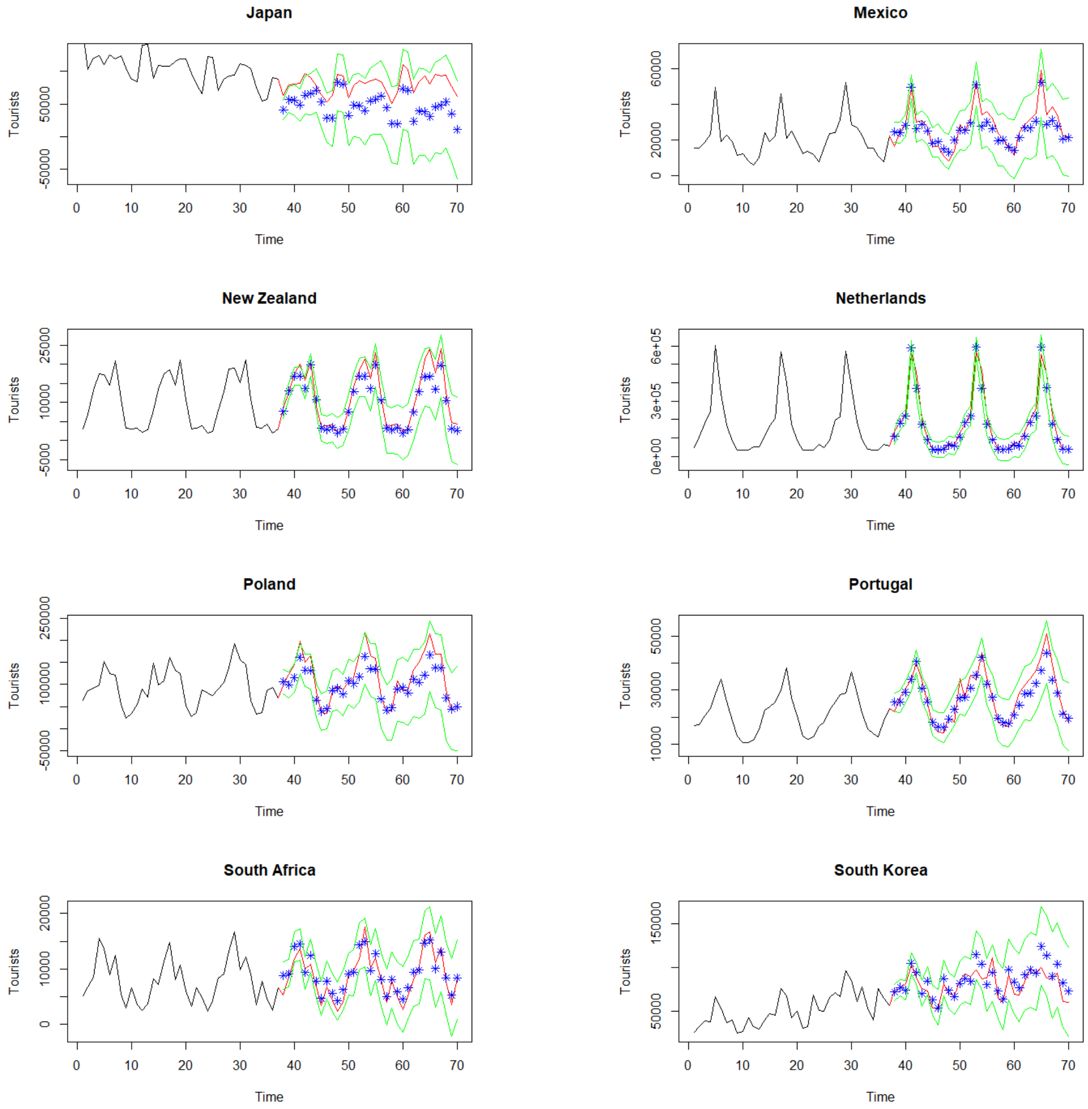
Backtesting Time-Series

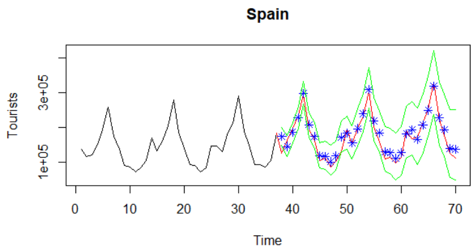
In this appendix are displayed the results of the backtesting presented in section 5.3. The time-series start from March 2013 and finish in December 2018 (In the pictures the two dates correspond to 0 and 70). The black lines represent the data included in the training set, the green ones are the upper and lower bounds of the confidence interval with a significance level of 5%, the red ones represent the actual values of the series (test set) while the blue star points represent the predictions. Each time series is the graphical representation of the inbound tourist in Italy, divided by country of origin. The series of Thailandese tourist that visited Italy is not shown because of lack of data, that does not allow me to implement the backtesting.

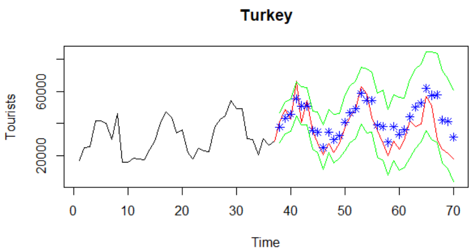


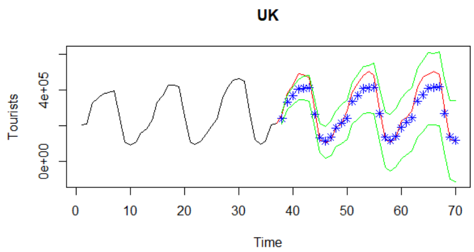


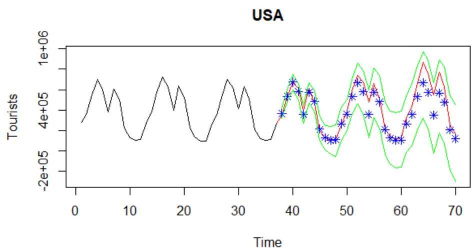

Belgium
Brazil
Canada
France
Germany
Greece
Indonesia
Israel
Tourists
Time


Japan
Tourists
Time
Mexico
Tourists
Time
New Zealand
Tourists
Time
Netherlands
Tourists
Time
Poland
Tourists
Time
Portugal
Tourists
Time
South Africa
Tourists
Time
South Korea
Tourists
Time


