'Phase transitions' in time series for arriving and household domestic tourism statistics

Sergey Borisenok

Department of Electrical and Electronics Engineering, Faculty of Engineering,

Abdullah Gül University, Sümer Campus, Kayseri 38080, Turkey

sergey.borisenok@agu.edu.tr

Abstract. The normalized time series over arriving and household domestic statistics for different types of expenditures (package tours, accommodation, food-beverage, health, tour services, GSM roaming, clothes-shoes, souvenirs, carpets-rugs) in the period 2003-2016 in Turkey touristic sector have been investigated by the application of nonlinear statistical analysis (entropy evaluation via Hurst exponents, fractal Hausdorff dimension in the state space). We demonstrate the existence of 'phase transitions' for the particular expenditures in different time periods. Our approach opens a gate for studying chaotic components in touristic time series and for developing nonlinear dynamical model in the multi-dimensional state space of touristic expenditure variables.

Keywords: Tourism, time series, nonlinear statistical analysis.

1 INTRODUCTION

Time series analysis is a powerful tool applied in many areas of natural and social sciences, including mathematical modeling of tourism statistics. As it was shown recently, particularly for tourist demand data, two approaches - autoregressive integrated moving average (ARIMA) and Chen's fuzzy time series (FTS) - outperform self-exciting threshold auto-regressions (SETAR), multivariate adaptive regression splines (MARS), multiple linear regression (MLR) and ANN models in most cases (Lim, Wang, 2005; Hanafiah, Harun, 2010; Lin, Chen, Lee, 2011; Lee, Nor, Suhartono, et al., 2012; Tularam, Wong, Nejad, 2012; Claveria, Torra, 2014; Akuno, Otieno, et al., 2015; Baldigara, Koić, 2015; Yılmaz, 2015), although there were some counterarguments against ARIMA for long-time perspective forecasts (Calantone, di Benedetto, Bojanic, 1988; Kon, 2002). This difference is statistically significant and varies from 40% up 70% (Claveria, Torra, 2014). FTS method has some advantage over ARIMA: it reduces the computation time and provides a distinct extra accuracy (Tsaur, Kuo, 2014).

Other important approaches for mathematical modeling involve modularity analysis and networks (Baggio, Scott, Cooper, 2010; Baggio, Sainaghi, 2016); complex dynamics (Baggio, 2013); topology of complex networks (Shih, 2006; Miguens, Mendes, 2008). The effect of external factors and their influence on tourism is discussed in (Proença, Soukiazis, 2005), while the important particular case of the correlation between time series for tourism and economic growth has been studied in (Bayramoğlu, Arı, 2015).

The review of models, including standard econometric and fuzzy time series, artificial neural networks (ANNs), rough set theory classificatory analysis and genetic algorithms one can find in (Song, Li, 2008). The specific case of Turkey has been described in (Aslan, Kaplan, Kula, 2008; Aktürk, Küçüközmen, 2012; Yılmaz, 2015).

1.1 Tourism time series as multidimensional dynamical system

In this paper we present a new paradigm to describe discrete-time step tourism time series in the terms of a multi-dimensional chaotic continuous-time dynamical system. Its variables represent different touristic expenditures normalized with the general tour expenditures to provide the same dimensionless units for all dynamical variables. They depend each on other and may be interpreted as a discretization of the continuous trajectory in multidimensional phase space. In the different time intervals of its evolution this trajectory has different fractal dimensions (reflecting its chaotic features) and other statistical characteristics that can be described as transitions between different statistical 'phases'.

The 'phase transitions' of such types reflect a complex combination of periodic and chaotic trends; and they cannot be restored from the linear analysis of the time series. The basic properties of the dynamical system could be derived from its Fourier power spectra and Hurst exponents as time-frequency domain analysis and fractal (Hausdorff) dimension for their topological features in the phase space.

