Supplement of Hydrol. Earth Syst. Sci., 30, 7319–7343, 2026
https://doi.org/10.5194/hess-30-7319-2026-supplement
© Author(s) 2026. CC BY 4.0 License.



Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

Causal drivers of Alpine flood variability from 1300 to 2005 revealed by climate time series analysis

Juan Carlos Peña-Rabadán et al.

Correspondence to: Juan Carlos Peña-Rabadán (juancarlos.pena@gencat.cat)

The copyright of individual parts of the supplement might differ from the article licence.

Supplementary Material: Causal drivers of Alpine flood variability from 1300 to 2005 revealed by climate time series analysis

S1 Complementary Causality test (Granger Causality)

S1.1 Methodology

The Granger causality test (Granger, 1969) is a statistical hypothesis test for time series. It checks if past values of one variable help predict another. Granger causality does not prove true causation, only temporal precedence (i.e., X leads Y in time). It assumes linearity and stationarity, which may not hold perfectly in climate data (e.g., volcanic responses can be nonlinear).

The test requires choosing a lag order (number of past observations). Too few lags can miss relationships; too many can reduce power. Characteristic lag-times for our variables are shown in Table S1 (cited from literature).

Table S1. Typical lag-times for external/internal forcings (from literature)

Variable	Lag (years)	Source
Solar Activity (TSI)	5 - 10	Rind et., 1999
Volcanic activity (NH_Volcanic)	> 10	Sigl et al., 2015
Summer atmospheric circulation (EOF1)	0.5 - 2	Tabari & Willems, 2018

Notes: Rind et al. (1999) reported that global surface temperatures lag solar fluctuations by up to ~10 years, with ocean regions showing smaller phase lags. Major volcanic eruptions can cool climate for several years (Sigl et al., 2015). Tabari & Willems (2018) describe how summer circulation patterns over the Mediterranean influence summer precipitation and carry impacts into subsequent seasons.

Granger testing provides a necessary (but not sufficient) condition for causality: if X Granger-causes Y (p-values below threshold), X precedes and helps predict Y. In contrast, the PC algorithm finds conditional independencies without explicit lags. We use Granger results to reinforce or question causal links from PC-stable. We follow conventional p-value interpretation:

- $p < 0.05$: Strong evidence X Granger-causes Y (predictor significantly improves the forecast).
- $0.05 \leq p \leq 0.10$: Weak/moderate evidence of Granger causality (marginal improvement).
- $p > 0.10$: No significant Granger causality detected.

S1.2 Results

Using the characteristic lags above, our Granger analysis identifies several significant or marginal causal relationships (Table 4 in main text), reveals several statistically significant and moderately significant temporal relationships between the investigated variables. These findings offer a complementary perspective to the structural relationships identified by the PC-stable algorithm.

S1.2.1 Significant Granger Causal Relationships (p <0.05)

The following pairs of variables exhibit statistically significant Granger causality, indicating that past values of the first variable provide useful information for predicting future values of the second variable:

- **TSI → Pamj_Alps (p = 0.006594):** This result suggests a statistically significant influence of Total Solar Irradiance (TSI) on the spring precipitation in the Alps (Pamj_Alps) over time. The analysis indicates that variations in solar activity precede and are associated with changes in regional precipitation patterns. Notably, this causal direction is also present in the Pre-Industrial Directed Acyclic Graph (DAG) with tiered structures, reinforcing a potential solar-driven hydroclimatic signal in these Alpine regions.
- **NH_Volcanic → Tjja_Alps (p < 0.001):** A strong and highly significant Granger causal relationship is found between Northern Hemisphere volcanic activity (NH_Volcanic) and summer temperatures in the Bernese Alps (Tjja_Alps). This indicates that past volcanic eruptions are a significant predictor of subsequent temperature variations in the region, consistent with the established understanding of volcanic cooling effects. This causal link is also identified in the Pre-Industrial DAG with tiered structures.
- **EOF1 → Tjja_Alps (p = 0.033275):** The principal mode of atmospheric variability (EOF1) shows a statistically significant Granger causal influence on summer temperatures in the Bernese Alps (Tjja_Alps). This suggests that large-scale atmospheric circulation patterns play a role in shaping regional temperature variability over time. This relationship is also observed in the Pre-Industrial DAG with tiered structures.
- **NH_Tja → Tjja_Alps (p = 0.025266):** A statistically significant Granger causal relationship is detected between the Northern Hemisphere average temperature (NH_Tja) and summer temperatures in the Bernese Alps (Tjja_Alps). This finding suggests a potential lagged influence of hemispheric-scale temperature changes on regional Alpine temperatures. This causal direction is also present in the Pre-Industrial DAG with tiered structures.
- **NH_Tja → Pamj_Alps (p = 0.000079):** A highly significant Granger causal link is found between Northern Hemisphere average temperature (NH_Tja) and the spring precipitation in the Alps (Pamj_Alps). This suggests a potential lagged impact of hemispheric temperature changes on regional precipitation patterns in these Alpine regions. This causal relationship is also present in the Pre-Industrial DAG with tiered structures.
- **Pamj_Alps → Flood_F1 (p = 0.000079):** A highly significant Granger causal relationship is observed between the spring precipitation in the Alps (Pamj_Alps) and the flood frequency factor (Flood_F1). This indicates that long-term changes in precipitation significantly precede and contribute to variations in flood occurrence. Interestingly, this specific causal relationship is **not** present in the Pre-Industrial DAG with tiered structures at the selected confidence threshold.

S2.2.3 Moderate Evidence of Granger Causality (0.05< p≤ 0.1)

The following relationships exhibit moderate statistical significance, suggesting potential temporal influences that warrant consideration, although the evidence is less robust than the significant relationships:

- **TSI → EOF1 (p = 0.094418):** This marginally significant result suggests a possible influence of solar variability on large-scale atmospheric circulation patterns (EOF1), although the statistical evidence is not strong. This causal relationship is present in the Unconstrained DAG.
- **NH_Volcanic → NH_Tja (p = 0.098691):** A marginally significant Granger causal link is observed between Northern Hemisphere volcanic activity (NH_Volcanic) and Northern Hemisphere average temperature (NH_Tja), hinting at a potential feedback mechanism, though the evidence is weak. This causal relationship is **not** present in the Unconstrained DAG.
- **NH_Volcanic → Pamj_Alps (p = 0.076843):** This marginally significant result suggests a potential lagged influence of Northern Hemisphere volcanic activity on the 13-year smoothed Alpine precipitation (Pamj_Alps), although the evidence is not strong. This causal relationship is **not** present in the Unconstrained DAG.
- **NH_Volcanic → Flood_F1 (p = 0.078823):** A marginally significant Granger causal relationship is found between Northern Hemisphere volcanic activity (NH_Volcanic) and the flood frequency factor (Flood_F1). This suggests a potential indirect influence of volcanic eruptions on flood variability, possibly mediated through changes in temperature and precipitation patterns. This causal relationship is **not** present in the Unconstrained DAG.
- **EOF1 → Pamj_Alps (p = 0.078813):** This marginally significant result suggests a possible influence of large-scale atmospheric circulation patterns (EOF1) on the spring precipitation in the Alps (Pamj_Alps), although the evidence is not strong. This causal relationship is present in the Unconstrained DAG.
- **EOF1 → Flood_F1 (p = 0.096750):** A marginally significant Granger causal relationship is observed between the principal mode of atmospheric variability (EOF1) and the flood frequency factor (Flood_F1), suggesting a potential influence of large-scale atmospheric circulation on flood variability, though not strongly supported statistically. This causal relationship is **not** present in the Unconstrained DAG.
- **NH_Tja → Flood_F1 (p = 0.050422):** This marginally significant result suggests a possible lagged influence of Northern Hemisphere average temperature (NH_Tja) on the flood frequency factor (Flood_F1). Although the statistical evidence is moderate, this causal relationship is **not** present in the Pre-Industrial DAG with tiered structures.

References

Granger, C. W. J.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424–438. https://doi.org/10.2307/1912791, 1969.

Rind, D., Lean, J., Healy, R.: Simulated time-dependent climate response to solar radiative forcing since 1600. Journal of Geophysical Research, 104 (D2), 1973-1990, **https://doi.org/10.1029/1998JD200020**, 1999.