Spain
Tourists
1e+05
3e+05
0
10
20
30
40
50
60
70
Time


Turkey
Tourists
20000
60000
0
10
20
30
40
50
60
70
Time


UK
Tourists
0e+00
4e+05
0
10
20
30
40
50
60
70
Time


USA
Tourists
-2e+05
4e+05
1e+06
0
10
20
30
40
50
60
70
Time

Appendix C

Restriction-Period Table

	Arg	Aus	Bel	Bra	Can	Fra	Ger	Gia	Gre	Isr	Ind	Ita
Arg		17 Mar 01 Aug	16 Mar 01 Aug	17 Mar 01 Dec	17 Mar 01 Nov	16 Mar 01 Aug	16 Mar 01 Aug	16 Mar 01 Oct	16 Mar 01 Aug	17 Mar 01 Sep	17 Mar 01 Jan	09 Mar 01 Aug
Aus	19 Mar 01 Jan		17 Mar 01 Aug	20 Mar 01 Dec	20 Mar 01 Sep	17 Mar 01 Aug	20 Mar 01 Aug	20 Mar 01 Oct	20 Mar 01 Aug	20 Mar 01 Sep	20 Mar 01 Jan	09 Mar 01 Aug
Bel	19 Mar 01 Jan	17 Mar 01 Aug		17 Mar 01 Dec	17 Mar 01 Sep	17 Mar 15 Jun	17 Mar 15 Jun	17 Mar 01 Oct	17 Mar 15 Jun	17 Mar 01 Sep	17 Mar 01 Jan	09 Mar 15 Jun
Bra	19 Mar 01 Jan	23 Mar 01 Aug	17 Mar 01 Aug		30 Mar 01 Aug	17 Mar 01 Aug	23 Mar 01 Aug	25 Mar 01 Oct	23 Mar 01 Aug	30 Mar 01 Sep	30 Mar 01 Jan	09 Mar 01 Aug
Can	19 Mar 01 Jan	18 Mar 01 Aug	17 Mar 01 Aug	18 Mar 01 Dec		17 Mar 01 Aug	18 Mar 01 Aug	18 Mar 01 Oct	18 Mar 01 Aug	18 Mar 01 Sep	18 Mar 01 Jan	09 Mar 01 Aug
Fra	17 Mar 01 Jan	17 Mar 01 Aug	17 Mar 01 Jun	17 Mar 01 Dec	17 Mar 01 Sep		17 Mar 01 Jun	17 Mar 01 Oct	17 Mar 01 Jun	17 Mar 01 Sep	17 Mar 01 Jan	09 Mar 01 Jun
Ger	17 Mar 01 Jan	19 Mar 01 Aug	17 Mar 03 Jun	19 Mar 01 Dec	19 Mar 01 Sep	17 Mar 03 Jun		19 Mar 01 Oct	23 Mar 03 Jun	19 Mar 01 Sep	19 Mar 01 Jan	09 Mar 03 Jun
Gia	17 Mar 01 Jan	23 Mar 01 Aug	17 Mar 01 Aug	03 Apr 01 Nov	03 Apr 01 Aug	17 Mar 01 Aug	23 Mar 01 Aug		23 Mar 01 Aug	03 Apr 01 Sep	28 Mar 01 Jan	09 Mar 01 Aug
Gre	19 Mar 01 Jan	19 Mar 01 Aug	17 Mar 01 Aug	19 Mar 01 Dec	19 Mar 01 Sep	17 Mar 01 Aug	19 Mar 01 Aug	19 Mar 01 Oct		19 Mar 01 Sep	19 Mar 01 Jan	09 Mar 01 Aug
Isr	13 Mar 01 Jan	13 Mar 01 Aug	13 Mar 01 Aug	13 Mar 01 Dec	13 Mar 01 Nov	12 Mar 01 Aug	12 Mar 01 Aug	12 Mar 01 Oct	13 Mar 01 Aug		13 Mar 01 Jan	09 Mar 01 Aug
Ind	19 Mar 01 Jan	23 Mar 01 Aug	17 Mar 01 Aug	01 Apr 01 Nov	01 Apr 01 Aug	17 Mar 01 Aug	19 Mar 01 Aug	01 Apr 01 Oct	23 Mar 01 Aug	01 Apr 01 Sep		09 Mar 01 Aug
Ita	09 Mar 01 Jan	09 Mar 01 Aug	09 Mar 03 Jun	09 Mar 01 Nov	09 Mar 01 Aug	09 Mar 03 Jun	09 Mar 03 Jun	09 Mar 01 Oct	09 Mar 03 Jun	09 Mar 01 Sep	09 Mar 01 Jan	
Mes	19 Mar 28 Jun	23 Mar 01 Aug	17 Mar 01 Jun	23 Mar 01 Jun	23 Mar 01 Jun	17 Mar 01 Jun	23 Mar 01 Jun	23 Mar 01 Jun	23 Mar 01 Jun	23 Mar 01 Aug	23 Mar 01 Jun	09 Mar 01 Jun
NZ	15 Mar 01 Jan	23 Mar 01 Aug	15 Mar 01 Aug	15 Mar 01 Dec	15 Mar 01 Nov	15 Mar 01 Aug	15 Mar 01 Aug	15 Mar 01 Oct	15 Mar 01 Aug	15 Mar 01 Sep	15 Mar 01 Jan	09 Mar 01 Aug
Ola	15 Mar 01 Jan	15 Mar 01 Aug	15 Mar 01 Aug	15 Mar 01 Dec	15 Mar 01 Sep	15 Mar 01 Aug	15 Mar 01 Aug	15 Mar 01 Oct	15 Mar 01 Aug	15 Mar 01 Sep	15 Mar 01 Jan	09 Mar 01 Aug
Pol	13 Mar 01 Jan	13 Mar 01 Aug	13 Mar 01 Aug	13 Mar 01 Dec	13 Mar 01 Nov	13 Mar 01 Aug	13 Mar 01 Aug	13 Mar 01 Oct	13 Mar 01 Aug	13 Mar 01 Sep	13 Mar 01 Jan	13 Mar 01 Aug
Por	19 Mar 01 Jan	19 Mar 01 Aug	17 Mar 17 Jun	19 Mar 01 Oct	19 Mar 01 Sep	17 Mar 17 Jun	19 Mar 17 Jun	19 Mar 01 Oct	19 Mar 17 Jun	19 Mar 01 Sep	19 Mar 01 Jan	09 Mar 17 Jun
SA	19 Mar 01 Jan	23 Mar 01 Aug	17 Mar 01 Aug	26 Mar 01 Nov	26 Mar 01 Aug	17 Mar 01 Aug	19 Mar 01 Aug	26 Mar 01 Oct	23 Mar 01 Aug	26 Mar 01 Sep	26 Mar 01 Jan	09 Mar 01 Aug
SC	19 Mar 01 Jan	19 Mar 01 Aug	17 Mar 01 Aug	19 Mar 01 Dec	19 Mar 01 Sep	16 Mar 01 Aug	16 Mar 01 Aug	16 Mar 01 Oct	18 Mar 01 Aug	19 Mar 01 Sep	19 Mar 01 Jan	09 Mar 01 Aug
Spa	14 Mar 01 Jan	14 Mar 01 Aug	14 Mar 21 Jun	14 Mar 01 Dec	14 Mar 01 Sep	14 Mar 21 Jun	14 Mar 21 Jun	14 Mar 01 Oct	14 Mar 21 Jun	14 Mar 01 Sep	14 Mar 01 Jan	09 Mar 21 Jun
Tai	19 Mar 01 Oct	23 Mar 01 Oct	17 Mar 01 Oct	25 Mar 01 Oct	25 Mar 01 Oct	17 Mar 01 Oct	23 Mar 01 Oct	25 Mar 01 Oct	23 Mar 01 Oct	25 Mar 01 Oct	25 Mar 01 Oct	09 Mar 01 Oct
Tur	19 Mar 28 Jun	23 Mar 01 Aug	14 Mar 12 Jun	23 Apr 12 Jun	23 Apr 12 Jun	14 Mar 12 Jun	14 Mar 12 Jun	23 Apr 12 Jun	23 Mar 12 Jun	23 Apr 01 Aug	23 Apr 12 Jun	09 Mar 12 Jun
UK	19 Mar 01 Jan	23 Mar 01 Aug	17 Mar 01 Aug	23 Mar 01 Dec	23 Mar 01 Aug	17 Mar 01 Aug	23 Mar 01 Aug	23 Mar 01 Oct	23 Mar 01 Aug	23 Mar 01 Sep	23 Mar 01 Jan	09 Mar 01 Aug
USA	19 Mar 01 Jan	23 Mar 01 Aug	13 Mar 01 Aug	27 May 01 Aug	20 Mar 31 Jul	13 Mar 01 Aug	13 Mar 01 Aug	23 Apr 12 Jun	13 Mar 01 Aug	02 Apr 01 Sep	No Restriction	09 Mar 01 Aug

