

A Novel Fuzzy Ensemble Clustering Approach Using Latent Gaussian State Space Models for Demarcating Homogeneous Precipitation Zones

Debdut Sengupta¹, Sreeparvathy Vijay²

^{1,2}Department of Civil Engineering, Indian Institute of Technology, Madras
(IIT Madras)

ce23d036@smail.iitm.ac.in

Abstract

Precipitation clustering is the process of making homogeneous precipitation groups that exhibit similar spatiotemporal patterns. The most widely used feature-based clustering techniques often involve identifying key attributes that change the precipitation patterns in the time domain. The extracted features might not describe the temporal dynamics and the non-stationarity in the precipitation time series, limiting overall clustering accuracy, which is a major concern in the era of climate change. While studies found model-based time series clustering effective, most existing methods model the linear association and fail to understand the complex non-stationary behaviour fully. A novel non-stationary model based clustering framework is developed to delineate homogeneous precipitation zones, using an advanced Latent Gaussian State Space Model (LGSSM) coupled with fuzzy ensemble clustering techniques. The LGSSM consists of a state equation and an observation equation in which the state equation describes the signal generation mechanism and the observation equation describe the signal observation mechanism, making it effective to capture the inherent non-stationarity in the timeseries. Monte Carlo simulation experiments were conducted on multiple realizations of a synthetically generated dataset, replicating the real-world precipitation behaviour, to assess the robustness of the proposed methodology. A range of existing time series modeling frameworks (e.g., Hidden Markov Models and ARIMA) were evaluated in combination with both hard and fuzzy clustering techniques. The results indicate that the LGSSM, when integrated with the Expectation Maximization (EM) algorithm, consistently produces optimal parameter representations for different precipitation realizations, leading to higher clustering accuracy even for complex non-stationary cases, compared with other model-clustering configurations. Further, the integration of ensemble fuzzy clustering technique minimizes the uncertainties associated with initialization sensitivity and the hard boundary problem, which are critical when demarcating precipitation zones. The potential of the proposed methodology was further demonstrated through a case study using historical precipitation data of India.

Keywords: Climate Change, Homogeneous Precipitation Zones, Regionalization, State Space Modeling, Ensemble Clustering