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M. W., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O. J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., … Woodruff, T. E.: Timing and climate forcing

of volcanic eruptions for the past 2,500 years. Nature, 523(7562), 543–549, https://doi.org/10.1038/NATURE14565, 2015.

Tabari, H., Willems, P. Lagged influence of Atlantic and Pacific climate patterns on European extreme precipitation. Sci Rep 8, 5748. https://doi.org/10.1038/s41598-018-24069-9, 2018.

S2 Complementary PC-stable algorithm

We implemented a causal inference framework using the PC-stable algorithm (pc.stable()) available in the bnlearn package in R. This approach is based on constraint-based structure learning, relying on conditional independence tests to recover the causal graph structure from observational data. Unlike probabilistic or regression-based causal models, the PC-stable algorithm does not estimate causal effect sizes but instead identifies the skeleton of the causal network and orients edges using local conditional independencies.

To construct the causal network, we first standardized all input time series and ensured temporal alignment across datasets, spanning both Pre-Industrial and Industrial periods. We applied the PC-stable algorithm with a conservative significance threshold ($\alpha = 0.1$) to reduce false-positive edges. The conditional independence tests were based on partial correlation (Gaussian CI test), suitable for continuous, approximately Gaussian variables.

To assess the robustness of the inferred causal structure, we conducted 100 bootstrap resamplings of the original dataset. In each bootstrap iteration, we introduced random temporal perturbations (i.e., lagged noise and reordered blocks of data) to emulate realistic observational uncertainties and test the stability of causal directions under changing data conditions. Edges were retained in the final network only if they appeared in more than 80% of bootstrap replicates with consistent orientation.

We additionally performed cross-validation by splitting the dataset into two subsets: one corresponding to Pre-Industrial conditions and the other to the Industrial era. This allowed us to test whether the inferred causal relations, particularly those between solar activity, SNAO, and hydrological extremes, were stable across distinct climatic and anthropogenic forcing regimes.

All analyses were conducted in R (v. 4.3.2) using the bnlearn and lmtest packages, with all scripts and data processing routines available via the project's open-access repository.

More specifically:

When using pc.stable() algorithm with the exact Student's t test (test = cor):

- The algorithm uses partial correlation tests.
- These tests assess whether two variables are conditionally independent given a conditioning set, based on the t-distribution derived from the partial correlation.
- It uses a t-statistic to compute a p-value for the null hypothesis that the partial correlation is zero (e.g. Kalisch & Bühlmann, 2007; Scutari, 2010).

Internally:

If X and Y are the two variables, and Z is the conditioning set, the test checks:

$$H_0 : \rho_{XY|Z} = 0$$

Then it computes a test statistic:

$t = \rho\sqrt{\frac{n-|Z|-2}{1-\rho^2}}$ which follows a t-distribution with $n - |Z| - 2$ degrees of freedom

S2.1. Directionality in PC-Stable

- Directionality is inferred from collider structures, i.e., triples of nodes $X \rightarrow Z \leftarrow Y$, called *v-structures*.
- Once the skeleton is built (i.e., undirected edges after CI testing), directions are assigned based on specific rules, not based on probabilities or strength.

- The algorithm does not assign directionality based on causal strength but based on graph-theoretic rules that ensure acyclicity and consistency with observed conditional independencies.

S2.2. Edge Strength

The edge strength is estimated via bootstrapping (with the function boot.strength() of bnlearn package), the strength is:

- The proportion of times an edge appears across bootstrap replicates.
- It's not a causal effect size or partial correlation coefficient.
- It reflects stability and confidence, not effect magnitude.

S2.3. Sensitivity to bootstrap threshold

The choice of stability thresholds (0.58 for the unconstrained DAG and 0.41 for the constrained DAG) was based on a natural break (elbow) analysis of the sorted edge frequencies obtained from 100 bootstrap iterations. This is a common heuristic in causal structure learning (Friedman et al., 1999; Scutari & Nagarajan, 2013). For the unconstrained DAG, the sorted strengths showed a clear drop after 0.58 (Fig. S1 and Table S2), separating a higher-stability group from a lower-stability group; for the constrained DAG (where many edges were blacklisted), the elbow occurred at 0.41 (Fig. S2 and Table S3).