	Mes	NZ	Ola	Pol	Por	SA	SC	Spa	Tai	Tur	UK	USA
Arg	17 Mar 01 Dec	17 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	16 Mar 01 Aug	17 Mar 01 Jan	16 Mar 01 Aug	14 Mar 01 Aug	17 Mar 01 Jan	17 Mar 01 Aug	16 Mar 01 Aug	16 Mar 01 Nov
Aus	20 Mar 01 Dec	20 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	19 Mar 01 Aug	20 Mar 01 Jan	05 Mar 01 Aug	14 Mar 01 Aug	20 Mar 01 Jan	20 Mar 01 Aug	20 Mar 01 Aug	20 Mar 01 Sep
Bel	17 Mar 01 Dec	17 Mar 01 Aug	15 Mar 15 Jun	13 Mar 15 Jun	17 Mar 15 Jun	17 Mar 01 Jan	17 Mar 01 Aug	14 Mar 15 Jun	17 Mar 01 Jan	17 Mar 01 Aug	17 Mar 01 Aug	17 Mar 01 Sep
Bra	23 Mar 01 Dec	26 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	19 Mar 01 Aug	26 Mar 01 Jan	25 Mar 01 Aug	14 Mar 01 Aug	25 Mar 01 Jan	30 Mar 01 Aug	25 Mar 01 Aug	25 Mar 01 Sep
Can	18 Mar 01 Dec	18 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	18 Mar 01 Aug	18 Mar 01 Jan	18 Mar 01 Aug	14 Mar 01 Aug	18 Mar 01 Jan	18 Mar 01 Aug	18 Mar 01 Aug	04 May 01 Sep
Fra	17 Mar 01 Dec	17 Mar 01 Aug	15 Mar 01 Jun	13 Mar 01 Jun	17 Mar 01 Jun	17 Mar 01 Jan	17 Mar 01 Aug	14 Mar 01 Jun	17 Mar 01 Jan	17 Mar 01 Aug	17 Mar 01 Aug	17 Mar 01 Sep
Ger	19 Mar 01 Dec	19 Mar 01 Aug	15 Mar 03 Jun	13 Mar 03 Jun	19 Mar 03 Jun	19 Mar 01 Jan	19 Mar 01 Aug	14 Mar 01 Aug	19 Mar 01 Jan	19 Mar 01 Aug	23 Mar 01 Aug	19 Mar 01 Sep
Gia	15 May 01 Dec	26 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	19 Mar 01 Aug	26 Mar 01 Jan	12 Mar 01 Aug	14 Mar 01 Aug	25 Mar 01 Jan	03 Apr 01 Aug	25 Mar 01 Aug	25 Mar 01 Sep
Gre	19 Mar 01 Dec	19 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	19 Mar 01 Aug	19 Mar 01 Jan	19 Mar 01 Aug	14 Mar 01 Aug	19 Mar 01 Jan	19 Mar 01 Aug	19 Mar 01 Aug	19 Mar 01 Sep
Isr	13 Mar 01 Dec	13 Mar 01 Aug	13 Mar 01 Aug	13 Mar 01 Jan	13 Mar 01 Aug	13 Mar 01 Jan	13 Mar 01 Aug	14 Mar 01 Aug	12 Mar 01 Jan	13 Mar 01 Aug	13 Mar 01 Aug	13 Mar 01 Nov
Ind	23 Mar 01 Dec	26 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	19 Mar 01 Aug	26 Mar 01 Jan	12 Mar 01 Aug	14 Mar 01 Aug	25 Mar 01 Jan	01 Apr 01 Aug	19 Mar 01 Aug	01 Apr 01 Sep
Ita	09 Mar 01 Dec	09 Mar 01 Aug	09 Mar 03 Jun	09 Mar 03 Jun	09 Mar 03 Jun	09 Mar 01 Jan	09 Mar 01 Aug	09 Mar 03 Jun	09 Mar 01 Jan	09 Mar 01 Aug	09 Mar 01 Aug	09 Mar 01 Sep
Mes		23 Mar 01 Jun	15 Mar 01 Jun	13 Mar 01 Jun	19 Mar 01 Jun	23 Mar 01 Jun	23 Mar 01 Jun	14 Mar 01 Jun	23 Mar 01 Jun	23 Mar 01 Jun	23 Mar 01 Aug	23 Mar 01 Jun
NZ	15 Mar 01 Dec		15 Mar 01 Aug	13 Mar 01 Jan	15 Mar 01 Aug	15 Mar 01 Jan	13 Mar 01 Aug	14 Mar 01 Aug	15 Mar 01 Jan	15 Mar 01 Aug	15 Mar 01 Aug	15 Mar 01 Nov
Ola	15 Mar 01 Dec	15 Mar 01 Aug		13 Mar 01 Jan	15 Mar 01 Aug	15 Mar 01 Jan	15 Mar 01 Aug	14 Mar 01 Aug	15 Mar 01 Jan	15 Mar 01 Aug	15 Mar 01 Aug	15 Mar 01 Sep
Pol	13 Mar 01 Dec	13 Mar 01 Aug	13 Mar 01 Aug		13 Mar 01 Aug	13 Mar 01 Jan	13 Mar 01 Aug	13 Mar 01 Aug	13 Mar 01 Jan	13 Mar 01 Aug	13 Mar 01 Aug	13 Mar 01 Sep
Por	19 Mar 01 Dec	19 Mar 01 Aug	15 Mar 17 Jun	13 Mar 17 Jun		19 Mar 01 Jan	19 Mar 01 Aug	14 Mar 01 Aug	19 Mar 01 Jan	19 Mar 01 Aug	19 Mar 01 Aug	19 Mar 01 Sep
SA	26 Mar 01 Dec	26 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	19 Mar 01 Aug		19 Mar 01 Aug	14 Mar 01 Aug	25 Mar 01 Jan	26 Mar 01 Aug	19 Mar 01 Aug	19 Mar 01 Sep
SC	19 Mar 01 Dec	19 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	18 Mar 01 Aug	19 Mar 01 Jan		14 Mar 01 Aug	19 Mar 01 Jan	19 Mar 01 Aug	16 Mar 01 Aug	19 Mar 01 Sep
Spa	14 Mar 01 Dec	14 Mar 01 Aug	14 Mar 21 Jun	14 Mar 21 Jun	14 Mar 21 Jun	14 Mar 01 Jan	14 Mar 01 Aug		14 Mar 01 Jan	14 Mar 01 Aug	14 Mar 01 Aug	14 Mar 01 Sep
Tai	25 Mar 01 Oct	25 Mar 01 Oct	15 Mar 01 Oct	13 Mar 01 Oct	19 Mar 01 Oct	25 Mar 01 Oct	12 Mar 01 Oct	14 Mar 01 Oct		25 Mar 01 Oct	25 Mar 01 Oct	25 Mar 01 Oct
Tur	23 Mar 12 Jun	26 Mar 12 Jun	15 Mar 12 Jun	13 Mar 12 Jun	19 Mar 12 Jun	26 Mar 12 Jun	12 Mar 12 Jun	14 Mar 12 Jun	25 Mar 12 Jun		23 Apr 01 Aug	23 Apr 12 Jun
UK	23 Mar 01 Dec	23 Mar 01 Aug	15 Mar 01 Aug	13 Mar 01 Jan	19 Mar 01 Aug	23 Mar 01 Jan	23 Mar 01 Aug	14 Mar 01 Aug	25 Mar 01 Jan	23 Mar 01 Aug		23 Mar 01 Aug
USA	23 Mar 01 Dec	26 Mar 14 May	13 Mar 01 Aug	13 Mar 01 Jan	13 Mar 01 Aug	26 Mar 01 Jan	23 Mar 01 Aug	13 Mar 01 Aug	25 Mar 01 Jan	23 Apr 01 Aug	17 Mar 01 Aug	

