EGU21-14519
https://doi.org/10.5194/egusphere-egu21-14519
EGU General Assembly 2021

Gaussianization for Multivariate, High-dimensional Earth Observation data Analysis

J. Emmanuel Johnson[1], Maria Piles[2], Valero Laparra[3], and Gustau Camps-Valls[4]
[1]Image Signal Processing Lab, University of Valencia, Valencia, Spain (Juan.johnson@uv.es)
[2]Image Signal Processing Lab, University of Valencia, Valencia, Spain (maria.piles@uv.es)
[3]Image Signal Processing Lab, University of Valencia, Valencia, Spain (valero.laparra@uv.es)
[4]Image Signal Processing Lab, University of Valencia, Valencia, Spain (gustau.camps@uv.es

Long-standing questions in multivariate statistics, information theory and machine learning reduce to estimating multivariate densities. However, this is still an unresolved problem and one of the biggest challenge in general, and for Earth system data analysis in particular, due to the high dimensionality (spatial, temporal and/or spectral) of the data streams. Gaussianization is a class of generative models (normalizing flows) that is effective in computing density estimates by using a sequence of composite invertible transformations which transform data from its original domain to a multivariate Gaussian distribution. The methodology in turn allows us to estimate information theory measures (ITMs), which are relevant for the analysis and characterization of Earth system data superseding the mean, variance and correlation, as higher order measures, thereby capturing more complexity and providing more insight into various problems. We show that our Rotation-Based Iterative Gaussianization (RBIG) method allows us to compute ITMs from multivariate (spatio-spectral-temporal) Earth data efficiently in both computation and memory terms, directly from the Gaussianizing transformation, while being robust to data dimensionality . We demonstrate how Gaussianization is useful in various Earth observation data analysis problems, from hyperspectral image analysis to drought detection in data cubes.