Discrete Multivariate Generalized Pareto Distribution with application to dry spells

Samira Aka[1,2], Marie Kratz[2], Philippe Naveau[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement (LSCE)
CEA-CNRS-UVSQ

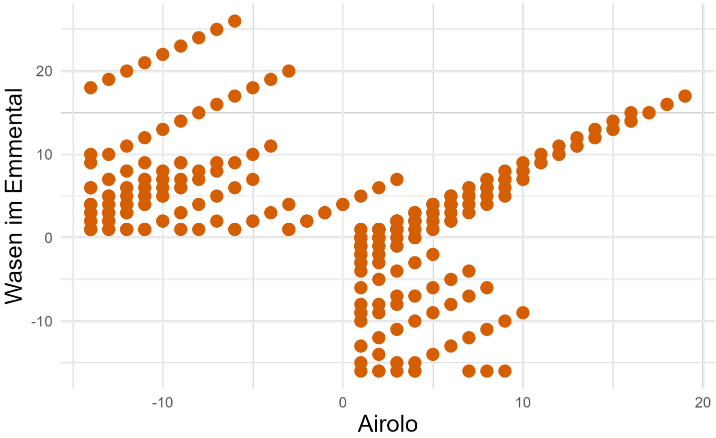
[2]Center for Research in Econo-Finance and Actuarial Sciences on Risk (CREAR)
ESSEC

AI & Environment, June 11, 2024

Motivation

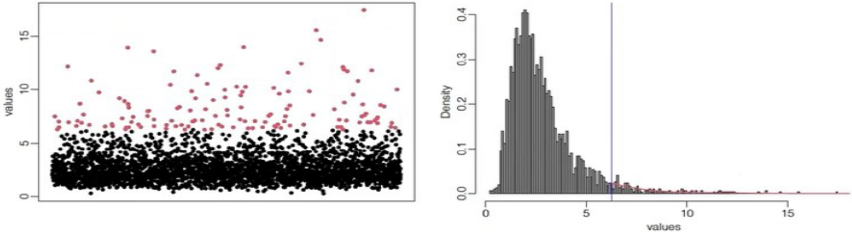
- How to model extremal dependence for multivariate discrete vectors ?
- Examples of interest in real life : insurance claims, number of fires, dry or wet spells.

Dry spells exceedances at two locations in Switzerland



Dry spells over 14 and 16 days

Exceedances in the continuous case



Generalized Pareto Distribution (GPD)

The distribution of exceedances above a large threshold u can be approximated as (Pickands-Balkema-de Haan) :

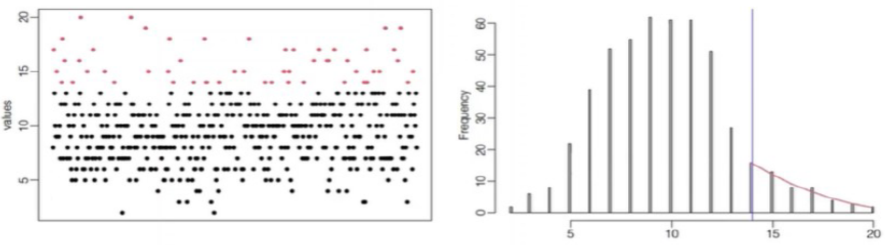
$$\mathbb{P}(Y - u > y \mid Y \geq u) \approx \overline{\text{GPD}}(y; \sigma_u, \xi)$$

where σ_u depends on u, and

$$\overline{\text{GPD}}(y; \sigma_u, \xi) = \left(1 + \xi \frac{y}{\sigma_u}\right)_+^{-\frac{1}{\xi}}, \text{ with } \sigma_u > 0,$$

is the survival function of the GPD.

Discrete GPD



Discrete generalized Pareto distribution (D-GPD) (see [Hitz et al., 2018]).

Discrete GPD definition

From [Hitz et al., 2018], the probability mass function p_{DGPD} of the discrete GPD is defined as, for $k \in \mathbb{N}$,

$$p_{\text{DGPD}}(k; \sigma, \xi) = \overline{\text{GPD}}(k; \sigma, \xi) - \overline{\text{GPD}}(k+1; \sigma, \xi).$$

Mutltivariate GPD

Definition ([Rootzén et al., 2018], Theorem 7)

$\boldsymbol{Z} \in \mathbb{R}^d$ follows a $MGPD(\mathbf{1}, \mathbf{0}, \boldsymbol{S})$ if:

- $\max(\boldsymbol{Z})$ unit exponential distribution,
- $\boldsymbol{S} = \boldsymbol{Z} - \max(\boldsymbol{Z})$ with $\boldsymbol{S}$ independent of $\max(\boldsymbol{Z})$.