Bibliography

[1] M. Emirbayer, "Manifesto for a relational sociology," *American Journal of Sociology*, vol. 103, no. 2, pp. 281–317, 1997.

[2] G. Featherstone, D. Simmel, and M. Frisby, *Simmel on culture: selected writings*. Sage, 1997.

[3] A. Marin and B. Wellman, "Social network analysis: An introduction," in *The SAGE handbook of social network analysis*, Sage, 2011, pp. 217–318.

[4] P. J. Carrington, J. Scott, and S. Wasserman, *Models and methods in social network analysis*. Cambridge University Press, 2005.

[5] N. Crossley, "The social world of the network. combining qualitative and quantitative elements in social network analysis," *Sociologica*, vol. 4, no. 1, 2010, 'doi: 10.2383/32049'.

[6] R. J. Trudeau, *Introduction to graph theory*. Courier Corporation, 2013.

[7] P. Fletcher, H. B. Hoyle, and C. W. Patty, *Foundations of discrete mathematics*. Brooks/Cole Publishing Co., 1990.

[8] A. Bavelas, "Communication patterns in task-oriented groups," *The Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 725–730, 1950.

[9] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, pp. 35–41, 1977.

[10] E. Landau, "Zur relativen wertbemessung der turnierresultate," *Deutsches Wochenschach*, vol. 11, pp. 366–369, 1895.

[11] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.

[12] D. Sullivan, "What is google pagerank? a guide for searchers & webmasters," *Search Engine Land*, 2007, Available at `http://searchengineland.com`.

[13] R. D. Luce and A. D. Perry, "A method of matrix analysis of group structure," *Psychometrika*, vol. 14, no. 2, pp. 95–116, 1949.

[14] A. Ben-Dor, R. Shamir, and Z. Yakhini, "Clustering gene expression patterns," *Journal of Computational Biology*, vol. 6, no. 3-4, pp. 281–297, 1999.

[15] N. Rhodes, P. Willett, A. Calvet, J. B. Dunbar, and C. Humblet, "Clip: Similarity searching of 3d databases using clique detection," *Journal of Chemical Information and Computer Sciences*, vol. 43, no. 2, pp. 443–448, 2003.

[16] D. Y. Kenett and S. Havlin, "Network science: A useful tool in economics and finance," *Mind & Society*, vol. 14, no. 2, pp. 155–167, 2015.

[17] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical Review Letters*, vol. 86, no. 14, p. 3200, 2001.

[18] R. Cohen, S. Havlin, and D. Ben-Avraham, "Efficient immunization strategies for computer networks and populations," *Physical Review Letters*, vol. 91, no. 24, p. 247 901, 2003.

[19] S. D. Reis, Y. Hu, A. Babino, J. S. Andrade Jr, S. Canals, M. Sigman, and H. A. Makse, "Avoiding catastrophic failure in correlated networks of networks," *Nature Physics*, vol. 10, no. 10, pp. 762–767, 2014.

[20] I. Sendiña–Nadal, Y. Ofran, J. A. Almendral, J. M. Buldú, I. Leyva, D. Li, S. Havlin, and S. Boccaletti, "Unveiling protein functions through the dynamics of the interaction network," *PLoS One*, vol. 6, no. 3, p. 17 679, 2011.

[21] K. Yamasaki, A. Gozolchiani, and S. Havlin, "Climate networks around the globe are significantly affected by el nino," *Physical review letters*, vol. 100, no. 22, p. 228 501, 2008.

[22] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási, "Structure and tie strengths in mobile communication networks," *Proceedings of the National Academy of Sciences*, vol. 104, no. 18, pp. 7332–7336, 2007.

[23] R. Pastor-Satorras and A. Vespignani, *Evolution and structure of the Internet: A statistical physics approach*. Cambridge University Press, 2007.

[24] J. Shao, S. V. Buldyrev, L. A. Braunstein, S. Havlin, and H. E. Stanley, "Structure of shells in complex networks," *Physical Review E*, vol. 80, no. 3, p. 036 105, 2009.

