Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

Complex Networks of the Air Passenger Traffic in Monterrey´s Airport

Olivia Sashiko Shirai Reyna[a] , Idalia Flores de la Mota[b]

[a,b] *Facultad de Ingeniería UNAM, Edificio Bernardo Quintana 3er Piso Departamento de Sistemas, UNAM, Cd. Universitaria, Del. Coyoacán, C.P. 04510, C.D.M.X., México*

Abstract

Nowadays, the air passenger traffic has been increasing, becoming an excellent, viable and reachable option for many people. This causes that airports may require an efficient organization to serve both, the companies that use the facilities and the passengers. In addition, it is important to consider that the amount of information that is generated may not be easy to analyse, sometimes because the managers don´t know all the information that they have, or they don´t know how much this information can help the business or what they can do with all these data. Therefore, in this work, we perform an analysis of the information obtained from Monterrey´s airport (domestic and international passengers), using the methodology of Complex Networks. Also, with the results obtained, we will seek to put forward improvements in the service of this type of facilities, and the infrastructure.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: complex networks; visibility; time series; forecast; air passengers.

1. Introduction

Since 2000, a record of all the passengers that arrive and depart at Mexican airports from domestic and international flights have been kept and have not been analysed because some of these data are not open data. Some of the data that may concern will be the number of airlines that operate in Mexican airports, the number of passengers, the number of routes, the number of available and occupied seats, delays on the flights and more information at domestic and international flights.

The main problem is that the strengths and weaknesses of Mexican airports as a whole system are not known or identified, causing, among other things, losses in business opportunities and system saturation. For this reason, it is important to perform an initial diagnosis. Then according to the results of this analysis we model the system. Then we

2352-1465
Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY

built mathematical models of the real data using complex networks, in order to put forward improvements for market strategies.

In order to improve overall business strategies, it is necessary to apply statistical models and mathematical models to analyse, find patterns and predict data behaviour. In this work, we will use different mathematical techniques such as network theory, complex networks, statistics, simulation and time series.

Mexico is the third country with the highest number of airports, aerodromes and flight tracks with a total of 1714, according to data published by the CIA (Central Intelligence Agency) World Factbook in 2013. So, it is important to perform a study of what happens in Mexican airports, how the flow of passengers in these airports behaves, and the problems they have. And therefore, propose alternatives to improve operations and infrastructure both in its operations and in its infrastructure. For this reason, we focused the study on one of the most important airports in Mexico, the Monterrey´s airport. Monterrey is the third largest city in Mexico and one of the most productive industrial centres in Mexico.

The airport is considered one of the most modern airports in North America as it has a design that is based on energy savings and sustainable awareness, serving more than 9 million passengers a year. Currently 85% of passenger traffic is local, mainly from Mexico City, Cancun, Guadalajara, Chihuahua, and Tijuana, and 15% of passenger traffic is international, mainly from the US cities of Miami, Dallas and Houston. It has more than 300 daily flights to more than 43 destinations in Mexico, the United States, Central America and the Caribbean and Asia. It is considered the fourth airport in the country in terms of passengers served and operations per year, after the International Airport of Mexico City (AICM), the International Airport of Cancun and the International Airport of Guadalajara. The Monterrey´s Airport ranks as OMA's (Grupo Aeroportuario Centro Norte) busiest airport.

2. Methodology

For this study we focused on the analysis of domestic, international and total passengers served by the Monterrey airport from January 2000 to December 2017. With this information, the first step was to perform a statistical analysis of the data, the second step was to work with all this information as time series, we did a decomposition of the time series, we obtained the ACF (Auto Correlation Function), PACF (Partial Auto Correlation Function) and we computed the differences, then with the results we modelled a time series as an ARIMA (Auto Regressive Integrated Moving Average) model. Finally, with this model we were able to do some forecast of the air passenger traffic for the next two years to analyse the strengths and the weaknesses of all the system in the future.

As a third step, after the time series analysis, we applied the visibility algorithm to transform the time series into a complex network. Once this network is obtained we used the methodology of complex networks, specially to analyze the topology or structure of the networks, for example, the clustering, the closeness, betweenness, assortativity, and more metrics of complex networks. Also, we can have the degree distribution of the networks and we can have a good approach of the networks behavior.

With all these results we were able to analyze and compare the domestic, international and total passenger networks, in addition, we could analyze the topology of these networks, which concludes what type of network model they belong to and what specific characteristics and properties they share.