LSCE

Standard MDGPD

Definition ([Aka et al., 2024])

$\boldsymbol{N} \in \mathbb{Z}^d$ follows a multivariate discrete Generalized Pareto Distribution $MDGPD(\mathbf{1}, \mathbf{0}, \boldsymbol{S})$ if :

- $\max(\boldsymbol{N})$ geometric distribution with parameter $1 - e^{-1}$

$$\mathbb{P}(\max(\boldsymbol{N}) \leq k) = 1 - e^{-k}, k \in \mathbb{N}^*$$

- $\boldsymbol{S} = \boldsymbol{N} - \max(\boldsymbol{N})$ with $\boldsymbol{S}$ independent of $\max(\boldsymbol{N})$.

LSCE

Threshold stability

Proposition ([Aka et al., 2024])

$\boldsymbol{N} \sim MDGPD(\mathbf{1}, \mathbf{0}, \boldsymbol{S})$, $\boldsymbol{A} = (a_{ij})$ *a matrix* $\in \mathbb{N}^{n \times d}$ *such that* $\mathbb{P}(\sum_{j=1}^{d} a_{ij} N_j > 0) > 0, \forall i = 1, \dots, n$, *and* $\boldsymbol{m} \in \mathbb{N}^n$, *then*

$$\mathcal{L}(\boldsymbol{AN} - \boldsymbol{m} | \boldsymbol{AN} \not\leq \boldsymbol{m}) = MDGPD(\mathbf{A1}, \mathbf{0}, \boldsymbol{S}_{\boldsymbol{m}}).$$

Generalization of the *MDGPD* by ratio on a Gamma

Proposition ([Aka et al., 2024])

$\boldsymbol{Z}$ MGPD$(\mathbf{1}, \mathbf{0}, \boldsymbol{S})$ and Λ a Gamma(α, β) random variable independent from $\boldsymbol{Z}$. Then,

$$\left\lfloor \frac{\boldsymbol{Z}}{\Lambda} \right\rfloor \text{ follows a MDGPD} \left(\frac{\boldsymbol{\beta}}{\boldsymbol{\alpha}}, \frac{\mathbf{1}}{\boldsymbol{\alpha}}, \boldsymbol{S} \right).$$

LSCE

Building **N** from the increments of a distribution

If $\boldsymbol{N} \sim MDGPD(\mathbf{1}, \mathbf{0}, \boldsymbol{S})$, then :

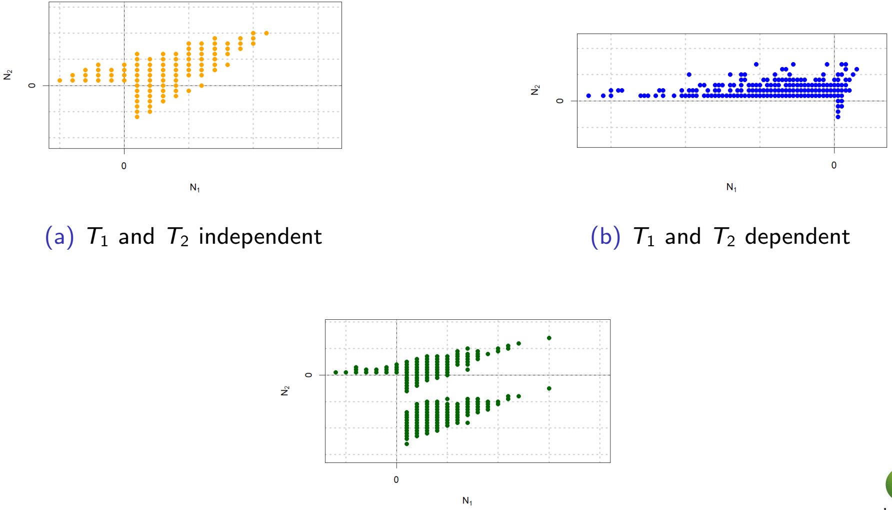
$$\boldsymbol{N} = \boldsymbol{T} - \max(\boldsymbol{T}) + \max(\boldsymbol{N})$$

with $\boldsymbol{T}$ discrete random vector.
In the bivariate case,

$$N_1 = \max(\boldsymbol{N}) + (T_1 - T_2)\mathbb{1}_{((T_1 - T_2) < 0)},$$
$$N_2 = \max(\boldsymbol{N}) - (T_1 - T_2)\mathbb{1}_{((T_1 - T_2) \geq 0)}.$$