[25] D. Li, B. Fu, Y. Wang, G. Lu, Y. Berezin, H. E. Stanley, and S. Havlin, "Percolation transition in dynamical traffic network with evolving critical bottlenecks," *Proceedings of the National Academy of Sciences*, vol. 112, no. 3, pp. 669–672, 2015.

[26] N. Harada, E. Herbst, and V. Wakelam, "A new network for higher-temperature gas-phase chemistry. i. a preliminary study of accretion disks in active galactic nuclei," *The Astrophysical Journal*, vol. 721, no. 2, p. 1570, 2010.

[27] P. Glasserman and H. P. Young, "How likely is contagion in financial networks?" *Journal of Banking & Finance*, vol. 50, no. 1, pp. 383–399, 2015.

[28] A. Garas, P. Argyrakis, C. Rozenblat, M. Tomassini, and S. Havlin, "Worldwide spreading of economic crisis," *New Journal of Physics*, vol. 12, no. 11, p. 113 043, 2010.

[29] M. Earls, "Advertising to the herd: How understanding our true nature challenges the ways we think about advertising and market research," *International Journal of Market Research*, vol. 45, no. 3, pp. 311–336, 2003.

[30] N. Scott, R. Baggio, and C. Cooper, *Network analysis and tourism: From theory to practice*. Channel View Publications, 2008.

[31] P. Björk and H. Virtanen, "What tourism project managers need to know about co-operation facilitators," *Scandinavian Journal of Hospitality and Tourism*, vol. 5, no. 3, pp. 212–230, 2005.

[32] H. Pechlaner, E. Smeral, and K. Matzier, "Customer value management as a determinant of the competitive position of tourism destinations," *Tourism Review*, vol. 57, no. 4, pp. 15–22, 2002.

[33] N. Leiper, "Tourist attraction systems," *Annals of Tourism Research*, vol. 17, no. 4, pp. 367–384, 1990.

[34] A. Palmer and D. Bejou, "Tourism destination marketing alliances," *Annals of Tourism Research*, vol. 22, no. 3, pp. 616–629, 1995.

[35] M. J. Dollinger, "The evolution of collective strategies in fragmented industries," *Academy of Management Review*, vol. 15, no. 2, pp. 266–285, 1990.

[36] R. Cross, S. P. Borgatti, and A. Parker, "Making invisible work visible: Using social network analysis to support strategic collaboration," *California Management Review*, vol. 44, no. 2, pp. 25–46, 2002.

[37] G. E. Box and G. M. Jenkins, *Time series analysis: Forecasting and control.* Holden-Day, 1970.

[38] A. C. Harvey, *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1990.

[39] M. West and J. Harrison, *Bayesian forecasting and dynamic models.* Springer, 2006.

[40] G. Kitagawa and W. Gersch, *Smoothness priors analysis of time series.* Springer, 1996.

[41] D. M. Grether and M. Nerlove, "Some properties of optimal seasonal adjustment," *Econometrica: Journal of the Econometric Society*, vol. 38, no. 5, pp. 682–703, 1970.

[42] A. Pagan, *Estimation of an evolving seasonal pattern as an application of stochastically varying parameter regression.* Princeton University, Econometric Research Program, 1973.

[43] ——, "A note on the extraction of components from time series," *Econometrica: Journal of the Econometric Society*, vol. 43, pp. 163–168, 1975.

[44] R. F. Engle, "Estimating structural models of seasonality," in *Seasonal analysis of economic time series*, NBER, 1979, pp. 281–308.

[45] J. Durbin and S. J. Koopman, *Time series analysis by state space methods.* Oxford University Press, 2012.

[46] J. J. Commandeur and S. J. Koopman, *An introduction to state space time series analysis.* Oxford University Press, 2007.

[47] G. B. Tawadros, "A structural time series test of the monetary model of exchange rates under four big inflations," *Economic Modelling*, vol. 25, no. 6, pp. 1216–1224, 2008.

[48] M. Marczak and T. Beissinger, "Real wages and the business cycle in germany," *Empirical Economics*, vol. 44, no. 2, pp. 469–490, 2013.

[49] T. Proietti, "Forecasting with structural time series models," in *A Companion to Economic Forecasting*, Blackwell Publishers, Oxford., 1991, pp. 105–132.

[50] S. Hylleberg, *Modelling seasonality.* Oxford University Press, 1992.

[51] P. H. Franses, "Forecasting seasonal time series," in *Econometric Forecasting And High-Frequency Data Analysis*, World Scientific, 2008, pp. 93–130.

[52] G. Einicke, *Smoothing, filtering and prediction: Estimating the past, present and future.* IntechOpen, 2012.

[53] N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications.* MIT Press, 1950.

[54] R. Bucy and P. Joseph, *Filtering for Stochastic Processes with Applications to Guidance.* Interscience-Wiley, New York, 1968.

[55] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.

[56] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. D, pp. 35–45, 1960.

[57] Y. Pei, S. Biswas, D. S. Fussell, and K. Pingali, "An elementary introduction to kalman filtering," *Communications of the ACM*, vol. 62, no. 11, pp. 122–133, 2019.

[58] Y. Kim and H. Bang, "Introduction to kalman filter and its applications," in *Introduction and Implementations of the Kalman Filter*, IntechOpen, 2018.

[59] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *American Institute of Aeronautics and Astronautics journal*, vol. 3, no. 8, pp. 1445–1450, 1965.

[60] S. Tian, C. Jiabin, S. Chunlei, and Y. Huan, "The application of rts smoothing algorithm in the post-processing of the integrated navigation," in *29th Chinese Control And Decision Conference (CCDC)*, IEEE, 2017, pp. 197–201.

[61] C. K. Wikle and L. M. Berliner, "A bayesian tutorial for data assimilation," *Physica D: Nonlinear Phenomena*, vol. 230, no. 1-2, pp. 1–16, 2007.

[62] R. Fletcher, "Sustaining tourism, sustaining capitalism? the tourism industry's role in global capitalist expansion," *Tourism Geographies*, vol. 13, no. 3, pp. 443–461, 2011.

[63] D. Greenwood, "Culture by the pound: An anthropological perspective on tourism as cultural commoditization.," in *Hosts and Guests, The Anthropology of Tourism*, University of Pennsylvania Press, 1989, pp. 171–185.

[64] D. MacCannell, *The tourist: A new theory of the leisure class.* University of California Press, 1999.

[65] N. H. Graburn, "Secular ritual: A general theory of tourism," in *Hosts and Guests Revisited: Tourism Issues of the 21st Century*, Waveland Press, Inc., 2004, pp. 23–34.

[66] J. Urry, *The tourist gaze.* Sage, 2002.

[67] M. Honey, *Ecotourism and sustainable development: Who owns paradise?* Island Press, 2008.

[68] I. Munt, "Eco-tourism or ego-tourism?" *Race & Class*, vol. 36, no. 1, pp. 49–60, 1994.