For all this statistical analysis, we used the R software, which is an open source programming language and software environment for statistical computing and graphics. For this work, we used specific R software packages, such as, igraph, networks, tkrplot, sand, sna, forecast, TimeSeries, TSA and others, this software allow us to generate graphs/networks, compute different network metrics like clustering or transitivity, different centrality metrics, plot networks, create mathematical models, forecast data and more functions.

First, we used all the data that we had, so, we plot this information as a time series. The next figures show the different time series for domestic, international, and total passengers of Monterrey´s Airport (respectively).

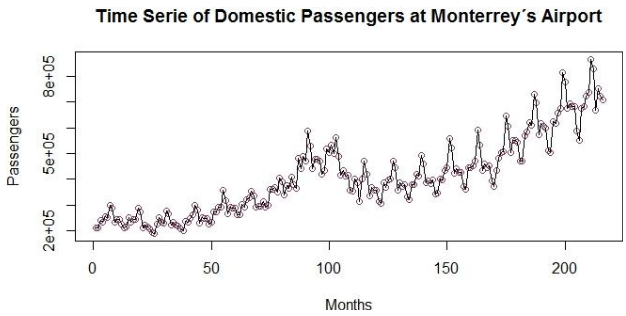

Figure 1 Time Series Domestic Passengers Monterrey´s Airport

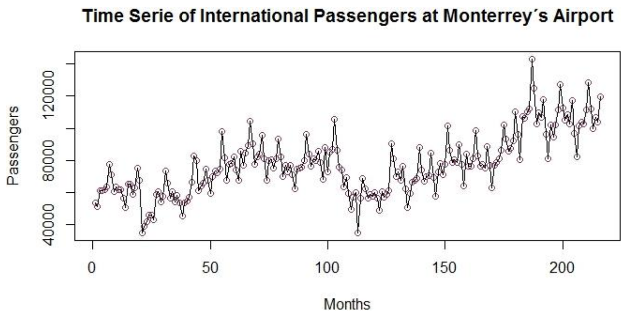

Figure 2 Time Series International Passengers Monterrey´s Airport

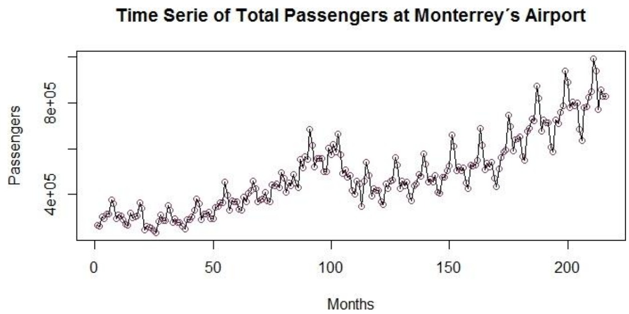

Figure 3 Time Series Total Passengers Monterrey´s Airport

Now, once we have the plots of the time series we use a technique to transform these time series into complex networks, so the best way to do this transformation is with the visibility algorithm (Lacasa L. et al 2008).

The main goal of this algorithm is to map a time series into a complex network, so, we want to study and analyze the complex networks with all the techniques and properties of network theory so it will be a more complete analysis not only the study of the time series with all their properties. Another important thing is that this network inherits several properties of the time series.

The criterion of this algorithm is to set up two arbitrary data values (t_a, y_a) and (t_b, y_b) that will have visibility, and consequently will become two connected nodes of the associated graph, if any other data (t_c, y_c) placed between them fulfills:

$$y_c < y_b + (y_a - y_b)\frac{t_b - t_c}{t_b - t_a} \quad (1)$$

So, we apply this algorithm to our time series and we obtain the next graphs.

