



## Effects of Zeros in Phase Space Reconstruction for Small and Large Solar Radiation Data Points during Wet and Dry Seasonal Modeling and Prediction

ADENIJI, AE

*Department of Physical Sciences, Bells University of Technology, Ota, Ogun State, Nigeria.*

*\*Corresponding Author Email: [emmaabidec@yahoo.com](mailto:emmaabidec@yahoo.com)*

**ABSTRACT:** The effect of zeros in the behaviour of nature system has been a major global concern which have been reported to bias the output of the analysis. This study examines the effect of zeros on small and large solar radiation data points in Nsukka from a nonlinear dynamic perspective. The solar radiation data used were collected from National Research for Space and Development Agency (NARSDA) and covers the period of two years (January 2012–December 2013). The influence of zeros on average mutual information method for delay time ( $\tau$ ), False nearest neighbour (FNN) for embedding dimension ( $m$ ), and phase space reconstruction is investigated by considering two different cases (one hour and five minutes interval for small and large data points on monthly basis respectively). The results reveal that the phase space trajectories of the raw and non-zero small data points for dry and wet seasons show evidence of an attractor in a well-defined region while raw and non-zero large data points have no attractor like shape but regular patterns and well-defined shapes are visible in dry and wet seasons. These imply low-dimensional and deterministic chaotic nature of the underlying dynamics of raw and non-zero data for small and large data points during wet and dry seasons. It is observed that there is little or no significant difference in the phase space reconstruction of raw and non-zero data for both small and large data points due to the low percentage of zeros in the time series data.

DOI: <https://dx.doi.org/10.4314/jasem.v25i3.26>

**Copyright:** Copyright © 2021 Adeniji. This is an open access article distributed under the Creative Commons Attribution License (CCL), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Dates:** Received: 12 December 2020; Revised: 26 January 2021; Accepted: 12 February 2021

**Keywords:** mutual information method, Phase space reconstruction, False nearest neighbour, chaotic nature.

Chaotic dynamical systems are ubiquitous in nature such as the tornado, stock market, turbulence, and weather. Their functions are different in different situations (Liu, 2010). Its nature has always been an inevitable problem in our efforts towards understanding and describing their evolutions. The significant variability of geophysical phenomena exhibits both in time and in space only further complicates this problem, since a particular mathematical model formulated (and found appropriate) for describing a phenomenon at one site and at one time-scale may not be appropriate for either another site or even at the same site at a different time-scale (Sivakumar, 2004). Dynamics of natural systems is often described by nonlinear equations. When those equations are unknown, we can reproduce the system dynamics through the reconstruction of the manifold from the time course of one of its variables (Packard *et al.*, 1980, Farmer and Sidorowich, 1987 and Ott, 1994). The most striking feature of chaos is the unpredictability of its future. This feature is usually called as the “sensitivity dependence on initial conditions” or “butterfly effect.” (Liu, 2010). The interaction between local Meteorological conditions such as temperature, humidity, precipitation (rain, snow, sleet, & hail), atmospheric pressure and wind

velocity with Solar radiation data constitute a strongly coupled multi-variable nonlinear system which is a difficult task to handle (Zeng, 2013 and Adeniji *et al.*, 2020). Geophysical, climatic, and meteorological observations over a period give us data in the form of time series. These are observations of physical quantities such as pressure, temperature, wind velocity, etc., at discrete time intervals. It is often found that due to various operational difficulties such time series contain some missing data. Often several consecutive values are found to be missing. Due to technical or maintenance issues, weather conditions, instrumental failures or apparatus errors throughout the data collection, human error during data entry, calibration process and/or a damage of data due to malfunctioning storing machinery, extended hydrometric data construction and organization become a hard task and, in time, gaps in the data set arise (Johnston, 1999, Goa, 2017, Tencaliec, 2017 and Peña-angulo *et al.*, 2019).

For instance, zero values are very common in meteorological parameters, sometimes for very long stretches of periods. They are also common even during rainy periods, especially when finer resolutions (e.g. hourly) are considered (Sivakumar, 2017). The presence of a large number of zeros (or any other

*\*Corresponding Author Email: [emmaabidec@yahoo.com](mailto:emmaabidec@yahoo.com)*

single value) in a time series may bias the outcomes of chaos methods, since the reconstructed hyper-surface in phase space will tend to a point (Tsonis *et al.*, 1994). As a result of this effect, this research work aimed at evaluating the effects of zeros in phase space reconstruction for small and large solar radiation data points during the wet and dry seasons for the purpose of modeling and prediction.