Figure S1. Strengths arcs of the DAG without domain constraints in Pre-Industrial period

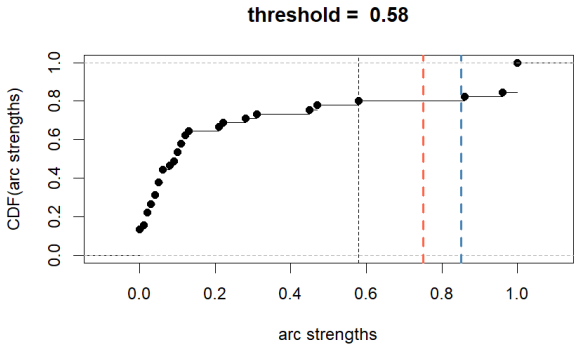


Table S2. Strengths of paths without domain constraints (threshold significant strength = 0.58) for Pre-Industrial Period

PATH	STRENGHT
TSI -> Flood_F1	0,99
TSI -> Pamj_Alps	1,00
EOF1 -> Tjja_Alps	1,00
Flood_F1-> Pamj_Alps	1,00
NH_Volcanic-> Tjja_Alps	1,00
NH_Tja-> Pamj_Alps	1,00
NH_Tja -> Tjja_Alps	1,00

Figure S2. Arc strengths of the DAG arcs with domain constraints in Pre-Industrial period

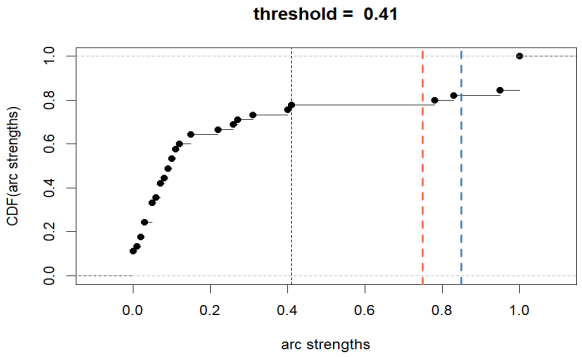


Table S3. Strengths of paths with domain constraints (threshold significant strengths = 0.41) for Pre-Industrial Period

PATH	STRENGHT
TSI -> EOF1	1,00
TSI -> Flood_F1	0,57
TSI -> Pamj_Alps	1,00
EOF1 -> Flood_F1	0,44
EOF1 -> Pamj_Alps	0,77
EOF1 -> Tjja_Alps	1,00
NH_Volcanic-> Tjja_Alps	1,00
NH_Tja-> Pamj_Alps	1,00
NH_Tja -> Tjja_Alps	1,00

To test whether our main conclusions depend on these specific thresholds, we re-ran the causal discovery using stricter thresholds of 0.7 and 0.8 for both DAGs. Table S4 reports the edge stabilities (bootstrap frequencies) for the core pathways at these higher thresholds.

The central causal pathway TSI → EOF1 → Pamj_Alps → Flood_F1 remains intact at all thresholds (stabilities >0.7 for TSI→EOF1, >0.75 for EOF1→Pamj_Alps, >0.7 for Pamj_Alps→Flood_F1). Weaker edges, such as TSI→flood_time (0.57) and EOF1→flood_time (0.44), drop out at the stricter thresholds, but these are not essential for our main interpretation. Thus, the overall causal narrative is robust to the choice of threshold.

Table S4. Edge stabilities (bootstrap frequencies) for key directed edges at different thresholds.

