

Causality in Hydrometeorological Systems: Process Understanding and Predictions

Name: Vivek Kumar Yadav

Email: viveky@iisc.ac.in

Supervisors: Prof. Murray Peel, Senior Lecturer Keirnan Fowler, Prof. Dongryeol Ryu,
Asst. Prof. Bramha Vishwakarma (IISc Bengaluru)

Domain: Environmental Hydrology and Water Resources



ABSTRACT

Identifying the drivers of a process or phenomena is central to understanding and predicting its future state. Machine learning methods have gained popularity in hydrology; however, they primarily rely on correlations rather than causation. In this work, we apply cause effect (causal) discovery methods for identifying drivers of surface soil moisture, in a synthetic environment. We contrast the results with a non-causal method, Pearson correlation coefficient. Results show causal discovery reveals a smaller, more robust set of drivers, enabling more reliable predictions under changing climatic conditions such as drought.

INTRODUCTION

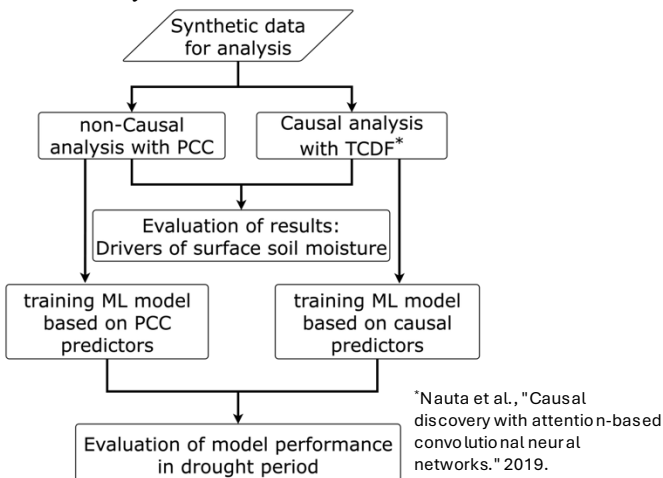
- Contemporary hydrological modelling increasingly relies on big data and machine learning.
- These methods rely on finding correlation between variables.
- However, correlation does not imply causation and such methods may fail under changing conditions like drought or climate change.

Research Questions

- Can the causal (cause-effect) drivers of a process be identified from observed data?
- Does finding causal drivers lead to better predictions? Especially under changing conditions like droughts or climate change

METHODOLOGY

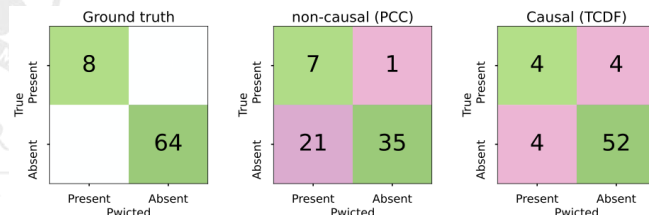
- We adopt a causal and a correlation-based method to identify the drivers of surface soil moisture.



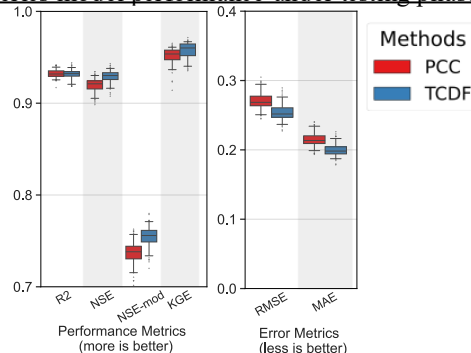
RESULTS & DISCUSSION

- The causal method identifies fewer false drivers of surface soil moisture compared to correlation-based method.
- Further, the causality-based prediction model performs satisfactorily under training phase and outperforms the correlation-based model under the testing phase of drought conditions.

Process understanding - Drivers of Surface Soil Moisture



Time-series model performance under testing phase



CONCLUSION

- Causality can help identify the drivers of a process.
- Identifying the causal drivers enables better predictions, especially under changing environments.
- Correlation based analysis suffers from high redundancy, which results in a poorer performance under changing conditions.