**MATERIALS AND METHODS**

In this study, solar radiation data used is collected from Nsukka station under the supervision of National Research for Space and Development Agency (NARSDA). The data used covers the period of two years starting from January, 2012 - December, 2013 with Campbell Scientific Automatic Weather Stations that generates real time data at five minutes update cycle throughout the day. The station lies between 6°51'28.14" latitude and 7°24'28.15"E longitude. The study area experiences two main climatic seasons yearly, namely; the rainy and dry season (Phil-Eze, 2004). The average daily minimum and maximum temperature of the area are about 23.3<sup>0</sup>C and 27<sup>0</sup>C respectively, (Inyang, 1978). The beginning and end of rainfall in the study area is always associated with violent thunderstorms (Anyadike, 2002).

Phase space reconstruction of the attractor is a powerful tool to investigate the natural phenomena in real systems. Embedding dimensions (*m*) are computed for each time series by using the false nearest neighbor method (FNN) (Kennel *et al.*, 1992). Average mutual information (AMI) is a well-known method for estimating the lag time (Fraser and Swinney, 1986). Time delays are estimated from the first minimum of the average mutual information (AMI) function. This time delay via AMI is calculated using equation (1).

$$I(x_t, x_{t+\tau}) = \sum_{x_t} \sum_{x_{t+\tau}} P(x_t, x_{t+\tau}) \log_2 \frac{P(x_t, x_{t+\tau})}{P(x_t)P(x_{t+\tau})} \quad (1)$$

Where the sum is extended over the total number of samples in the time series,  $P(x_t)$  and  $P(x_{t+\tau})$  are the marginal probabilities for measurements  $x_t$  and  $x_{t+\tau}$  and  $P(x_t, x_{t+\tau})$  is their joint probability. The optimal time delay  $\tau$  minimizes the value of the function  $I(x_t, x_{t+\tau})$  for  $t = \tau$ .

According to Takens' delay embedding theorem (Takens, 1981), given a single-variable series,  $d_t$  where  $t = 1, 2, \dots, N$ , a multi-dimensional phase space can be reconstructed as follows:

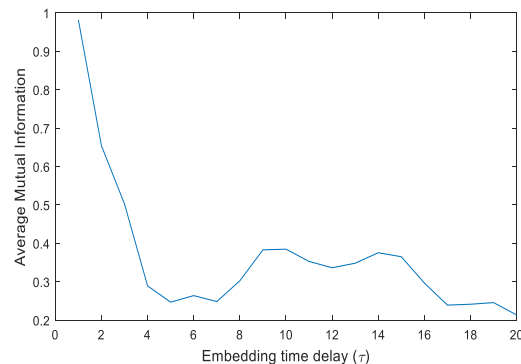
$$D_t = \{x_t, x_{t+\tau}, x_{t+2\tau}, \dots, x_{t+(m-1)\tau}\}, \quad (2)$$

$$t = 1, 2, 3, \dots, N$$

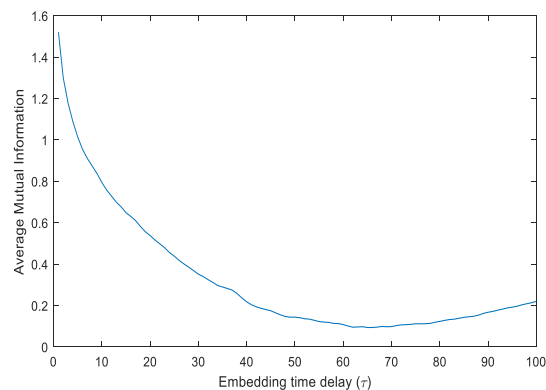
Where  $D_t$  is a vector of the solar radiation data of  $\{x_t\}_{t=1,2,\dots,N}$ .  $N$  is the number of recorded solar radiation data points,  $m$  is the embedding dimension and  $\tau$  is the time delay.

**RESULTS AND DISCUSSION**

Figures 1 and 2 show the time delay values obtained using the average mutual information (AMI) method (equation 1) for small and large data points of solar radiation data covering the period of two years. The delay times obtained is 5 for raw and non-zero small data points while 34 for raw and non-zero large data points of solar radiation data from January, 2012 to December, 2013 for each month. The average mutual information, which is a measure of how far the pairs of random variables of a time series data are dependent to each other gives the value of time lag,  $\tau$ . (Renjini *et al.*, 2020). This value must not be too small or large and also must be chosen with utmost care in order not to lose vital information about the dynamics of the system. In the AMI method,  $\tau$  is chosen to coincide with the first minimum of the mutual information (Fraser and Swinney, 1986).

