Title: Fusing Spatio-Temporal Data to Estimate Streamflow via a Bayesian Network

Abstract:
Estimating parameters that exhibit complex spatial and temporal dependencies using multi-modal measurements (features) poses a significant challenge. One important context in which this challenge arises is in the prediction of streamflow in watersheds, in which numerous measurements of the underlying system are collected from different sources and comprise different domains (spatial, temporal, and spatio-temporal). Most data fusion algorithms explored to date work with data samples belonging to a single "family" possessing one-to-one sample mappings, such as images from a given set of satellites and hence cannot be directly applied here. A robust framework is therefore needed to effectively merge multi-modal datasets measured over different domains.

We develop an estimation framework based on Bayesian networks that provides a mechanism to incorporate a variety of spatial, temporal and spatio-temporal features, and in doing so maintains a manageable computational complexity. This network is used to estimate streamflow throughout the Santa Fe River Watershed in North-central Florida. Optimal graph topologies for Bayesian networks are traditionally obtained rigorously under assumptions of plentiful training data and computational time. This requirement is often the limiting factor in the use of graphical models in applications plagued either by the "sparse-data" problem, real-time constraints, or both. In this work, the graph topology is specified by employing simple physical models characterizing the underlying random process.

In addition, we evaluate the efficacy of our network via several uncertainty measures and test its performance under different training data scenarios. We are working towards a comprehensive uncertainty analysis framework that will allow hydrologists to investigate the impacts of the availability and density of various data types on streamflow estimation accuracy. This, in turn, will facilitate decision making on how to parsimoniously instrument watersheds and thus manage costs.

Keywords: Bayesian networks; hydrology; feature selection; information theory; spatio-temporal estimation

Challenges addressed: Climate variability and climate change factors impacting water resource sustainability

Issues addressed: Water availability and allocation