1.2 Statistical data used for the model

In our numerical analysis we used time series over arriving and household domestic statistics for different types of expenditures (package tours, accommodation, food-beverage, health, tour services, GSM roaming, clothes-shoes, souvenirs, carpets-rugs) in the period 2003-2016 in Turkey touristic sector following the data presented on the www.turkstat.gov.tr website. Those series have the season time discretization, i.e. one time step corresponds to 3 months. Definitely, one can observe the clear time scale reflecting the season trends. Additionally, the absolute values for time series vary from year to year depending on the external factors.

To minimize the factors mentioned above we normalize here the particular expenditures (package tours, accommodation, food-beverage and others) with the general tour expenditures and, thus, we reproduce the dimensionless time series showing the rates of the particular expenditures, see the typical example series on Fig.1 for four variables: food and beverage (FB), accommodation (Ac), souvenirs (So) and organized tour services (TS) normalized with the general tour expenditures (TE). Below we describe the case of arriving citizens statistics in more details, similar results are observed for the household domestic tourism statistics.

On Fig.1 one year corresponds to 4 time steps, such that the season trends must have a period 4. Nevertheless, we see that for the normalized series the periodicity of trends does not follow only the year period 4, but can include other components varying for different expenditures and for different years. It may include the deterministic chaotic components as well, because the total dimension of the modeling dynamical system is high: to possess deterministic chaos a time-continuous dynamical system must have its dimension $D = 3$ or greater (Medio, Lines, 2003).

2 POWER SPECTRUM AND HURST EXPONENTS OF TOURISM TIME SERIES

2.1 Fourier power spectrum

Fourier power spectrum is the basic tool to analyze the typical features of the time series (Leon, 2012). For the touristic time series plotted on Fig.1 in dimensionless variables it is represented on Fig.2.

One can see that the expected one-year periodical trend is not the only and leading one, and there are also some chaotic components.

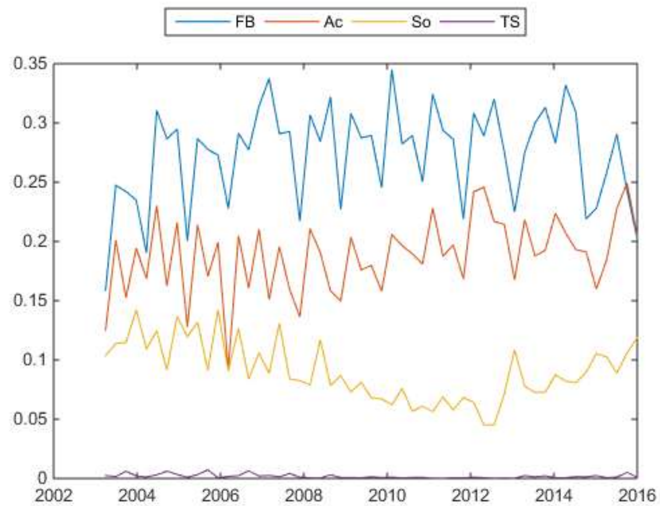

Fig.1. Typical time series for four types of expenditures in normalized units.

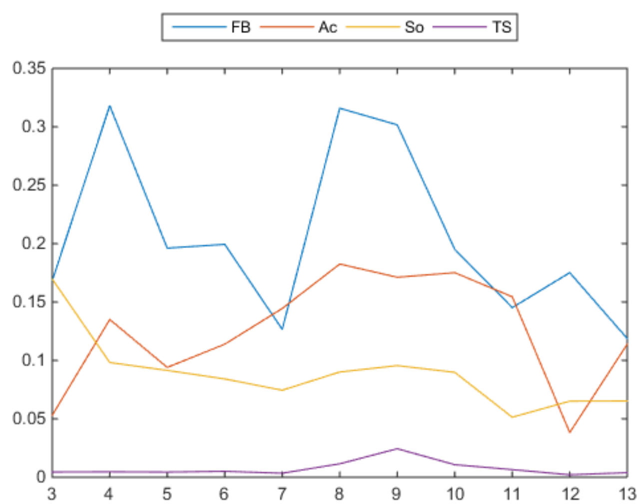

Fig.2. Dimensionless Fourier power spectrum for the time series presented on Fig.1.