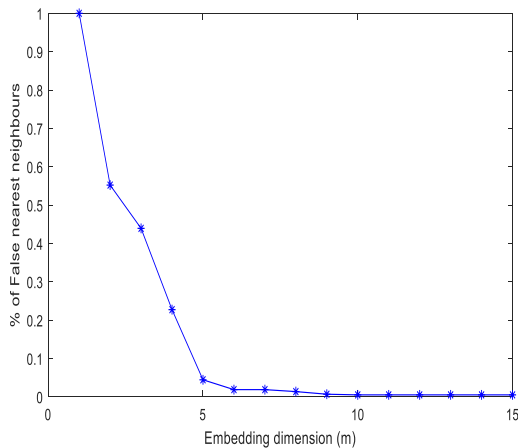


**Fig 1 :** Average mutual information (AMI) for raw and non-zero small data points.



**Fig 2:** Average mutual information (AMI) for raw and non-zero large data points.

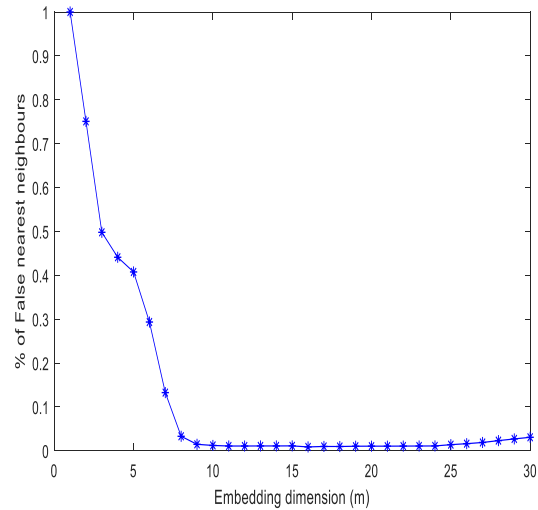
In order to determine the optimum embedding dimension  $m$ , false nearest neighbors (FNN) method proposed by Kennel *et al.*, 1992 was used. The FNN method is based on the idea that when the correct embedding dimension is determined, the percentage of false nearest neighbors (FNN) drops to zero. Figures 3 and 4 depict the false nearest neighbour of raw and non-zero data for both small and large data points of solar radiation. It is observed that any choice of dimension  $\geq 10$  and  $\geq 25$  are the suitable choices for raw and non-zero small and large data size, respectively. The higher embedding dimension is chosen from various samples to avoid under-embedding one sample than the other following the concept in Wallot, 2017 and Adeniji *et al.*, 2018. To take into account the issue of delay time on the FNN results, four different delay time values are considered in this study, especially to have varying intervals in the elements of the reconstructed phase space:  $\tau = 5$  hours (to represent raw and non-zero small data points for one hour interval for the period of one month) and  $\tau = 34$  minutes (to represent raw and non-zero large data points for five minutes interval for the period of one month).



**Fig 3:** False nearest neighbour (FNN) dimension for raw and non-zero small data points.

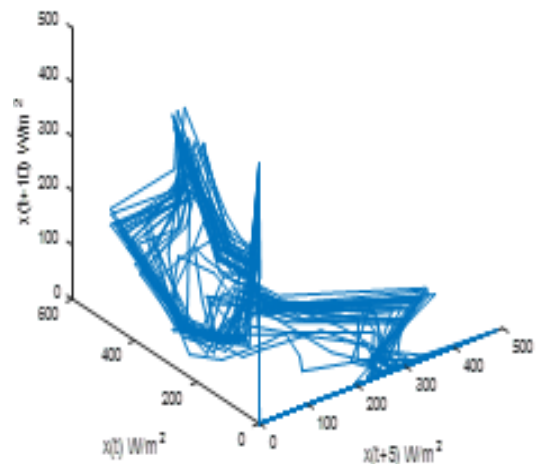
State space reconstructions of the raw and non-zero data for small and large data points during the wet and dry seasons are executed by using embedding technique (Mane, 1981) (equation 2). A representative of the phase space plots for 24 months of the two years is given in figures 5 – 8 and 9-12. The evolution of the system from some initial state is described by the trajectories of the phase space diagram which is assumed to be known, therefore representing the history of the system. The figures correspond to reconstruction in three dimensions with the projection of the attractor on the plane as  $\{x_i, x_{i+\tau}, x_{i+2\tau}\}$  with the time lag of 5 hours and 34 minutes for small and large

data points during the wet and dry seasons, respectively.