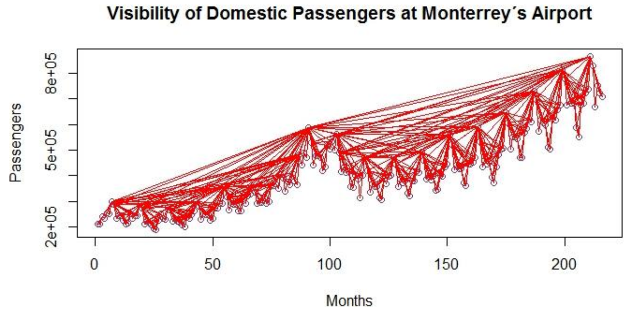

Figure 4 Visibility Domestic Passengers Monterrey´s Airport

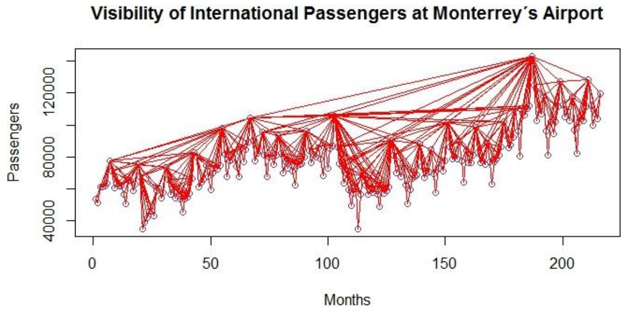

Figure 5 Visibility International Passengers Monterrey´s Airport

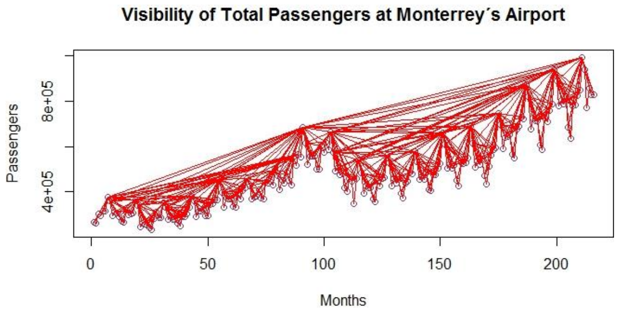

Figure 6 Visibility Total Passengers Monterrey´s Airport

We can easily check that by means of the present algorithm, the associated graph extracted from a time series is always:

1. Connected: each node sees at least its nearest neighbours (left and right), the first and the last one at least sees one.

2. Undirected: the way the algorithm is built up, there is no direction defined in the links.

3. Invariant under affine transformations of the series data: the visibility criterion is invariant under rescaling of both horizontal and vertical axes, and under horizontal and vertical translations.

After, we applied the visibility algorithm, we can obtain the visibility graph, so now, we plot the time series as complex networks, so, with all the complex networks are what we are going to use to compute specific network metrics to study the topology, that is what we will see in the next figures.

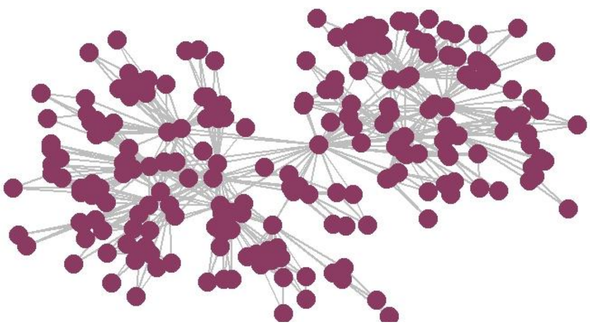

Figure 7 Network Domestic Passengers Monterrey´s Airport

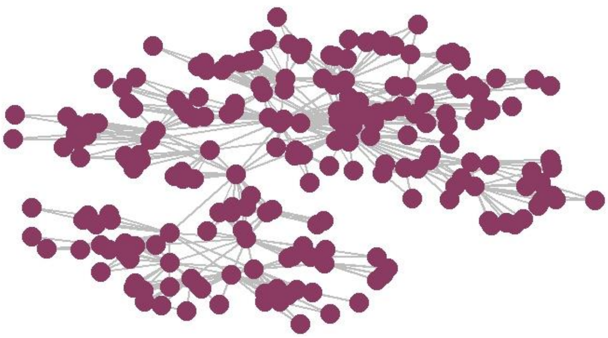

Figure 8 Network International Passengers Monterrey´s Airport

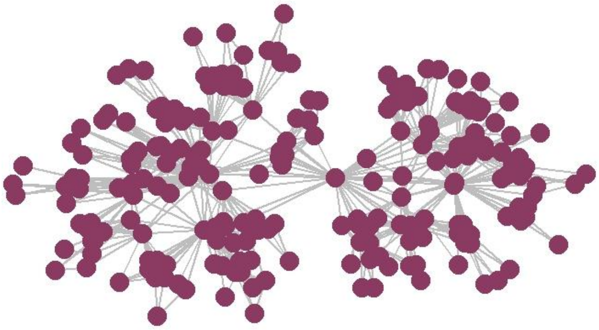

Figure 9 Network Total Passengers Monterrey´s Airport

Now, with the graphs above we are going to give some important definitions. A complex network is a graph or network with non-trivial topological features, with patterns of connection between their elements that are neither purely regular nor purely random. Such features include a heavy tail in the degree distribution, a high clustering coefficient, assortativity or disassortativity among vertices, community structure, and hierarchical structure. that do not occur in simple networks such as lattices or random graphs but often occur in graphs modelling of real systems (Newman, 2010).