2.2 Hurst exponents

As a rule, the q-th order correlation persistent follows the scaling law (Morales, Di Matteo, *et al*., 2012):

$$\frac{\left\langle \left| S(t+\tau) - S(t) \right|^q \right\rangle}{\left\langle \left| S(t) \right|^q \right\rangle} \propto \tau^{qH(q)} , \qquad (1)$$

where τ can vary between 1 and the maximal interval, and the brackets <…> denotes the average over the time-window. The particular case of q = 2 defines the autocorrelation function. If time series are not multifractal, the scaling parameter H is a constant, and it is called the generalized *Hurst exponent*. The fractal dimension d of the time series can be restored with the Hurst exponent as $d = 2 – H$.

The Hurst exponent defined by Eq.(1) provides a good tool to track unstable periods in time series. For instance, the sharp increasing of its fluctuation corresponds to a ‘phase transition’ regime, like financial crisis in economy (Morales, Di Matteo, *et al*., 2012). For its numerical evaluation we use here the Average Wavelet Coefficient (AWC) method (Weron, 2014) with the time window shift. The window interval is equal to 30 points. The corresponding fractal dimensions of the expenditure time series over the average Hurst exponents are represented on Fig.3. On this plot we cannot cover the epochs before 2006 and after 2012 due to the large half-size of the sliding window (around 4 years).

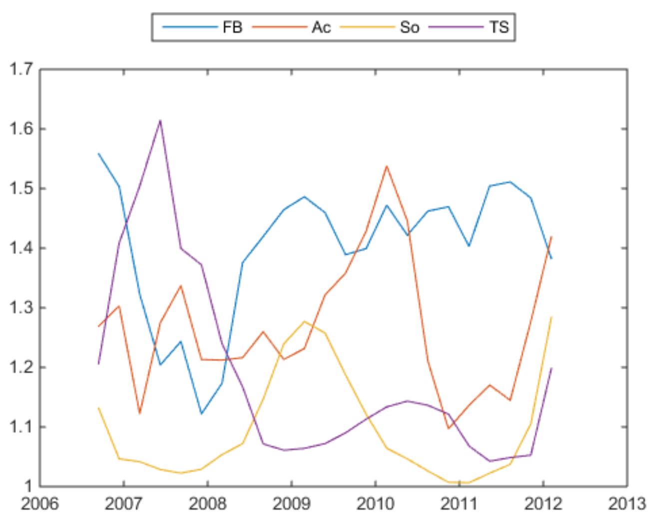

Fig.3. Time-dependent averaged fractal dimensions $d = 2 – H$ calculated over the Hurst exponents for the time series presented on Fig.1.

On Fig.3 one can clearly see the ‘phase transitions’ in the touristic time series. For instance, after the year 2008 we observe the increasing of chaotization for food-beverage expenditures (their Hausdorff dimension becomes more fractal) and the decreasing of chaotic features for the organized tour services (their dynamics follow more regular behavior). The dynamics of buying souvenirs is oscillating between almost regular easily predictable regime in 2007-2008 and in 2011 and distinctly irregular in the period 2009-2010.

3 DATA CLOUDS AND THEIR TOPOLOGICAL DIMENSIONS

A data cloud is the phase space representation of time series, where each independent parameter (like FB, Ac, So, TS on Fig.1) for the same time moment corresponds to a Cartesian coordinate of the cloud point in the multi-dimensional (4-dimensional in the case of Fig.1) space. For the sake of visualization those data clouds are projected on the planes of 3-dimensional hyperplanes in phase space coordinates (see Fig. 4).

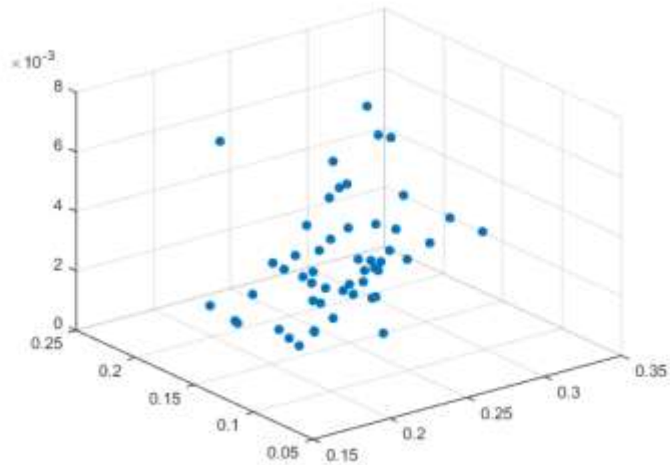

Fig.4. Data cloud in the coordinate space of normalized variables (FB, Ac, TE) for the time series presented on Fig.1.