[69] C.-C. Lee and C.-P. Chang, "Tourism development and economic growth: A closer look at panels," *Tourism Management*, vol. 29, no. 1, pp. 180–192, 2008.

[70] B. Archer, "Demand forecasting and estimation," in *Travel, Tourism and Hospitality Research.* John Wiley & Sons, Inc., 1987, pp. 77–85.

[71] B.-Q. Wu, J.-F. Wu, X.-T. Shi, T.-G. Zhang, C.-C. Deng, and S.-S. Wu, "Visiting probability model: A new method for tourist volume forecasting," *Asia Pacific Journal of Tourism Research*, vol. 24, no. 12, pp. 1155–1168, 2019.

[72] M. Batty, "The size, scale, and shape of cities," *Science*, vol. 319, no. 5864, pp. 769–771, 2008.

[73] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[74] V. Belik, T. Geisel, and D. Brockmann, "Natural human mobility patterns and spatial spread of infectious diseases," *Physical Review X*, vol. 1, no. 1, p. 011 001, 2011.

[75] A. Wesolowski, N. Eagle, A. J. Tatem, D. L. Smith, A. M. Noor, R. W. Snow, and C. O. Buckee, "Quantifying the impact of human mobility on malaria," *Science*, vol. 338, no. 6104, pp. 267–270, 2012.

[76] M. Tizzoni, P. Bajardi, A. Decuyper, G. K. K. King, C. M. Schneider, V. Blondel, Z. Smoreda, M. C. González, and V. Colizza, "On the use of human mobility proxies for modeling epidemics," *PLoS Computational Biology*, vol. 10, no. 7, e1003716, 2014.

[77] E. X. Jiao and J. L. Chen, "Tourism forecasting: A review of methodological developments over the last decade," *Tourism Economics*, vol. 25, no. 3, pp. 469–492, 2019.

[78] H. Song, R. T. Qiu, and J. Park, "A review of research on tourism demand forecasting," *Annals of Tourism Research*, vol. 75, pp. 338–362, 2019.

[79] G. Athanasopoulos, R. J. Hyndman, H. Song, and D. C. Wu, "The tourism forecasting competition," *International Journal of Forecasting*, vol. 27, no. 3, pp. 822–844, 2011.

[80] G. Laber, "Determinants of international travel between canada and the united states," *Geographical Analysis*, vol. 1, no. 4, pp. 329–336, 1969.

[81] C. A. Martin and S. F. Witt, "Tourism demand forecasting models: Choice of appropriate variable to represent tourists' cost of living," *Tourism Management*, vol. 8, no. 3, pp. 233–246, 1987.

[82] S. Ongan and G. Gozgor, "Tourism demand analysis: The impact of the economic policy uncertainty on the arrival of japanese tourists to the usa," *International Journal of Tourism Research*, vol. 20, no. 3, pp. 308–316, 2018.

[83] V. S. Lin, A. Liu, and H. Song, "Modeling and forecasting chinese outbound tourism: An econometric approach," *Journal of Travel & Tourism Marketing*, vol. 32, no. 1-2, pp. 34–49, 2015.

[84] E. Smeral, "International tourism demand and the business cycle," *Annals of Tourism Research*, vol. 39, no. 1, pp. 379–400, 2012.

[85] C. Goh, "Exploring impact of climate on tourism demand," *Annals of Tourism Research*, vol. 39, no. 4, pp. 1859–1883, 2012.

[86] E. Smeral, "The impact of the financial and economic crisis on european tourism," *Journal of Travel Research*, vol. 48, no. 1, pp. 3–13, 2009.

[87] ——, "Impacts of the world recession and economic crisis on tourism: Forecasts and potential risks," *Journal of Travel Research*, vol. 49, no. 1, pp. 31–38, 2010.

[88] T. C. Syriopoulos, "A dynamic model of demand for mediterranean tourism," *International Review of Applied Economics*, vol. 9, no. 3, pp. 318–336, 1995.

[89] W. Long, C. Liu, and H. Song, "Pooling in tourism demand forecasting," *Journal of Travel Research*, vol. 58, no. 7, pp. 1161–1174, 2019.

[90] G. Li, K. K. Wong, H. Song, and S. F. Witt, "Tourism demand forecasting: A time varying parameter error correction model," *Journal of Travel Research*, vol. 45, no. 2, pp. 175–185, 2006.

[91] H. Song, S. F. Witt, and T. C. Jensen, "Tourism forecasting: Accuracy of alternative econometric models," *International Journal of Forecasting*, vol. 19, no. 1, pp. 123–141, 2003.

[92] P. F. Bangwayo-Skeete and R. W. Skeete, "Can google data improve the forecasting performance of tourist arrivals? mixed-data sampling approach," *Tourism Management*, vol. 46, pp. 454–464, 2015.

[93] W. R. Moore, "The impact of climate change on caribbean tourism demand," *Current Issues in Tourism*, vol. 13, no. 5, pp. 495–505, 2010.

[94] S. Saha and G. Yap, "The moderation effects of political instability and terrorism on tourism development: A cross-country panel analysis," *Journal of Travel Research*, vol. 53, no. 4, pp. 509–521, 2014.

[95] S. Page, H. Song, and D. C. Wu, "Assessing the impacts of the global economic crisis and swine flu on inbound tourism demand in the united kingdom," *Journal of Travel Research*, vol. 51, no. 2, pp. 142–153, 2012.

[96] C. Bonham, C. Edmonds, and J. Mak, "The impact of 9/11 and other terrible global events on tourism in the united states and hawaii," *Journal of Travel Research*, vol. 45, no. 1, pp. 99–110, 2006.

[97] H. Song, S. Lin, S. F. Witt, and X. Zhang, "Impact of financial/economic crisis on demand for hotel rooms in hong kong," *Tourism Management*, vol. 32, no. 1, pp. 172–186, 2011.

[98] T. Dergiades, E. Mavragani, and B. Pan, "Google trends and tourists' arrivals: Emerging biases and proposed corrections," *Tourism Management*, vol. 66, pp. 108–120, 2018.

[99] X. Yang, B. Pan, J. A. Evans, and B. Lv, "Forecasting chinese tourist volume with search engine data," *Tourism Management*, vol. 46, pp. 386–397, 2015.

[100] B. Peng, H. Song, and G. I. Crouch, "A meta-analysis of international tourism demand forecasting and implications for practice," *Tourism Management*, vol. 45, pp. 181–193, 2014.

[101] C. A. Martin and S. F. Witt, "Forecasting tourism demand: A comparison of the accuracy of several quantitative methods," *International Journal of Forecasting*, vol. 5, no. 1, pp. 7–19, 1989.

