Institute for Advanced Simulation
Jülich Supercomputing Centre

IAS Seminar

Topic: **Identification of Climate Signals using Novel Robust Unsupervised Machine Learning Method**

Speaker: Velimir V. Vesselinov, PhD, Computational Earth Sciences Group (EES-16), Los Alamos National Laboratory, USA

Contents: The integration of large datasets and powerful computational capabilities has resulted in the widespread use of machine learning (ML) methods in many research areas including geosciences. ML analyses allow for rigorous analyses of large datasets and complex model outputs with unprecedented efficiency and details. Recently, we have developed a novel unsupervised ML method based on Non-negative Tensor Factorization coupled with custom clustering algorithm (called NTFk). NTFk allows for extracting meaningful and understandable features from multi-dimensional datasets. Here, we demonstrate applications of NTFk for various types of exploratory analyses related to earth sciences and hydrology.

For example, NTFk was applied to extracts interpretable features of model predicted air-temperature and water-table depth based on the climate simulations of Europe. NTFk identifies three dominant signals in the air-temperature tensor output: storm, winter, and summer. NTFk also identifies three dominant signals in the water-table tensor output: spring snowmelt, summer storms and seasonality fluctuations. NTFk also identifies additional temporal and spatial signals in the model predictions. The estimated signals are applied to find cross correlation between the model inputs and model predicted climate outputs.

NTFk applications are not limited to geosciences. NTFk can readily be applied to any observed or simulated datasets that can be represented as a tensor (multi-dimensional array) and have separable latent signatures or features. Frequently, these signatures/features are related to underlying physical processes governing the spatial and temporal variability in the data. Extracting these signatures and structure-preserving features can help the development of conceptual models, reduced-order models, and simplified closed-form mathematical expressions, which can then be used to predict the system behavior. Furthermore, NTFk analyses can be applied to detect anomalies or disruptions in the data, which might be caused by additional phenomena not detected in past. Lastly, unsupervised machine learning methods such as NTFk can be coupled with supervised machine learning methods to perform deep learning.

Time: Wednesday, 20 June 2018, 14:00

Venue: Jülich Supercomputing Centre, Hörsaal, building 16.3, room 222