In general case, the distribution of points in the cloud is fractal, and it can be evaluated numerically with the Hausdorff dimension. This topological characteristic of the data cloud is very stable under the perturbation of the external noise, as it was shown numerically for the EEG data without selection of artifacts (Borisenok, 2015).

If we consider the data cloud as a bounded set with the minimum number N of multi-dimensional covering balls of radius ε, then the Hausdorff dimension d_H is given by (Mandelbrot, 1982):

$$d_H = -\lim_{\varepsilon \to 0} \frac{\ln N(\varepsilon)}{\ln \varepsilon}, \qquad (2)$$

For non-fractal objects the Hausdorff dimension (2) is a non-negative integer. Particularly, for the 3-dimensional distribution on Fig.4 by the numerical *box counting procedure* we get the fractal dimension d_H = 1.6 that again proves the chaotic character of the corresponding nonlinear dynamical system. We cannot evaluate efficiently Eq.(2) for the shorter time intervals due to the small number of points in the cloud.

4 CONCLUSIONS

- The presence of chaotic components in the touristic time series together with 'phase transitions' among different dynamical regimes makes impossible modeling these data with one type of fitting, moreover with the linear one.
- The approach presented in the paper is a beginning step into developing new dynamical model for tourism expenditures in Turkey. First, it must be reformulated for a more precise time series (with the time step of one month at least, or even weekly) and supported by the analysis of statistical (in)dependency of different expenditures to clarify the basic number of independent components. It will help to define the minimum dimension of the corresponding dynamical system.

- This model based on the nonlinear statistical analysis is able to track not only linear components in the season trends, but it covers also nonlinear and deterministic chaotic components.
- It can be also reformulated for the case of time series in other areas of social sciences.

References

Aktürk, T., Küçüközmen, C. C. (2012). Tourism demand for Turkey: Models, analysis and results. http://www.coskunkucukozmen.com/wp-content/uploads/2012/02/g5.pdf

Akuno, A. O., Otieno, M. O., Mwangi, Ch. W., Bichanga, L. A. (2015). Statistical models for forecasting tourists' arrival in Kenya. *Open Journal of Statistics*, *5*, 60-65.

Aslan, A., Kaplan, M., Kula, F. (2008). International tourism demand for Turkey: A dynamic panel data approach. https://mpra.ub.uni-muenchen.de/10601

Bayramoğlu, T., Arı, Y. O. (2015).The relationship between tourism and economic growth in Greece economy: a time series analysis. *Computational Methods in Social Sciences*, 3(1), 89-93.

Baggio, R. (2013). Studying complex tourism systems: a novel approach based on networks derived from a time series. Moscow: XIV April International Academic Conference on Economic and Social Development.

Baggio, R., Sainaghi, R. (2016). Mapping time series into networks as a tool to assess the complex dynamics of tourism systems. *Tourism Management*, *54*, 23-33.

Baggio, R., Scott, N., Cooper, Ch. (2010). Network science: a review focused on tourism. *Annals of Tourism Research*, *37*(3), 802-827,

Baldigara, T., Koić, M. (2015). Modelling the international tourism demand in Croatia using a polynomial regression analysis. *Turističko poslovanje*, *15*, 29-38.

Borisenok, S. (2015). Fractal properties of EEG with non-removed artifacts. Istanbul: 2nd International Conference on Recent Advances in Pure and Applied Mathematics (ICRAPAM 2015), 80.

Calantone, R. J., di Benedetto, A., Bojanic, D. C. (1988). Multimethod forecasts for tourism analysis. *Annals of Tourism Research*, *15*, 387-406.