[102] S. F. Witt, G. D. Newbould, and A. J. Watkins, "Forecasting domestic tourism demand: Application to las vegas arrivals data," *Journal of Travel Research*, vol. 31, no. 1, pp. 36–41, 1992.

[103] J. T. Coshall, "Combining volatility and smoothing forecasts of uk demand for international tourism," *Tourism Management*, vol. 30, no. 4, pp. 495–511, 2009.

[104] J. T. Coshall and R. Charlesworth, "A management orientated approach to combination forecasting of tourism demand," *Tourism Management*, vol. 32, no. 4, pp. 759–769, 2011.

[105] R. Fildes, Y. Wei, and S. Ismail, "Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures," *International Journal of Forecasting*, vol. 27, no. 3, pp. 902–922, 2011.

[106] S. Kim and H. Song, "Analysis of inbound tourism demand in south korea: A cointegration and error correction approach," *Tourism Analysis*, vol. 3, no. 1, pp. 25–41, 1998.

[107] M. D. Geurts and I. Ibrahim, "Comparing the box-jenkins approach with the exponentially smoothed forecasting model application to hawaii tourists," *Journal of Marketing Research*, vol. 12, no. 2, pp. 182–188, 1975.

[108] Y.-M. Chan, "Forecasting tourism: A sine wave time series regression approach," *Journal of Travel research*, vol. 32, no. 2, pp. 58–60, 1993.

[109] G. S. Dharmaratne, "Forecasting tourist arrivals in barbados," *Annals of Tourism Research*, vol. 22, no. 4, pp. 804–818, 1995.

[110] C. Lim and M. McAleer, "A seasonal analysis of asian tourist arrivals to australia," *Applied Economics*, vol. 32, no. 4, pp. 499–509, 2000.

[111] M. G. Millán, M. d. P. Pablo-Romero, and J. Sánchez-Rivas, "Oleotourism as a sustainable product: An analysis of its demand in the south of spain (andalusia)," *Sustainability*, vol. 10, no. 1, p. 101, 2018.

[112] F. Chu, "A fractionally integrated autoregressive moving average approach to forecasting tourism demand," *Tourism Management*, vol. 29, no. 1, pp. 79–88, 2008.

[113] ——, "Analyzing and forecasting tourism demand with arar algorithm," *Tourism Management*, vol. 29, no. 6, pp. 1185–1196, 2008.

[114] C. Goh and R. Law, "Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention," *Tourism Management*, vol. 23, no. 5, pp. 499–510, 2002.

[115] F. Chan, C. Lim, and M. McAleer, "Modelling multivariate international tourism demand and volatility," *Tourism Management*, vol. 26, no. 3, pp. 459–471, 2005.

[116] C. Holt, "Forecasting trends and seasonal by exponentially weighted moving averages," *ONR Memorandum*, vol. 52, no. 2, 1957.

[117] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management Science*, vol. 6, no. 3, pp. 324–342, 1960.

[118] L. R. Weatherford and S. E. Kimes, "A comparison of forecasting methods for hotel revenue management," *International Journal of Forecasting*, vol. 19, no. 3, pp. 401–415, 2003.

[119] L. W. Turner and S. F. Witt, "Forecasting tourism using univariate and multivariate structural time series models," *Tourism Economics*, vol. 7, no. 2, pp. 135–147, 2001.

[120] A. Guizzardi and M. Mazzocchi, "Tourism demand for italy and the business cycle," *Tourism Management*, vol. 31, no. 3, pp. 367–377, 2010.

[121] A. Clewer, A. Pack, and M. T. Sinclair, "Forecasting models for tourism demand in city dominated and coastal areas," *Papers of the Regional Science Association*, vol. 69, no. 1, pp. 31–42, 1990.

[122] H. Li, C. Goh, K. Hung, and J. L. Chen, "Relative climate index and its effect on seasonal tourism demand," *Journal of Travel Research*, vol. 57, no. 2, pp. 178–192, 2018.

[123] B. Pan and Y. Yang, "Forecasting destination weekly hotel occupancy with big data," *Journal of Travel Research*, vol. 56, no. 7, pp. 957–970, 2017.

[124] W. H. K. Tsui, H. O. Balli, A. Gilbey, and H. Gow, "Forecasting of hong kong airport's passenger throughput," *Tourism Management*, vol. 42, pp. 62–76, 2014.

[125] S. Park, J. Lee, and W. Song, "Short-term forecasting of japanese tourist inflow to south korea using google trends data," *Journal of Travel & Tourism Marketing*, vol. 34, no. 3, pp. 357–368, 2017.

[126] M. Uysal and M. S. El Roubi, "Artificial neural networks versus multiple regression in tourism demand analysis," *Journal of Travel Research*, vol. 38, no. 2, pp. 111–118, 1999.

[127] H. Constantino, P. O. Fernandes, and J. P. Teixeira, "Tourism demand modelling and forecasting with artificial neural network models: The mozambique case study," *Tékhne*, vol. 14, no. 2, pp. 113–124, 2016.

[128] C.-F. Chen, M.-C. Lai, and C.-C. Yeh, "Forecasting tourism demand based on empirical mode decomposition and neural network," *Knowledge-Based Systems*, vol. 26, pp. 281–287, 2012.

[129] S. Cang and H. Yu, "A combination selection algorithm on forecasting," *European Journal of Operational Research*, vol. 234, no. 1, pp. 127–139, 2014.

[130] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[131] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy Sets and Systems*, vol. 54, no. 3, pp. 269–277, 1993.

[132] C.-H. Wang and L.-C. Hsu, "Constructing and applying an improved fuzzy time series model: Taking the tourism industry for example," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2732–2738, 2008.

[133] C. Goh, R. Law, and H. M. Mok, "Analyzing and forecasting tourism demand: A rough sets approach," *Journal of Travel Research*, vol. 46, no. 3, pp. 327–338, 2008.

[134] D. Ju-Long, "Control problems of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.

[135] X. Sun, W. Sun, J. Wang, Y. Zhang, and Y. Gao, "Using a grey–markov model optimized by cuckoo search algorithm to forecast the annual foreign tourist arrivals to china," *Tourism Management*, vol. 52, pp. 369–379, 2016.

[136] P.-F. Pai, K.-C. Hung, and K.-P. Lin, "Tourism demand forecasting using novel hybrid system," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3691–3702, 2014.

[137] W. Lijuan and C. Guohua, "Seasonal svr with foa algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow," *Knowledge-Based Systems*, vol. 110, pp. 157–166, 2016.

[138] J. Shahrabi, E. Hadavandi, and S. Asadi, "Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series," *Knowledge-Based Systems*, vol. 43, pp. 112–122, 2013.

[139] K.-Y. Chen, "Combining linear and nonlinear model in forecasting tourism demand," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10 368–10 376, 2011.

[140] G. Li, H. Song, and S. F. Witt, "Recent developments in econometric modeling and forecasting," *Journal of Travel Research*, vol. 44, no. 1, pp. 82–99, 2005.

