Utrecht University

Exploring Global and Local Water Scarcity Dynamics through Causal Discovery and Structural Causal Models

NATIONAL GEOGRAPHIC SOCIETY

Myrthe Leijnse[1], Marc F.P. Bierkens[1,2], and Niko Wanders[1]

1 Utrecht University, Department of Physical Geography, Utrecht, Netherlands, 2 Deltares, Utrecht, Netherlands

Figure 1 *The 21 water scarcity hotspots based on the underlying water gap values at 5 arcmin resolution obtained from PCR-GLOBWB 2 between 2010-2019. Hotspot boundaries are determined by water provinces with a water gap exceeding 0.015 m y⁻¹ (Leijnse et al., in press). The causal graph shows the results from the JPCMCI conditional independence test. The graphs on the sides show the true GRACE Equivalent Water Thickness (in cm) and the prediction by Lasso regression.*

Motivation

- **Water scarcity** is driven by **diverse natural and anthropogenic factors** representing a critical global challenge.
- **Structural Causal Models** are powerful tools to reveal the intricate **interactions among social, ecological and hydrological components** within human-water systems affected by water scarcity.
- This study integrates causal thinking into **statistical and data-driven hydrological modelling**, offering a different perspective on understanding system dynamics affecting water resources in water-scarce regions, the so-called water scarcity hotspots (Figure 1).
- **We aim to identify the most important global causal relations of water scarcity using causal discovery (JPCMCI)** and to predict GRACE (Equivalent water thickness) through **Lasso regression** on global and local (hotspot) scale.

Methods

- **JPCMCI** (Günther et al., 2023) - Causal discovery method to find the causal graph of (observational) variables related to water scarcity at global hotspots. JPCMCI includes latent (unobserved) variables and tests conditional independencies between each variable.
- **Data preprocessing** - **1)** Global gridded dataset timeseries **clipping** according to the respective hotspot area and calculation of the **zonal sum or average** of each variable. **2) Removal of monthly trends** (subtraction of long-term monthly mean).
- **Lasso regression** (Scikit-learn, 2011) – with optimal alpha is applied to the compiled datasets for each hotspots to predict GRACE equivalent water thickness. Evaluation metrics are found in Table 2.

Table 1 *Data inputs for JPCMCI and Lasso regression*

Variable	Dataset	Spatial resolution
Total precipitation	ERA5	0.5 degrees
2m temperature	ERA5	0.5 degrees
Cropland area	MODIS	500 m
Equivalent water thickness	GRACE	3 degrees
Population	Worldpop	1 km
LAI	MODIS	0.1 degrees
EVI	MODIS	0.5 degrees
Discharge	PCR-GLOBWB 2	5 arcmin

Conclusion

- JPCMCI: system memory, temperature and precipitation have causal link with GRACE equivalent water thickness at hotspots.
- Lasso regression: System memory, population and discharge most important variables to affect GRACE equivalent water thickness at hotspots.
- **Open for discussion: How to implement in local scale modelling and apply interventions?**

References

Günther, W., Ninad, U., & Runge, J. (2023, July). Causal discovery for time series from multiple datasets with latent contexts. In Uncertainty in Artificial Intelligence (pp. 766-776). PMLR.
Leijnse, M., Bierkens, M.F.P., Gommans, K.H.M., Lin, D., Tait, A., Wanders N. (in press). Key drivers and pressures of global water scarcity hotspots. *Environmental Research Letters*. DOI10.1088/1748-9326/ad3c54
Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Table 2 *Evaluation metrics Lasso regression*

Hotspot	Alpha	r²	MAE	MSE
Arabian Peninsula	0.003	0.983	0.4	0.3
California	0.053	0.940	1.8	5.5
Central Chile	0.019	0.521	2.1	8.9
Coastal Peru	0.095	0.647	0.8	1.1
Ganges	0.000	0.951	1.5	3.3
Greece	0.006	0.655	2	5.9
Indus	0.015	0.883	1.3	2.7
Iran	0.048	0.970	1	1.3
Italy	0.002	0.726	1.5	3.7
Japan	0.215	0.792	1.8	5.9
Java	0.000	0.620	1.8	4.7
Mexico	0.005	0.879	0.7	0.6
Murray-Darling	0.433	0.943	1.3	2.6
Nile Delta	0.017	0.785	0.4	0.3
North China Plain	0.152	0.916	1.1	2.5
Spain	0.003	0.686	1.8	4.7
Thailand	0.076	0.869	2.5	9.2
Turkey	0.007	0.886	1.3	3.1
US High Plains	0.008	0.899	1.1	2.2
Vietnam	0.006	0.668	2.4	9.7
White Nile	0.000	0.779	0.9	1.3

Missed out on chatting or want to know more? m.leijnse@uu.nl → | Check out the World Water Map yourself → | More on Global Water Gap Hotspots →