Proceeding Series of the Brazilian Society of Computational and Applied Mathematics
Preprint

Analysis of Rainfall Networks Generated from Different Similarity Measures

Iuri S. Diniz[1]
Federal University of Ouro Preto, Ouro Preto, MG, Brazil
Leonardo B. L. Santos [2]
Center for Monitoring and Early Warning of Natural Disasters, São José dos Campos, SP, Brazil
Leonardo N. Ferreira [3]
Max Planck Institute for Human Development, Berlin, Germany
Vander L. S. Freitas [4]
Federal University of Ouro Preto, Ouro Preto, MG, Brazil

Network science allows the investigation of several complex systems. From a set of nodes (vertices) and links (edges), it is possible to abstract the characteristics of these systems and the interactions between their components, enabling, for example, the study of the dynamics of meteorological systems via their time series on global [4], and local [2, 5] scales.

The climate networks generation process involves the transformation of spatiotemporal climate data into networks. One possible way of transformation involves the extraction of a set of time series with well-defined geographic positions that gives rise to the nodes of a complex network. The next two steps are [4]: to evaluate the similarity between the pairs of time series (network nodes) and to quantify the similarity thresholds that allow for connections (definition of a correlation threshold). Considering the first step, some authors use Pearson Correlation [2, 5] and Mutual Information [3]. Ferreira et al. [4] investigated 29 measures of similarity between global near-surface air temperature series. The definition of a threshold, on the other hand, depends on the evaluation of network metrics since very restrictive thresholds will only maintain the links between highly correlated series (few links), and less restrictive thresholds will impair the identification of possible patterns in the network [3].

The aim of this work is to analyze the networks generated from the same rainfall data used by Ceron et al. [2] but using different ways to quantify the similarity between the series and, consequently, build different networks. In addition to Pearson Correlation (PC), we investigate Mutual Information (MI), Dynamic Time Warping (DTW)[1], and Event Synchronization (ES)[6].

Figure 1 presents an example of our process to build the meteorological network. On a reduced scale, there is a grid with a spatial resolution of $1km$. The overlapped points represent time series with a time resolution of 1 hour. By testing the similarity between all pairs of time series, it is possible to generate a network connecting the series whose similarity satisfies a threshold criterion.

The data used in this work have the same spatiotemporal resolution seen in Fig. 1. There are 1755 time series containing 240 data points each. As the nodes represent the time series, the networks have 1755 nodes; the threshold applied to each network seeks to maintain a link density of 0.01, consistent with the PC network [2]. Our results show that the PC and the MI networks share similar metrics, such as the average path length: 8.12 and 8.79; clustering coefficient: 0.59 and 0.54; number of components: 29 and 28 and size of the giant component: 1675 and 1660. The

[1]iuri.diniz@aluno.ufop.edu.br
[2]santoslbl@gmail.com
[3]ferreira@leonardonascimento.com
[4]vander.freitas@ufop.edu.br

ES network presents a similar clustering coefficient of 0.58 but has a higher level of fragmentation (121 components). The DTW network is the one with the lowest average path length (6.05) and clustering (0.48), but it has the highest fragmentation (578 components). Ceron et al. [2] showed that the community structures in the network and the land use (e.g., vegetation, farming, urban area) in the region are correlated, which might be helpful when drawing disaster risk scenarios. Thus, our next step is to assess the correlations between land use and community structures in the obtained networks. We hope to use the constructed networks as a contribution to describe the propagation behavior of the rainfall events analyzed here.

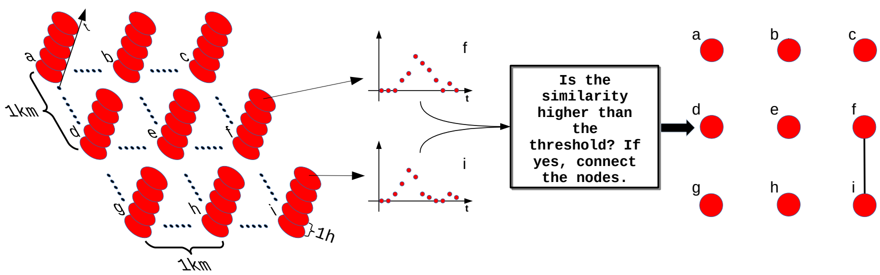

Figure 1: The methodological approach used to construct climate networks.

Acknowledgements

Iuri S. Diniz and Vander L. S. Freitas thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Grant number 122625/2020-0.

References

[1] Donald J Berndt and James Clifford. "Using dynamic time warping to find patterns in time series." In: **KDD workshop**. Vol. 10. 16. Seattle, WA, USA: 1994, pp. 359–370.

[2] Wilson Ceron et al. "Community detection in very high-resolution meteorological networks". In: **IEEE Geoscience and Remote Sensing Letters** 17.11 (2019), pp. 2007–2010. DOI: 10.1109/LGRS.2019.2955508.

[3] Jonathan F Donges et al. "Complex networks in climate dynamics". In: **The European Physical Journal Special Topics** 174.1 (2009), pp. 157–179. DOI: 10.1140/epjst/e2009-01098-2.

[4] Leonardo N Ferreira et al. "The effect of time series distance functions on functional climate networks". In: **The European Physical Journal Special Topics** 230.14 (2021), pp. 2973–2998. DOI: 10.1140/epjs/s11734-021-00274-y.

[5] Aurelienne AS Jorge, Izabelly C Costa, and Leonardo BL Santos. "Geographical Complex Networks applied to describe meteorological data". In: **Proceedings XXI GEOINFO**. 2020, pp. 258–263.

[6] R Quian Quiroga, Thomas Kreuz, and Peter Grassberger. "Event synchronization: a simple and fast method to measure synchronicity and time delay patterns". In: **Physical review E** 66.4 (2002), p. 041904. DOI: 10.1103/PhysRevE.66.041904.