After we have our data as networks we can compute the different network´s metrics. We compute and compare metrics of our three networks, such as the number of edges, minimum, maximum and mean degree, diameter, mean distance, the number of cliques, the density of the network, if it is assortativity or disassortativity among nodes, the global clustering of the network, the mean local clustering, closeness centrality, degree centrality and betweenness centrality. We compute these metrics to study the topology of the networks and to understand and identify different patterns and to classify the networks in the different complex networks models that we have (Random networks, Scale-free networks and Small World networks).

Table 1 Metric Results

Results	Domestic	International	Total
Nodes	216	216	216
Edges	836	723	815
Max. Degree	41	31	35
Min. Degree	2	2	2
Mean Degree	7.740741	6.694444	7.5463
Diameter	7	8	7
Mean Distance	3.385702	3.885271	3.490913
Cliques	8	8	9
Density	0.03600345	0.03113695	0.03509905
Assortativity	0.1008781	0.03829409	0.08140715
Global Clustering	0.4541985	0.4435291	0.4629545
Mean Local Clustering	0.4107809	0.7684028	0.7834424
Closeness Centrality	0.7714598	0.306989	0.3881132
Degree Centrality	0.1546942	0.1130491	0.1276916
Betweenness Centrality	0.5414731	0.4442892	0.5418021

First, we need to understand the numbers that we have in the Table 1. We notice that there is a lot of difference between the minimum, maximum and the mean degree, which is because there are just few nodes with a lot of links and there are many nodes with just a few links. The diameter is similar in the three networks. The mean distance is small that means that we can go from any node from the network to another node in a few steps, in this case approximately 3 steps. The number of cliques in the network tells us the number of complete subgraphs in our network. The network density describes the portion of potential connections in a network that are actual connected, so that means that our networks don´t have a lot of actual connections.

To continue with our analysis, our three networks are assortative so that means that there is a preference for a network´s nodes to attach to others that are similar in some way, so assortativity is often operationalized as a correlation between nodes. The global clustering coefficient is based on triplets of nodes and the local clustering coefficient of a node in a network quantifies how close its neighbours are to be a clique (Bollobás, 1998). In our networks, the global clustering and the mean local clustering are not so high approximately 0.4 but we have a high mean local clustering in the cases of the for international and total passengers.

We also computed the centrality metrics, the closeness centrality of a node is a measure of centrality in a network, calculated as the sum of the length of the shortest paths between the node and all other nodes in the network. Thus, the more central a node is, the closer it is to all other nodes, so, in our networks the closeness centrality is not so high only for the domestic passengers we have a high closeness centrality (Caldarelli et. al. 2012). The degree centrality is defined as the number of links incident upon a node (i.e., the number of ties that a node has), in our case, the networks have a low degree centrality (Newman, 2010). The betweenness centrality is a measure of centrality in a network based on shortest paths. For every pair of vertices in a connected network, there exists at least one shortest path between the vertices such that either the number of edges that the path passes through so, the betweenness centrality for each node is the number of these shortest paths that pass through the node; in our networks, we have betweenness centrality among .54 that is not to low but not so high (Caldarelli et. al. 2012).

