SUPPLEMENTARY MATERIAL TO
“CAUSAL MODELLING OF HEAVY-TAILED VARIABLES AND
CONFOUNDERS WITH APPLICATION TO RIVER FLOW”

Olivier C. Pasche
Research Center for Statistics, University of Geneva, Switzerland,
Institute of Mathematics, EPFL, 1015 Lausanne, Switzerland,
olivier.pasche@unige.ch

Valérie Chavez-Demoulin
Faculty of Business and Economics, University of Lausanne, Switzerland,
valerie.chavez@unil.ch

Anthony C. Davison
Institute of Mathematics, EPFL, 1015 Lausanne, Switzerland,
anthony.davison@epfl.ch

S.1 Variables with Comparable Tails

S.1.1 Non-Parametric Causal Tail Coefficient Estimator

Figure S.1 shows the sample distributions of the non-parametric estimators $\hat{\Gamma}_{1,2}$ and $\hat{\Gamma}_{2,1}$ for all four causal structures, for the $t_4$, Pareto(1,2) and LogN(0,1) noise distributions, respectively. The true coefficient values $\Gamma_{1,2}$ and $\Gamma_{2,1}$ are obtained using (2). Figure S.2 shows the sample distribution of the coefficient difference estimator $\hat{\Delta}_{1,2} := \hat{\Gamma}_{1,2} - \hat{\Gamma}_{2,1}$ for the $t_4$ noise distribution.

S.1.2 LGPD Causal Tail Coefficient with Post-Fit and Constrained Fit Corrections

Figure S.3 shows the sample distribution of $\hat{\Gamma}_{1,2|H}^{GPD}$ and $\hat{\Gamma}_{2,1|H}^{GPD}$ with the constrained fit, for a comparable confounder tail. Figure S.4 shows the sample distribution of $\hat{\Gamma}_{1,2|H}^{GPD}$ and $\hat{\Gamma}_{2,1|H}^{GPD}$ with post-fit correction for all four causal configurations, for the $t_4$, Pareto(1,2) and LogN(0,1) noise distributions.

S.2 Application Results for Competitors

Table S.1 shows the causal coefficients between the discharge station pairs estimated using ICA-LiNGAM, with and without considering the average catchment precipitation variable.
Figure S.1: Histograms of $\hat{\Gamma}_{1,2}$ (turquoise) and $\hat{\Gamma}_{2,1}$ (blue) for $t_4$ (top four panels), Pareto(1, 2) (middle four panels) and LogN(0, 1) (bottom four panels) distributed noise variables, for the four causal configurations. Half-lines (black) indicate $\Gamma_{1,2}$ and $\Gamma_{2,1}$. 
CAUSAL MODELLING OF HEAVY-TAILED VARIABLES AND CONFOUNDERS...

Figure S.2: Histogram of $\hat{\Delta}_{1,2}$ for $t_4$ distributed noise variables, for the four causal configurations. Lines indicate $\Delta_{1,2} = \Gamma_{1,2} - \Gamma_{2,1}$.

Figure S.3: Histograms of $\hat{\Gamma}_{1,2|H}^{\text{GPD}}$ (turquoise) and $\hat{\Gamma}_{2,1|H}^{\text{GPD}}$ (blue) with constrained fit for $t_4$ distributed noise variables, for the four causal configurations. Half-lines (black) indicate $\Gamma_{1,2}$ and $\Gamma_{2,1}$.
Figure S.4: Histograms of $\hat{\Gamma}_{1,2|H}^{\text{GPD}}$ (turquoise) and $\hat{\Gamma}_{2,1|H}^{\text{GPD}}$ (blue) with post-fit correction for $t_4$ (top four panels), Pareto(1, 2) (middle four panels) and LogN(0, 1) (bottom four panels) distributed noise variables, for the four causal configurations. Half-lines (black) indicate $\Gamma_{1,2}$ and $\Gamma_{2,1}$. 
Table S.1: Linear causal coefficients for the discharge station pairs estimated with the ICA-LiNGAM algorithm using either the station pair only (LiNGAM, two variables) or the station pair and precipitation (LiNGAM-H, three variables). Non-null values indicate significant causal effects. The arrows indicate the estimated direct causal directions between the stations.

| Stations | Pair type | LiNGAM | LiNGAM-H |
|----------|-----------|--------|----------|
| 43-62    | causal    | 1.92↑  | 2.02↑    |
| 42-63    | causal    | 2.08↑  | 2.21↑    |
| 36-63    | causal    | 3.29↑  | 3.61↑    |
| 24-61    | causal    | 2.96↑  | 3.03↑    |
| 44-61    | causal    | 2.66↑  | 2.83↑    |
| 22-38    | causal    | 2.35↑  | 2.35↑    |
| 22-35    | causal    | 2.55↑  | 2.55↑    |
| 30-45    | non-caus. | 0.84↑  | 0.87↑    |
| 36-39    | non-caus. | 0.66↑  | 0.66↑    |
| 42-34    | non-caus. | 1.39↑  | 1.29↑    |
| 42-34*   | non-caus. | 1.39↑  | 1.39*    |
| 32-33    | non-caus. | 0.59↑  | 0.54*    |
| 62-63    | non-caus. | 1.02↑  | 1.05↑    |
| 57-60    | non-caus. | 0.68↑  | 0.67↑    |
| 13-14    | non-caus. | 0.50↑  | 1.10*    |
| 17-22    | non-caus. | 1.80↑  | 1.69↑    |
| 12-21    | non-caus. | 1.04↑  | 1.08↑    |
| 26-28    | non-caus. | 0.75↑  | 0.72↑    |
| 27-31    | non-caus. | 0.54↑  | 0.66↑    |
| 23-39    | non-caus. | 0.25↑  | 0.18↑    |
| 23-35    | non-caus. | 0.42↑  | 0.36↑    |