A Neural Granger Causality Inference Approach to Identify Meteorological and Socio-Demographic Drivers of Urban Water Demand in the Contiguous United States

Wenjin Hao, Andrea Cominola, Andrea Castelletti

Urban water demands across spatio-temporal scales are influenced by a variety of socio-demographic, climatic, and urban form determinants. Inferring the key determinants of urban water use is a prerequisite for accurately predicting future water demands to design informed demand management policies and pursue water security. Previous studies have predominantly relied on Input Variable Selection (IVS) techniques to determine key variables potentially influencing urban water use. However, most rely on methods that only reveal correlations or mutual information among variables without inferring any causal relationship.

In this study, we scrutinize predictive determinants with causal interpretability of urban water consumption throughout the Contiguous United States (CONUS). We utilize open access data of monthly municipal water consumption from 126 US cities gathered from 2010 to 2017, along with associated socio-demographic, meteorological, and building characteristics variables. Data sources include the U.S. Census Bureau, the American Community Survey, and the PRISM climate dataset. We implement a neural Granger model to discover causal relationships. Granger causality describes whether past values of a variable that varies in time could predict time-lagged values of another variable, assuming causal effects are ordered in time (i.e., cause before effect). We model urban water demand as a function of meteorological and socio-demographic determinants for both the CONUS as a whole and three climate regions, i.e., arid, continental, and temperate regions. Numerical results show that some meteorological variables, i.e., vapor pressure deficit and temperature, are the most relevant external determinants of water demand for the CONUS across all climate regions. Some socio-demographics and building characteristics variables, including population, household characteristics, labor force situation, and building construction years also contribute to the overall effect. The relevance of maximum temperatures is more pronounced in arid regions than in other areas. Water demand prediction models developed using only the drivers identified by Granger causal discovery achieve high accuracy for both the CONUS and the different climate regions (average R^2 equal to 0.85).