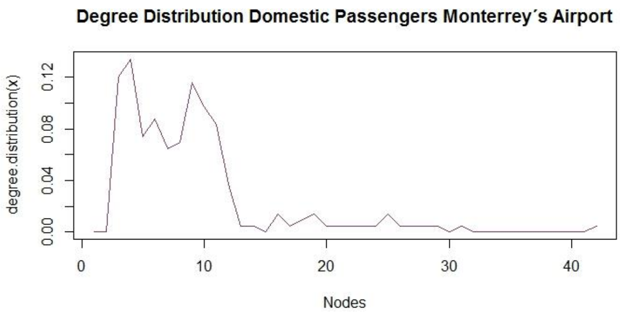

Figure 10 Degree Distribution Domestic Passengers Monterrey´s Airport

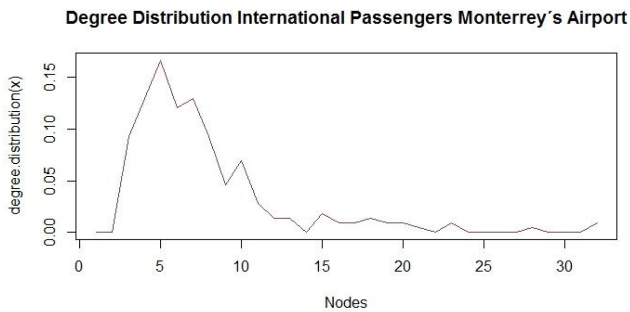

Figure 11 Degree Distribution International Passengers Monterrey´s Airport

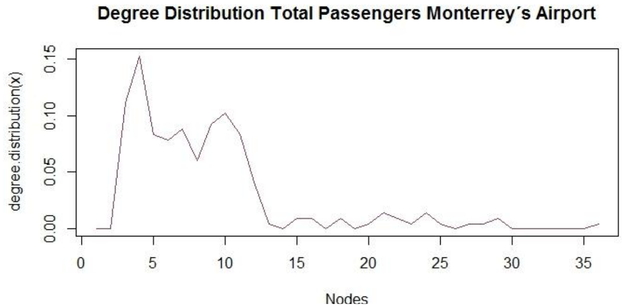

Figure 12 Degree Distribution Total Passengers Monterrey´s Airport

With the degree distribution graphs above we notice that it seems that our networks have a power law distribution. Maybe, it is not so clear but there are a few nodes with a high degree and it decays fast, so, there are a lot of nodes with a low degree.

As part of the statistical analysis, we create mathematical models (ARIMA – Auto Regressive Integrated Moving Average models) to forecast the time series so we can have an idea of how is going to be the pattern of growth, the tendency and cycle of our passengers in the next 24 months, also we have a confidence interval of 80% and 95%.

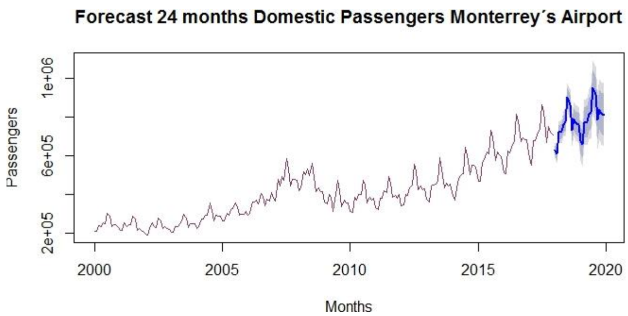

Figure 13 Forecast 24 months Domestic Passengers Monterrey´s Airport

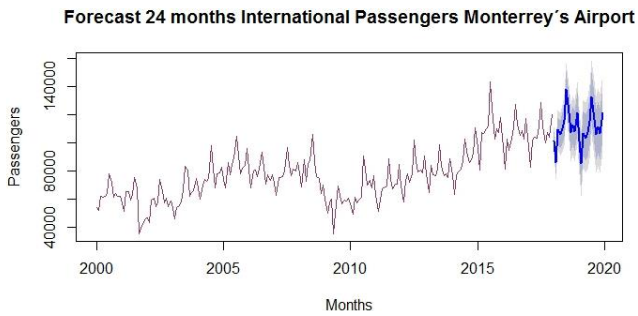

Figure 14 Forecast 24 months International Passengers Monterrey´s Airport

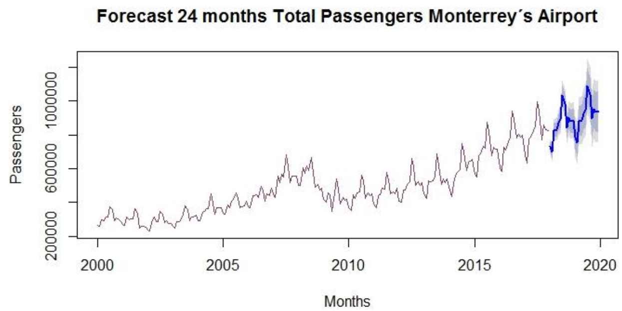

Figure 15 Forecast 24 months Total Passengers Monterrey´s Airport

The Monterrey´s airport recently opened the Air Cargo Terminal, which has an area of 60,000 m^2 for operations. No other airport in Mexico has facilities of such magnitude and much less designed specifically for the transport of cargo by air.