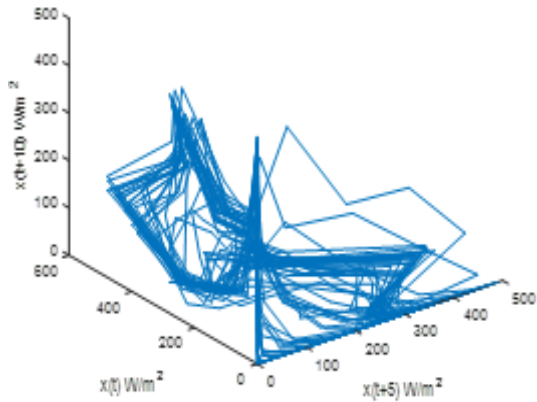


**Fig 4:** False nearest neighbour (FNN) dimension for raw and non-zero large data points.

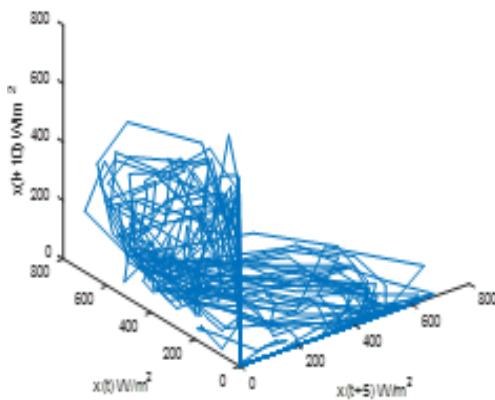
Figures 5 – 8 presents the phase space plots for raw and non-zero data for small data points during the wet and dry seasons. The trajectories in figures 5 and 6 representing the dry month shows presence of an attractor in a well-defined region and several loops close to each other. This indicating a ‘simple’ and ‘deterministic’ nature of the underlying dynamics and potentially of a low-dimensional and possibly chaotic system (Sivakumar, 2017). Similarly, figures 7 and 8 representing the wet month shows an attractor in a slightly well-defined region with the trajectories spreading out indicating low deterministic and high chaotic nature of the underlying dynamics. It is observed that there is no significant difference in the trajectories of the raw and non-zero data for the small data points.



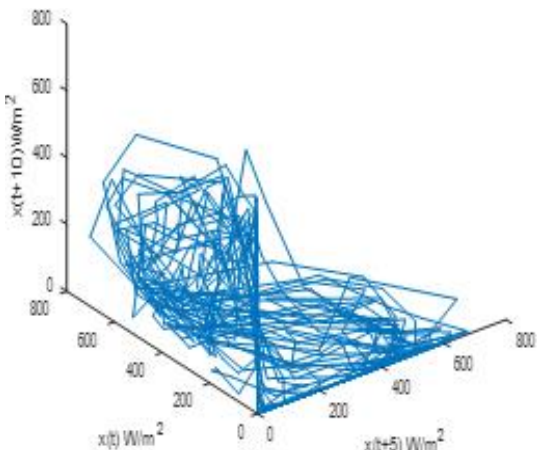
**Fig 5:** Phase space plots for raw small data points during the dry season.



**Fig 6:** Phase space plots for non-zero small data points during the dry season.



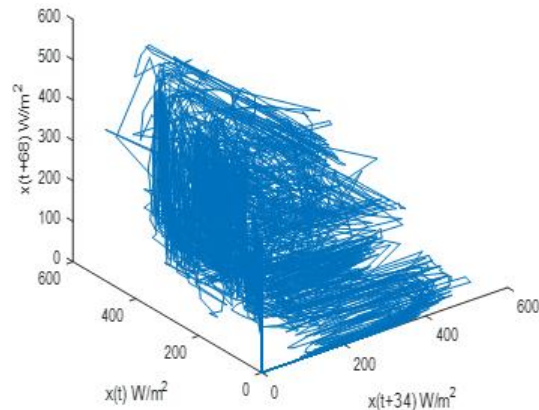
**Fig 7:** Phase space plots for raw small data points during the wet season.