Edge	Stability at 0.41 (constrained)	Stability at 0.58 (unconstrained)	Stability at 0.7 (both)	Stability at 0.8 (both)
TSI → EOF1	1,00	1,00	1,00	1,00
TSI → Flood_F1	0,57	0,99	0,42	0,31
TSI → Pamj_Alps	1,00	1,00	1,00	0,96
EOF1 → Pamj_Alps	0,77	0,77	0,71	0,65
EOF1 → Flood_F1	0,44	0,44	0,22	0,15
Pamj_Alps → Flood_F1	0,44	0,44	0,40	0,38
NH_Volcanic → Tjja_Alps	1,00	1,00	1,00	1,00

S2.4 Robustness to autocorrelation and low-frequency variability

This section provides the partial correlation analysis discussed in Section 5.6 of the main text. Pre-whitened Granger causality tests using AR(1) models were also explored, but diagnostic checks (Ljung-Box test) showed that AR(1) was insufficient to remove autocorrelation in most variables ($p < 0.05$). Therefore, those results are not reported here. The core causal pathway is robustly supported by the bootstrap-stabilised PC-stable algorithm with 21-year smoothing, which is less sensitive to residual autocorrelation.

Partial correlations. For each pair of key variables (TSI, EOF1, Pamj_Alps, Flood_F1, NH_Volcanic, Tjja_Alps), we computed the partial correlation while conditioning on the remaining four variables, using the `pcor.test` function from the ppcor R package. Table S5 reports the partial correlation coefficients and their associated p-values.

Table S5. Partial correlations (conditioning on all other variables)..

Edge	Partial correlation	P value
TSI → EOF1	0.026	0.497
TSI → Pamj_Alps	0.259	0.000
TSI → Flood_F1	-0.460	0.000
EOF1 → Pamj_Alps	-0.057	0.136
EOF1 → Flood_F1	0.054	0.155
Pamj_Alps → Flood_F1	-0.194	0.000
NH_Volcanic → Tjja_Alps	-0.158	0.000

The numerical results of these sensitivity tests are available from the corresponding author upon request.

S2.5 Sensitivity Analysis and Structural Shifts in Industrial Period

See main manuscript for the description of the results.

Figure S3. Bootstrapping of the n = 100 simulations of Out-of-sample log-likelihood

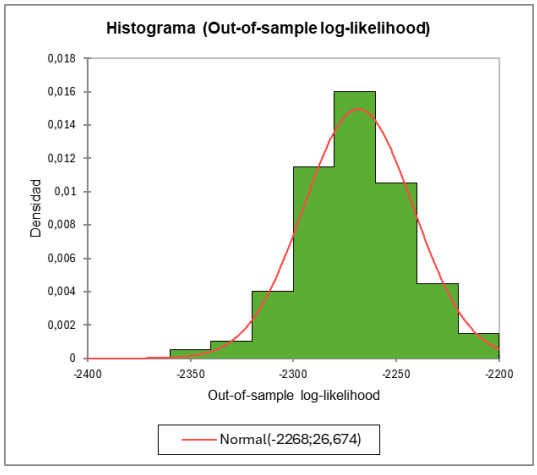


Figure S4. Constrained Industrial Network Structure (edges with > 41% confidence) with the strengths of the edges: cold colours (weak strengths) to warm colours (strong strengths).

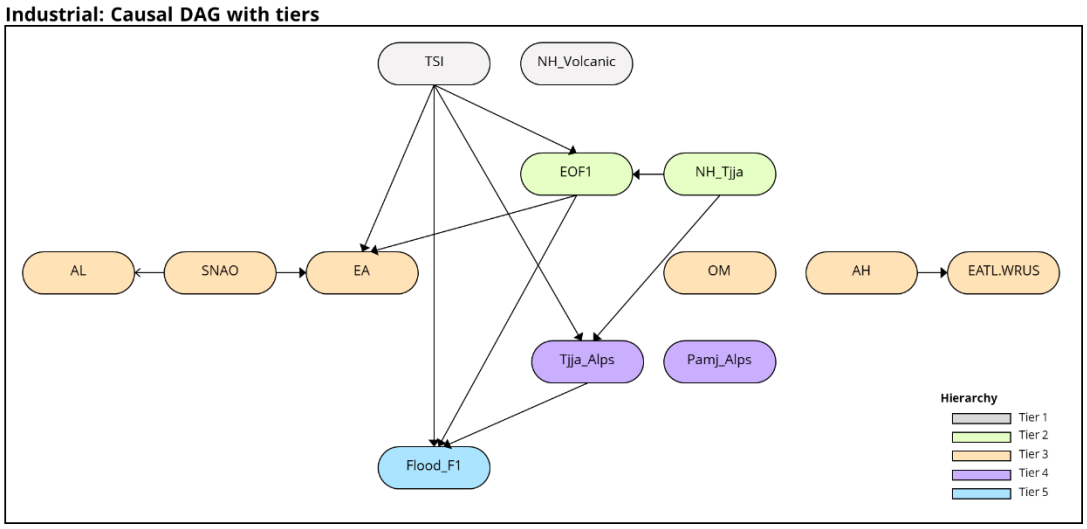


Table S6. Edge Stability Under Dating Uncertainty for Pre-Industrial vs Industrial periods

