Geophysical Research Abstracts
Vol. 21, EGU2019-9988, 2019
EGU General Assembly 2019

K-gaps: clustering regional climates from uneven time series

Fernando Jaume Santero (1,2), Leopoldo Carro Calvo (3), Ricardo García Herrera (1,2), David Barriopedro Cepero (2), and Sancho Salcedo Sanz (4)

(1) Department of Earth Physics and Astrophysics, Universidad Complutense de Madrid, Madrid, Spain, (2) Geosciences Institute (IGEO), (CSIC/UCM), Madrid, Spain, (3) Department of Signal Theory and Communications, Universidad Rey Juan Carlos, Fuenlabrada, Spain, (4) Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, Madrid, Spain

With the increasing availability of climate data during the last few decades, the characterization of regional climates by means of clustering techniques has provided important insights into the complex diversity of Earth's climate system. Although classical clustering algorithms have a good performance under continuous monitoring of well sampled regions, they are not usually suited for non-homogeneous data, hampering the regionalization of long-term climate records.

Within this context, k-gaps, a novel clustering technique has been developed to perform robust climate regionalizations with discontinuous time series covering different temporal periods. Its structure is based on widely used methods such as the k-means algorithm, with some particular differences that allow for the selection and attachment of records with different temporal lengths. The skill of k-gaps in clustering time series has been validated with 500 synthetic case studies replicating realistic spatial distribution of sparse climate records.

Here, we show two main applications retrieved from the clusterization of uneven time series using the k-gaps algorithm. On one side, due to the use of the entire temporal length of each record k-gaps allows for the reconstruction of long-term regional climate trends beyond the scope of classical techniques. On the other hand, the clusterization of time series according to their variance allows for the detection of regions where extreme climate events took place decades to centuries ago. The results identified similar climate regions when compared with complete data clusterizations, revealing the good performance of k-gaps for regional analyses with sample-starved historical and paleoclimate datasets. Therefore, this new clustering technique brings a new approach to study and characterize the climate of the past from sparse climate datasets.