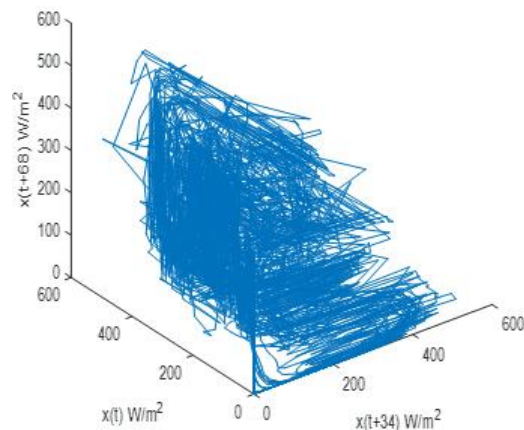
**Fig 8:** Phase space plots for non-zero small data points during the wet season.

Figures 9-12 present the phase space plots for raw and non-zero large data points during the wet and dry seasons. The trajectories in figures 9 and 10 (raw and non-zero data points for dry season) show evidence of more regular patterns and well-defined shapes than figures 11 and 12 (raw and non-zero data points for wet season) but no attractor like shape which is due to large data points causing the state space clustering

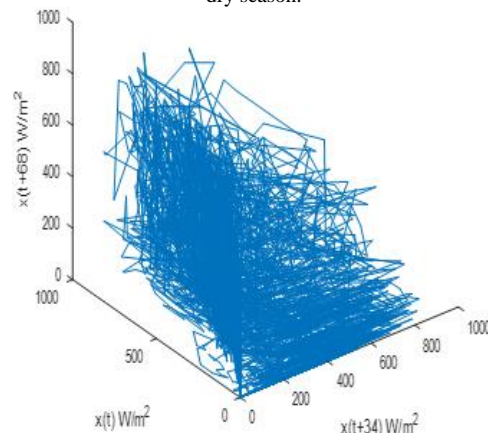
along the three axes. The patterns in Figures 9 and 10 suggest high deterministic and low chaotic nature of the underlying dynamics while figures 11 and 12 suggest low deterministic and high chaotic nature of the underlying dynamics. In general, trajectories in figures 5-12 show no significant difference for the raw and non-zero data for both small and large data points during wet and dry seasons.



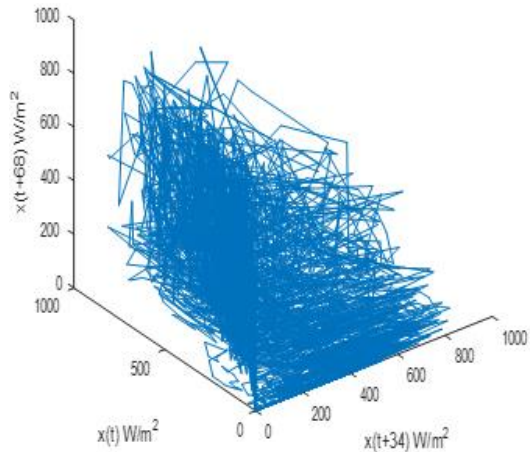
**Fig 9:** Phase space plots for raw large data points during the dry season.



**Fig 10:** Phase space plots for non-zero large data points during the dry season.



**Fig 11:** Phase space plots for raw large data points during the wet season.



**Fig 12:** Phase space plots for non-zero large data points during the dry season.

**Conclusion:** The work reported in the paper elucidates the potential of phase space reconstruction in revealing the effects of zeros in raw and non-zero data for small and large data points during wet and dry seasons with embedding dimension and time lag. A low-dimensional and deterministic chaotic nature of the underlying dynamics are observed. It was observed that wet month exhibit high deterministic chaotic behaviour than the dry month for the two years. No significant difference was observed for the raw and non-zero data for both small and large data points due to the low percentage of zeros in the time series data.

**Acknowledgement:** The author is grateful to National Space Research and Development Agency (NASRDA) for making their data available for this work. We sincerely thank the Centre for Atmospheric Research and their partners for promoting high standards of atmospheric observatory practice ([www.carnasrda.com](http://www.carnasrda.com)).