Edge from	Edge to	Pre-Industrial	Industrial
TSI	EOF1	TRUE	TRUE
TSI	Pamj_Alps	TRUE	FALSE
TSI	Flood_F1	TRUE	FALSE
TSI	Tjja_Alps	FALSE	TRUE
TSI	EA	FALSE	TRUE
EATL.WRUS	EA	TRUE	FALSE
EOF1	Tjja_Alps	TRUE	FALSE
EOF1	Pamj_Alps	TRUE	FALSE
EOF1	NH_Tja	FALSE	TRUE
EOF1	EA	FALSE	TRUE
NH_Volcanic	Tjja_Alps	TRUE	TRUE
SNAO	EA	TRUE	TRUE
NH_Tja	Tjja_Alps	TRUE	TRUE
NH_Tja	Pamj_Alps	TRUE	FALSE
NH_Tja	Flood_F1	FALSE	TRUE

S2.6 Sensitivity to blacklist constraints

To assess the influence of the expert-defined blacklist on the inferred causal structure, we compared three scenarios:

1. No blacklist → the unconstrained DAG (all edges allowed).
2. Full blacklist → the tiered hierarchy described in Section 4.5 (prohibits retrocausality, local → hemispheric edges, and links among external forcings).
3. Reduced blacklist → same as full blacklist but allowing physically unrealistic edges (e.g., temperature affecting volcanic activity).

Table S6 reports the number of edges in the resulting DAGs, and the stability of the core pathway edges (TSI→EOF1, EOF1→Pamj_Alps, Pamj_Alps→Flood_F1) under each scenario.

Table S7. Comparison of network properties under different blacklist constraints.

Blacklist scenario	Total edges	Core pathway edges (stability ≥0.41)	Edges removed relative to no blacklist	Notes
No blacklist	42	Present (stabilities >0.9)	–	Unconstrained skeleton
Full blacklist	32	All present	10 edges (23 %)	Only physically plausible edges retained
Reduced blacklist	37	All present	5 edges (12 %)	Allows unrealistic edges; core pathway unchanged

The core causal pathway (TSI→EOF1→Pamj_Alps→Flood_F1) remains intact and stable (stabilities >0.7) in all three scenarios, confirming that the main conclusions are robust to the specific blacklist design. The full blacklist removes 23 % of spurious edges without affecting the core pathway. These results are discussed in Section 5.4 of the main text.

References

Gray, L. J., Beer, J., Geller, M., Haigh, J. D., Lockwood, M., Matthes, K., Cubasch, U., Fleitmann, D., Harrison, G., Hood, L., Luterbacher, J., Meehl, G. A., Shindell, D., van Geel, B., and White, W.: Solar influences on climate, Rev. Geophys., 48, RG4001, https://doi.org/10.1029/2009RG000282, 2010.

Ineson, S., Scaife, A. A., Knight, J. R., Manners, J. C., Dunstone, N. J., Gray, L. J., and Haigh, J. D.: Solar forcing of winter climate variability in the Northern Hemisphere, Nat. Geosci., 4, 753–757, https://doi.org/10.1038/ngeo1282, 2011.

Kalisch, M., & Bühlmann, P.: Estimating high-dimensional directed acyclic graphs with the PC-algorithm. Journal of Machine Learning Research, 8, 613–636, https://doi.org/10.48550/arXiv.math/0510436, 2007.

Scutari, M.: Learning Bayesian Networks with the bnlearn R Package. Journal of Statistical Software, 35(3), 1-22, https://doi.org/10.18637/jss.v035.i03, 2010.