Claveria, A., Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models. *Economic Modelling*, *36*, 220-228.

Hanafiah, M. H. M., Harun, M. F. M. (2010). Tourism Demand in Malaysia: A cross-sectional pool time-series analysis. *International Journal of Trade, Economics and Finance*, *1*(1), 80-83.

Kon, S. Ch. (2002). Forecasting tourist demand to Singapore: A modern time-series approach. Singapore: Victoria University.

Lee, M. H., Nor, M. E., Suhartono, Sadaei, H. J., Abd Rahman, N. H., & Kamisan, N. A. B. (2012). Fuzzy time series: An application to tourism demand forecasting. American *Journal of Applied Sciences*, *9*(1), 132-140.

Leon, C. (2012). The Singular Spectrum Analysis as a New Time Series Method in Tourism and Hospitality Research. 6th Advances in Hospitality and Tourism Marketing and Management conference. Guangzhou, China.

Lim, C., Wang, Y. (2005). A time series analysis of Chinese outbound tourism to Australia. International Congress on Modelling and Simulation MODSIM05, 2246-2252. Abstracts. Melbourne, Australia: University of Melbourne.

Lin, Ch., Chen, H., Lee, T. (2011). Forecasting tourism demand using time series, artificial neural networks and multivariate adaptive regression splines: Evidence from Taiwan. *International Journal of Business Administration*, *2*(2), 14-24.

Liu, X., Cai, Z., Chen, R. (2015). Functional coefficient seasonal time series models with an application of Hawaii tourism data. *Computational Statistics*, *30*(3), 719-744.

Mandelbrot, B. B. (1982). The Fractal Geometry of Nature. Henry Holt and Co.

Morales, R., Di Matteo, T., Gramatica, R., Aste, T. (2012). Dynamical generalized Hurst exponent as a tool to monitor unstable periods in financial time series. *Physica A: Statistical Mechanics and its Applications*, *391*(11), 3180-3189.

Medio, A., Lines, M. (2003). Nonlinear Dynamics. Cambridge, Cambridge University Press.

Miguens, J. I. L., Mendes, J. F. F. (2008). Travel and tourism: into a complex network. *Physica A: Statistical Mechanics and its Applications*, *387*(12), 2963-2971.

Proença, S. A., Soukiazis, E. (2005). Demand for tourism in Portugal: A panel data approach. Impresso na Secção de Textos da FEUC, Discussion Paper 29. Coimbra: Centro de Estudos da União Europeia.

Shih, H. (2006). Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tourism Management*, *27*, 1029-1039.

Song, H., Li, G. (2008). Tourism demand modelling and forecasting - A review of recent research. *Progress in Tourism Management*, *29*(2), 203-220.

Tsaur, R., Kuo, T. (2014). Tourism demand forecasting using a novel high-precision fuzzy time series model. *International Journal of Innovative Computing, Information and Control*, *10*(2), 695-701.

Tularam, G. A., Wong, V. S. H., Nejad, S. A. Sh. (2012). Modeling tourist arrivals using time series analysis: Evidence from Australia. *Journal of Mathematics and Statistics*, *8*(3), 348-360.

Weron, R. (2014). AWC_HURST: MATLAB function to compute the Hurst exponent using the Average Wavelet Coefficient (AWC) method. http://econpapers.repec.org/software/wuuhscode/m14002.htm

Yılmaz, E. (2015). Forecasting tourist arrivals to Turkey. *Tourism: An International Interdisciplinary Journal*, *63*(4), 435-445.

Sergey Borisenok is an associate professor at Faculty of Engineering, Abdullah Gül University (Kayseri, Turkey). He received his BS/MS degrees in physics from the Department of Statistical Physics, St. Petersburg State University (Russia) in 1993 and his PhD degree in theoretical and mathematical physics from the same university in 1996. He is the author of more than 30 journal papers and has published six books. His current research interests include nonlinear series analysis and control theory in natural and social sciences, mathematical models for quantum engineering, brain dynamics. He is a member of IPACS (International Physics and Control Society).