[141] UNWTO, *Tourism highlights: 2017 edition. international tourism trends*, Available at `https://people.unica.it/carlamassidda/files/2017/06/UNWTO_Tourism-Highlights_2017.pdf`, 2017.

[142] ——, *Tourism highlights: 2018 edition. international tourism trends*, Available at `https://www.e-unwto.org/doi/pdf`, 2018.

[143] Instituto Nacional de Estatistica y Censos, *Turismo receptivo y emisivo: Llegada y salida de turistas*, Available at `https://www.indec.gob.ar`, 2013-2019.

[144] Knoema, *Visitor arrivals and resident returns by country, australia*, Available at `https://knoema.com`, 2010-2020.

[145] StatBel, *Touriste selon pays de provenance*, Available at `https://statbel.fgov.be`, 2015-2018.

[146] Ministério do Turismo Brasilero, *Anuário estatístico de turismo*, Available at `http://www.dadosefatos.turismo.gov.br`, 2012-2019.

[147] Statistics Canada, *Trips to canada by travellers from abroad*, Available at `https://www150.statcan.gc.ca`, 1993-2020.

[148] Institut national de la statistique et des études économiques, *Enquête de fréquentation hôtelière*, Available at `https://www.insee.fr`, 2011-2019.

[149] DESTATIS, *Monatserhebung im tourismus*, Available at `https://www.destatis.de`, 2018-2020.

[150] Hellenic Statistical Authority, *Arrivals of non-residents from abroad, by country of residence*, Available at https://www.statistics.gr, 2008-2016.

[151] Badan Pusat Statistik, *Perkembangan pariwisata dan transportasi nasional*, Available at https://www.bps.go.id, 2018-2020.

[152] Israel Central Bureau of Statistics, *Monthly bulletin of statistics*, Available at https://old.cbs.gov.il, 2012-2020.

[153] ISTAT, *Movimento dei clienti negli esercizi ricettivi*, Available at http://dati.istat.it, 2008-2018.

[154] JNTO, *Visitor arrivals to japan and japanese overseas travelers*, Available at https://statistics.jnto.go.jp, 1990-2020.

[155] Instituto Nacional de Estadística y Geografía, *Visitor arrivals to mexico by airport*, Available at https://en.www.inegi.org.mx, 2012-2020.

[156] Centraal Bureau voor de Statistiek, *Visitor in accomodation*, Available at https://www.cbs.nl/en-gb, 2012-2020.

[157] Figure.nz, *Travel - visitor arrivals by country of residence, purpose and nz port (monthly) apr 1978–mar 2020*, Available at https://figure.nz, 1978-2020.

[158] Glowny Urzad Statystyczny, *Nights spent by tourists (domestic and foreign) in tourist accommodation*, Available at https://stat.gov.pl, 2017-2020.

[159] Instituto Nacional de Estadística, *Statistics portugal, holiday campers stays in holiday camps survey*, Available at http://www.ine.pt, 2008-2018.

[160] STAT SA, *Report on tourism*, Available at http://www.statssa.gov.za, 2017-2019.

[161] Korea Tourism Organization, *Korea, monthly statistics of tourism*, Available at https://kto.visitkorea.or.kr, 2014-2020.

[162] Instituto Nacional de Estadística, *Detaile monthly results*, Available at https://www.ine.es/buscar, 2015-2020.

[163] National Statistics Office, *International tourist arrivals to thailand by nationality: 2009 - 2018*, Available at http://statbbi.nso.go.th, 2009-2018.

[164] Turkish Statistical Institute, *Number of tourists by nationality*, Available at http://www.turkstat.gov.tr, 2006-2020.

[165] visitbritain.org, *Inbound nation, region and county data*, Available at https://www.visitbritain.org, 2002-2018.

[166] National Travel and Tourism Office, *I-94 arrivals based on country of residence (cor)*, Available at https://travel.trade.gov/research/programs/i94/index.asp, 2011-2020.

[167] UNWTO, *Outbound tourism: Trips abroad by resident visitors to countries of destination*, Available at https://www.e-unwto.org, 2008-2018.

[168] J. N. Albrecht, "Networking for sustainable tourism–towards a research agenda," *Journal of Sustainable Tourism*, vol. 21, no. 5, pp. 639–657, 2013.

[169] C. Casanueva, Á. Gallego, and M.-R. García-Sánchez, "Social network analysis in tourism," *Current Issues in Tourism*, vol. 19, no. 12, pp. 1190–1209, 2016.

[170] M. G. Chung, A. Herzberger, K. A. Frank, and J. Liu, "International tourism dynamics in a globalized world: A social network analysis approach," *Journal of Travel Research*, vol. 59, no. 3, pp. 387–403, 2020.

[171] T. Haryanto, "Covid-19 pandemic and international tourist demand," *JDE (Journal of Developing Economies)*, vol. 5, no. 1, pp. 1–5, 2020.

[172] S. Gössling, D. Scott, and C. M. Hall, "Pandemics, tourism and global change: A rapid assessment of covid-19," *Journal of Sustainable Tourism*, pp. 1–20, 2020, doi.org/10.1080/09669582.2020.1758708.

[173] L. Peng, W. Yang, D. Zhang, C. Zhuge, and L. Hong, "Epidemic analysis of covid-19 in china by dynamical modeling," *arXiv preprint arXiv:2002.06563*, 2020.

[174] Y. Gu, *Covid-19 projection*, Available at https://covid19-projections.com, 2020.

[175] C. T. Bergstrom, *Covid-19 projection analysis*, Available at https://twitter.com/CT_Bergstrom, 29-04-2020.

[176] B. W. Ritchie, "Chaos, crises and disasters: A strategic approach to crisis management in the tourism industry," *Tourism Management*, vol. 25, no. 6, pp. 669–683, 2004.

[177] J. Stacey, "Tourism policy responses to the coronavirus (covid-19)," *OECD Policy Responses to Coronavirus (COVID-19)*, 2020, Available at http://www.oecd.org/coronavirus/policy-responses/tourism-policy-responses-to-the-coronavirus-covid-19-6466aa20/.

[178] J. Wen, M. Kozak, S. Yang, and F. Liu, "Covid-19: Potential effects on chinese citizens' lifestyle and travel," *Tourism Review*, 2020, doi: 10.1108/TR-03-2020-0110.

[179] Y. Zheng, E. Goh, and J. Wen, “The effects of misleading media reports about covid-19 on chinese tourists’ mental health: A perspective article,” *Anatolia*, vol. 31, no. 2, pp. 337–340, 2020.

[180] O. Coibion, Y. Gorodnichenko, and M. Weber, “Labor markets during the covid-19 crisis: Apreliminary review,” National Bureau of Economic Research, Tech. Rep. w227017, 2020.