The Monterrey´s airport terminal has been recently remodeled in its image and facilities in order to provide its users with pleasant and efficient facilities. Among other innovations stand out the automatic counters for self-documentation of passengers, internet access modules, new VIP lounges and for passengers on transit and regional flights. With these efforts, the Monterrey´s airport is at the forefront in the development of airport facilities in Mexico.

Within the main OMA projects, the capacity of the Monterrey international airport is expected to be expanded in order to avoid saturation and increase its capacity to mobilize up to 20 million passengers per year. This increase can be seen in the forecast that we made.

The expansion program will consist of three stages:

- 2015 to 2020. Adaptation of roads and platforms to allow the expansion of terminal buildings.
- 2020 to 2025. Expansion works for passenger terminals.
- 2025 to 2030. Integration of terminals in a single airport complex and construction of a second runway (1,500 meters from the current one) to enable simultaneous operations

3. Conclusions

With all these results, we can conclude that our networks follow a scale-free model. The most notable characteristic in a scale-free network is the relative commonness of nodes with a degree that greatly exceeds the average. The scale-free property strongly correlates with the network's robustness to failure. It turns out that the major hubs are closely followed by smaller ones. This implies that the low-degree nodes belong to very dense sub-networks and those sub-networks are connected to each other through hubs (Barabási et. al. 2003). For the time series and forecast analysis, we have a growth in the passengers but it´s supposed that the tendency and cycle will be the same as the have seen in the last years, so the action plan seems to be working good for the next years.

Acknowledgments

To CONACYT scholarship for the first author, and DGAPA-PAPIIT project IN 102117 Accesibilidad y movilidad del transporte público urbano en la ciudad de México, el caso de la delegaci6n Tlalpan.

References

Barabási A.-L., Albert R., 1999. Emergence of scaling in random networks. Science 286 (5439): 509-512.

Barabási A.-L., Bonabeau E., 2003. Scale-Free Networks. Scientific American. 50-59.

Boccaletti S., Latora V., Moreno Y., Chavez., Hwang D.-U., 2006. Complex networks: Structure and dynamics. Physics Reports. 424. 175-308.

Bollobás B., 1998. Modern Graph Theory. Springer.

Buchanan M., 2002. Nexus: Small Worlds and the Groundbreaking Science of Networks. New York, N.Y.: W.W. Norton.

Caldarelli G., Catanzaro M., 2012. Networks: A Very Short Introduction. Oxford University Press, Oxford.

Caldarelli G., Vespignani A.,2007. Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Science. World Scientific Publishing Co., Inc., USA.

Dalgaard P., 2008. Introductory Statistics with R. Springer.

Dorogovtsev S.N., 2010. Lectures on Complex Networks. Oxford University Press, Inc. New York, NY, USA.

Kolaczyk E.D., Csárdi G., 2014. Statistical Analysis of Network Data with R. Springer, New York, USA.

Lacasa L., Luque B., Ballesteros F., Luque J., Nuño J.C., 2008. From time series to complex networks: The visibility graph. PNAS. Vol. 5. No, 13. 4972-4975.

Lacasa, L. et al., 2015. Network structure of multivariate time series. Sci. Rep. 5, 15508; doi: 10.1038/srep15508.

Lütkepohl H., 2005. New Introduction to Multiple Time Series Analysis. Springer.

Newman M.E.J., 2010. Networks: An Introduction. Oxford University Press, Oxford.

Newman M.E.J., 2003. The Structure and Function of Complex Networks. SIAM Review 45,167-256.

Newman M.E.J., Barabási A.-L., Watts D.J., 2006. The Structure and Dynamics of Networks. Princeton University Press, Princeton, NJ, USA.

Newman M.E.J., Peixoto T.P., 2015. Generalized Communities in Networks. Physical Review Letters 115, 08870. American Physical Society.

Wei W. W. S., 2006. Time Series Analysis Univariate and Multivariate Analysis. Pearson.

Zanin M., Lacasa L., Cea M., Crístobal S., 2009. A dynamical model for the air transportation network. Proceedings 23rd European Conference on Modelling and Simulation.

Zanin M., Lillo F., 2013. Eur. Phys. J. Spec. Top. 215:5. doi:10.1140/epist/e2013-01711-9