LSCE

Simulate *MDGPD* from known T

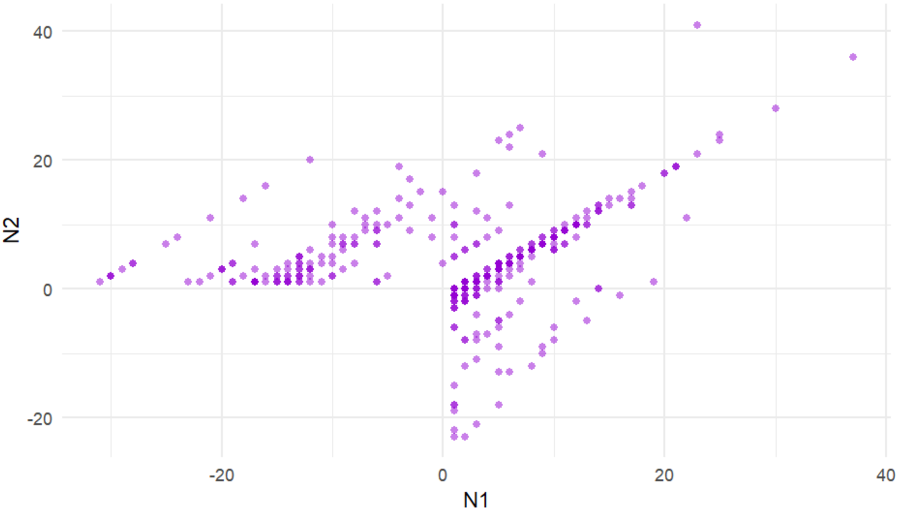


(a) T_1 and T_2 independent

(b) T_1 and T_2 dependent

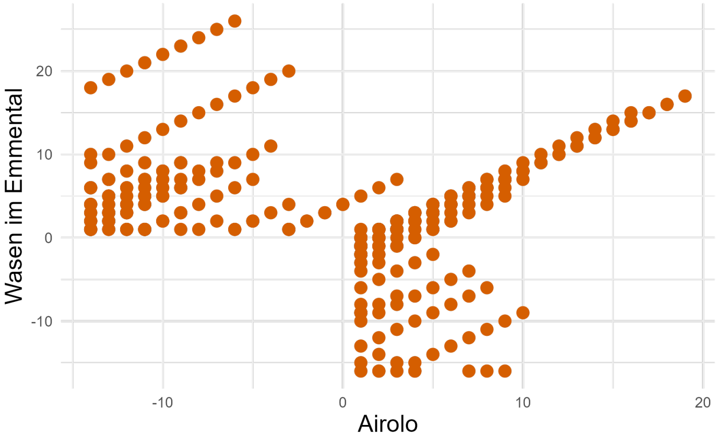
(c) T_1 and T_2 Bimodal

Simulate *MDGPD* from unknown *T*



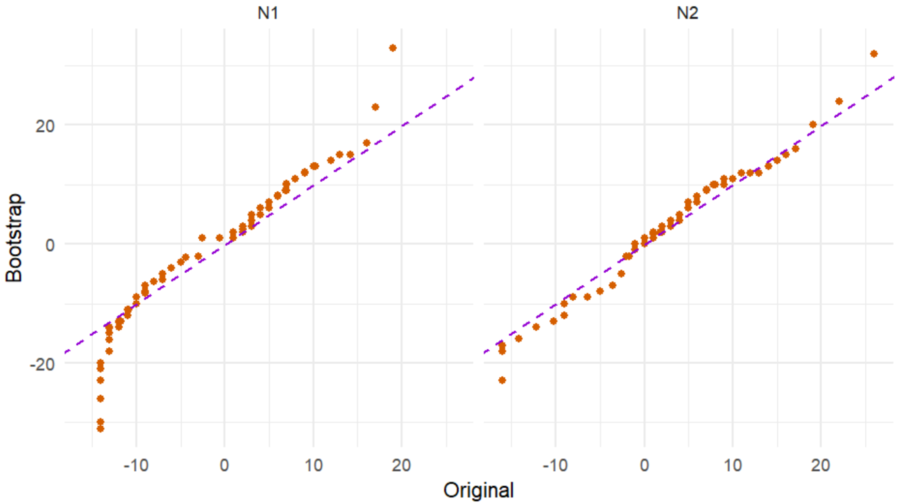
Realization of $\boldsymbol{N}$ using dry spells data

Dry spells exceedances at two locations in Switzerland



Dry spells over 14 and 16 days

Quantile-Quantile plot of marginals



Quantile-Quantile plot of marginals

Future directions

- Fit the *MDGPD* on dry spells for drought assessments,
- Perform regressions with *MDGPD*.

References

Aka, S., Kratz, M., and Naveau, P. (2024).
Discrete multivariate generalized pareto distribution with application to dry spells.

Balkema, A. A. and de Haan, L. (1974).
Residual life time at great age.
The Annals of Probability, 2.

Hitz, A. S., Davis, R. A., and Samorodnitsky, G. (2018).
Discrete extremes.

Pickands, J. (1975).
Statistical inference using extreme order statistics.
The Annals of Statistics, 3.

Rootzén, H., Segers, J., and Wadsworth, J. L. (2018).
Multivariate generalized pareto distributions: Parametrizations, representations, and properties.
Journal of Multivariate Analysis, 165:117–131.