MathDataLab postdoc project

Discovering and analyzing teleconnections with neural networks

The terrestial climate system involves a variety of temporal and spatial scales and complex feedback effects. As a result, the climate is highly nonlinear and many seemingly unrelated meteorological variables are connected in a very complicated fashion. Long-range spatial relations between meteorological variables are known as *teleconnections* and are very important for the understanding of climate dynamics. Often, teleconnections can neither be appropriately described by dynamical climate models, which are unable to resolve them due to limited spatial and temporal resolution, nor by standard statistical methods such as correlation analysis, due to assumptions of linearity, independence, and normality etc.

The objective of this project is find, analyze, and predict teleconnections using statistical methods, such as neural networks, with the overall aim of increasing predictability of the climate system. One statistical method that could be used in this context is nonhomogeneous hidden Markov models (HMM), see Hughes and Guttorp (1994), where the atmospheric state is assumed to belong to one of a finite number of different regimes. Each regime is equipped with a distinct probability distribution for the meteorological variables of interest. The transition probabilities between regimes are not held fixed, as for classical HMM, but are allowed to evolve in time as a function of a small number of large-scale atmospheric exogenous predictors. An alternative statistical method is provided by nonlinear methods for time series analysis, such as the SETAR-GARCH model with exogenous inputs, which assumes the dynamics underlying a time series to be state-dependent with a number of different regimes. Moreover, the noise of the system is assumed to be correlated over time. These two classes of methods can provide some insight to teleconnections, but are rather restrictive in that they assume a certain type of nonlinearity and/or a given number of fixed atmospheric states.

Neural networks offer a natural alternative to these models, due to their inherent nonlinear connections and flexibility, and have been shown to be skillful at encoding subtle, nonlinear, statistical relationships between multiple variables. In this project, we will develop neural networks that are tailor-made to be trained on meteorologic time series and use these networks to discover yet unknown relations between different meteorological variables. More specifically, we will investigate the following two meteorological questions:

- By what mechanisms does the current rapid warming of the Arctic affect the climate in the midlatitudes?
- By what mechanisms are typical climate indices such as the El Niño-Southern Oscillation or the North Atlantic Oscillation related to observations of other meteorological variables?

As a starting point, these two questions will be investigated with nonlinear methods for time series analysis. The next step is then to extend the analysis using neural networks. The results obtained and the performance of the two classes of methods will be compared and evaluated in terms of significance, robustness, interpretability, and ability to reproduce extremes. The neural networks used in these investigations can also be augmented by pattern recognition and feature extraction methods to allow for identification of additional useful relations.

Causality of the relations found will be investigated by the cross convergent mapping index described in Sugihara et al. (2012), which has been shown to be related to the ability of neural networks to learn connections between meteorological variables, see e.g. Hirata et al. (2016).