Geophysical Research Abstracts
Vol. 19, EGU2017-16748, 2017
EGU General Assembly 2017

Data driven analysis of rain events: feature extraction, clustering, microphysical /macro physical relationship

Mohamed Djallel Dilmi, Cécile Mallet, Laurent Barthes, and Aymeric Chazottes
LATMOS/IPSL, UVSQ Université Paris-Saclay, UPMC Univ. Paris 06, CNRS, Guyancourt, France

The study of rain time series records is mainly carried out using rainfall rate or rain accumulation parameters estimated on a fixed integration time (typically 1 min, 1 hour or 1 day). In this study we used the concept of rain event. In fact, the discrete and intermittent natures of rain processes make the definition of some features inadequate when defined on a fixed duration. Long integration times (hour, day) lead to mix rainy and clear air periods in the same sample. Small integration time (seconds, minutes) will lead to noisy data with a great sensibility to detector characteristics. The analysis on the whole rain event instead of individual short duration samples of a fixed duration allows to clarify relationships between features, in particular between macro physical and microphysical ones. This approach allows suppressing the intra-event variability partly due to measurement uncertainties and allows focusing on physical processes.

An algorithm based on Genetic Algorithm (GA) and Self Organising Maps (SOM) is developed to obtain a parsimonious characterisation of rain events using a minimal set of variables. The use of self-organizing map (SOM) is justified by the fact that it allows to map a high dimensional data space in a two-dimensional space while preserving as much as possible the initial space topology in an unsupervised way. The obtained SOM allows providing the dependencies between variables and consequently removing redundant variables leading to a minimal subset of only five features (the event duration, the rain rate peak, the rain event depth, the event rain rate standard deviation and the absolute rain rate variation of order 0.5).

To confirm relevance of the five selected features the corresponding SOM is analyzed. This analysis shows clearly the existence of relationships between features. It also shows the independence of the inter-event time (IETp) feature or the weak dependence of the Dry percentage in event (Dd%e) feature. This confirms that a rain time series can be considered by an alternation of independent rain event and no rain period.

The five selected feature are used to perform a hierarchical clustering of the events. The well-known division between stratiform and convective events appears clearly. This classification into two classes is then refined in 5 fairly homogeneous subclasses.

The data driven analysis performed on whole rain events instead of fixed length samples allows identifying strong relationships between macrophysics (based on rain rate) and microphysics (based on raindrops) features. We show that among the 5 identified subclasses some of them have specific microphysics characteristics.

Obtaining information on microphysical characteristics of rainfall events from rain gauges measurement suggests many implications in development of the quantitative precipitation estimation (QPE), for the improvement of rain rate retrieval algorithm in remote sensing context.