## REFERENCES

- Adeniji, AE; Njah, AN; Olusola, OI (2020). Regional and Seasonal Variation of Chaotic Features in Hourly Solar Radiation Based on Recurrence Quantification Analysis. *J. Appl. Nonlinear Dyn.* 9(2): 175-187.
- Adeniji, AE; Olusola, OI; Njah, AN (2018). Comparative study of chaotic features in hourly wind speed using recurrence quantification analysis. *AIP Adv.* 8: 2158- 3226.
- Anyadike, RNC (2002). Climate and Vegetation. In: Oformata G E K (ed). A survey of Igbo Nation. African Publisher Limited, Onitsha, p. 73-89.
- Farmer, JD; Sidorowich, JJ (1987). Predicting chaotic time series. *Phys. Rev. Lett.* 59(8): 845–848.
- Fraser, AM; Swinney, HL (1986). Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* 33: 1134–1140.
- Goa, Y (2017). Dealing with missing data in hydrology - Data analysis of discharge and groundwater time series in Northeast Germany. Freie Universität Berlin.
- Inyang, P E B (1978). Climate of Nsukka and Environs. In Oformata, G E K (ed). The Nsukka Environment. Enugu, Fourth dimension publishers.
- Johnston, CA (1999). Development and evaluation of infilling methods for missing hydrologic and chemical watershed monitoring data. Virginia Polytechnic Institute.
- Kennel, M B; Brown, R; Abarbanel, H D I (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* 45: 3403.
- Li, J; Kong, K; Cui, C; Zhang, Z (2020). Rainfall Data Reconstruction Based on Chaotic Characteristics of Meteorological Factors. *IOP Conf. Series: Earth and Environmental Sci.* 598: 012027.
- Liu, Z (2010). Chaotic Time Series Analysis. *Math. Probl. Eng.* p. 31. doi:10.1155/2010/720190
- Mane, R (1981). On the dimension of the compact invariant sets of certain nonlinear maps, Dynamical systems and turbulence (Lecture Notes in Mathematics) Springer-Verlag, Berlin, 898: 230.
- Ott, E; Sauer, T D; Yorke, J A (1994). Coping with Chaos: Analysis of Chaotic Data and the Exploitation of Chaotic Systems, John Wiley & Sons, New York, NY, USA.
- Packard, N H; Crutchfield, J P; Farmer, J D; Shaw, R S (1980). Geometry from a time series. *Phys. Rev. Lett.* 45(9): 712–716.
- Peña-angulo, D; Nadal-romero, E; González-hidalgo, J C; Albaladejo, J; Andreu, V; Bagarello, V; Barhi, H; Batalla, R J; Bernal, S; Bienes, R; Campo, J; Campobescós, M A; Canatarío-duarte, A; Cantón, Y; Casali, J; Castillo, V; Cerdà, A; Cheggour, A; Cid, P; Cortesi, N; Zorn, M (2019).

- Spatial variability of the relationships of run off and sediment yield with weather types throughout the Mediterranean basin. *J. Hydrol.* 571: 390 – 405.
- Phil-Eze, P O (2001). The vegetation of the Nsukka Plateau: Empirical Analysis of Forms and Patterns. In Nwafor J C; Bob- Duru, R C; Anyadike, R N C (ed). Geographical Report of the University of Nigeria, Nsukka. Enugu: JAMOE Enterprise, Enugu, Nigeria, p. 42-53.
- Renjini, A; Raj, V; Swapna, M S; Sreejyothi, S; Sankararaman, S (2020). Phase Portrait for High Fidelity Feature Extraction and Classification: A Surrogate Approach. *Chaos* 30:113122.
- Sivakumar, B (2004). Chaos theory in geophysics: past, present and future. *Chaos, Solitons Fractals.* 19: 441–462.
- Sivakumar, B (2017). Chaos in Hydrology Bridging Determinism and Stochasticity. Springer Science +Business Media Dordrecht.
- Takens, F (1981). Detecting strange attractors in turbulence. In Dynamical Systems and Turbulence, Warwick; Springer: Berlin/Heidelberg, Germany, p. 366–381.
- Tencaliec, P (2017). Developments in statistics applied to hydrometeorology: Imputation of streamflow data and semiparametric precipitation modeling. Universite Grenoble Alpes.
- Tsonis AA; Triantafyllou GN; Elsner JB; Holdzkom JJ; Kirwan Jr, AD (1994). An investigation on the ability of nonlinear methods to infer dynamics from observables. *Bull Amer Meteor Soc* 75:1623–1633.
- Wallot, S. (2017). Recurrence Quantification Analysis of processes and products of discourse: A tutorial in R. *Discourse Processes* (Taylor and Francis Group): *A multidisciplinary J.* 54: 382-405.
- Zeng, Z; Yang, H; Zhao, R; Meng, J (2013). Nonlinear characteristics of observed solar radiation data. *Solar Energy* 87: 204-218.