

CONTEXTE

THÉORIE DE LA CONVERGENCE ASYMPTOTIQUE DES VALEURS EXTRÊMES

DISTRIBUTION DE PARETO GÉNÉRALISÉE (GPD)

(1) GPD CLASSIQUE

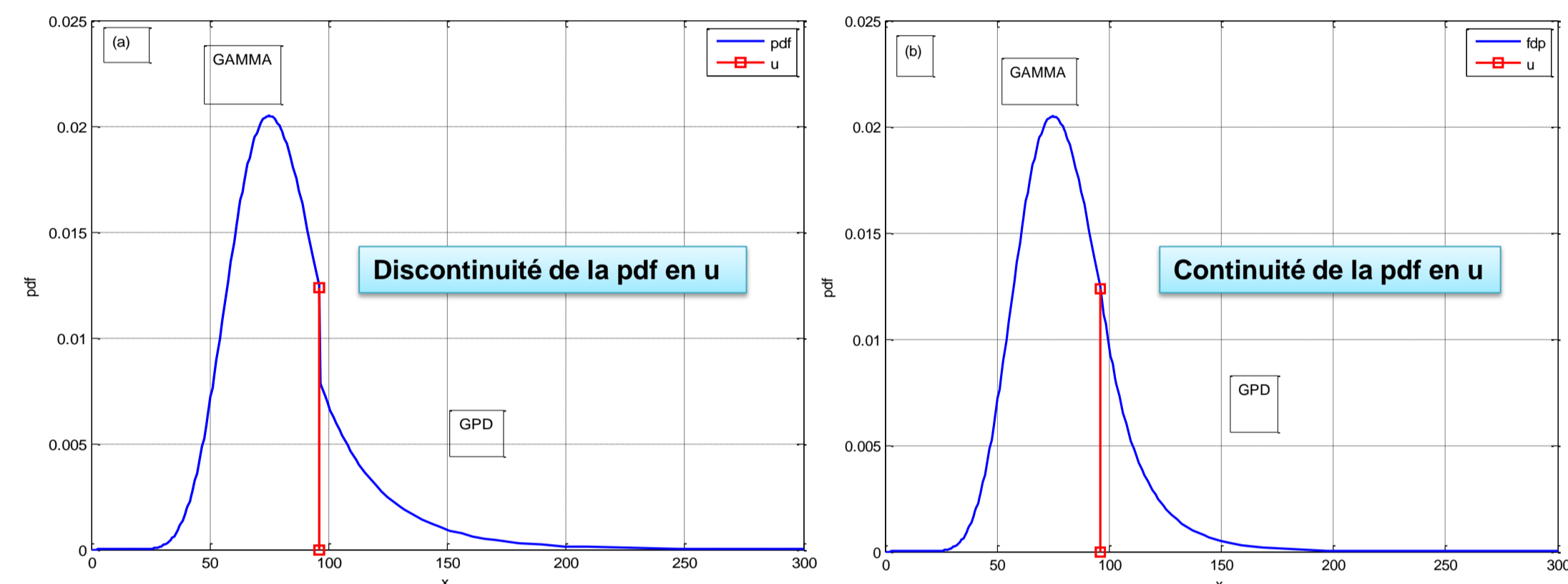
(2) GPD HYBRIDE

(1) → Fonction de densité de probabilité (fdp ou pdf) du modèle M0

$$g(x|u, \kappa, \beta) = \begin{cases} \frac{1}{\beta} \left(1 + \kappa \left(\frac{x_i - u}{\beta} \right) \right)^{-1-1/\kappa} & \kappa \neq 0; x_i \geq u \\ \frac{1}{\beta} e^{-(x_i - u)/\beta} & \kappa = 0; x_i \geq u \end{cases}$$

Calcul des quantiles: inverse de G, la fonction de répartition (cdf) de M0

(2) → Combinaison d'une distribution paramétrique (ex: Gamma, Weibull, Normale) pour C1 | $x_i < u$, avec la GPD pour C2 | $x_i \geq u$, de sorte que le seuil u est estimé au point de convergence de C1 et C2.



DÉVELOPPEMENT MÉTHODOLOGIQUE

(2.1) → Modèle Gamma-GPD sans contrainte de continuité de la fdp f en u (modèle M1)

$$f(x_i) = [h(x_i) \cdot \mathbf{1}_{\{x_i < u\}}] + [((1 - H(u)) g(x_i - u)) \cdot \mathbf{1}_{\{x_i \geq u\}}]$$

h est la pdf de Gamma sachant $\theta_1=(a, b)$.

$$h(x_i|\theta_1) = \frac{1}{b^a \Gamma(a)} x_i^{a-1} e^{-x_i/b}; x_i > 0, \quad a > 0, \quad b > 0$$

$(1 - H(u))$ garantit la condition d'utilisation de la GPD (i.e., $\Pr(X > u)$).

La fonction de vraisemblance L de M1 sachant θ_1 et $\theta_2=(\kappa, \beta)$, est :

$$L(\underline{x}|\theta_1; \theta_2) = \prod_{x_i \in C1} h(x_i|\theta_1) \cdot \prod_{x_i \in C2} (1 - H(u|\theta_1)) g(x_i - u|\theta_2)$$

(2.2) → Modèle Gamma-GPD avec contrainte de continuité de la fdp f en u (modèle M2)

La Continuité de f en u implique que

$$h(u|\theta_1) = (1 - H(u|\theta_1)) \cdot g(u|\theta_2)$$

Une solution explicite à cette équation est $\beta = \frac{(1 - H(u|\theta_1))}{h(u|\theta_1)}$

(2.1) & (2.2) → Calcul des quantiles: ajustement de la probabilité annuelle au non-dépassement $P_a | P_j = P a^{\frac{1}{365}}$ (données journalières)

→ si $P_j < P_u$, quantile=inverse de $H(P_j, \hat{a}, \hat{b})$ avec $P_u = F(\hat{u}, \hat{a}, \hat{b})$

→ sinon, quantile=inverse de $G(PP, \hat{\kappa}, \hat{\beta}, 0)$ avec $PP = (P_j - P_u)/(1 - P_u)$

INFÉRENCE BAYÉSIENNE

*Distributions a priori non-informatives pour $u, (a, b), \kappa, \beta$.

*Algorithme de Métropolis-Hastings pour obtenir

$$P(\theta|\underline{x}) = \frac{L(\underline{x}|\theta)\pi(\theta)}{\int L(\underline{x}|\theta)\pi(\theta)d(\theta)}$$

*Analyse de sensibilité: valeurs théoriques et initiales de u, a et κ .

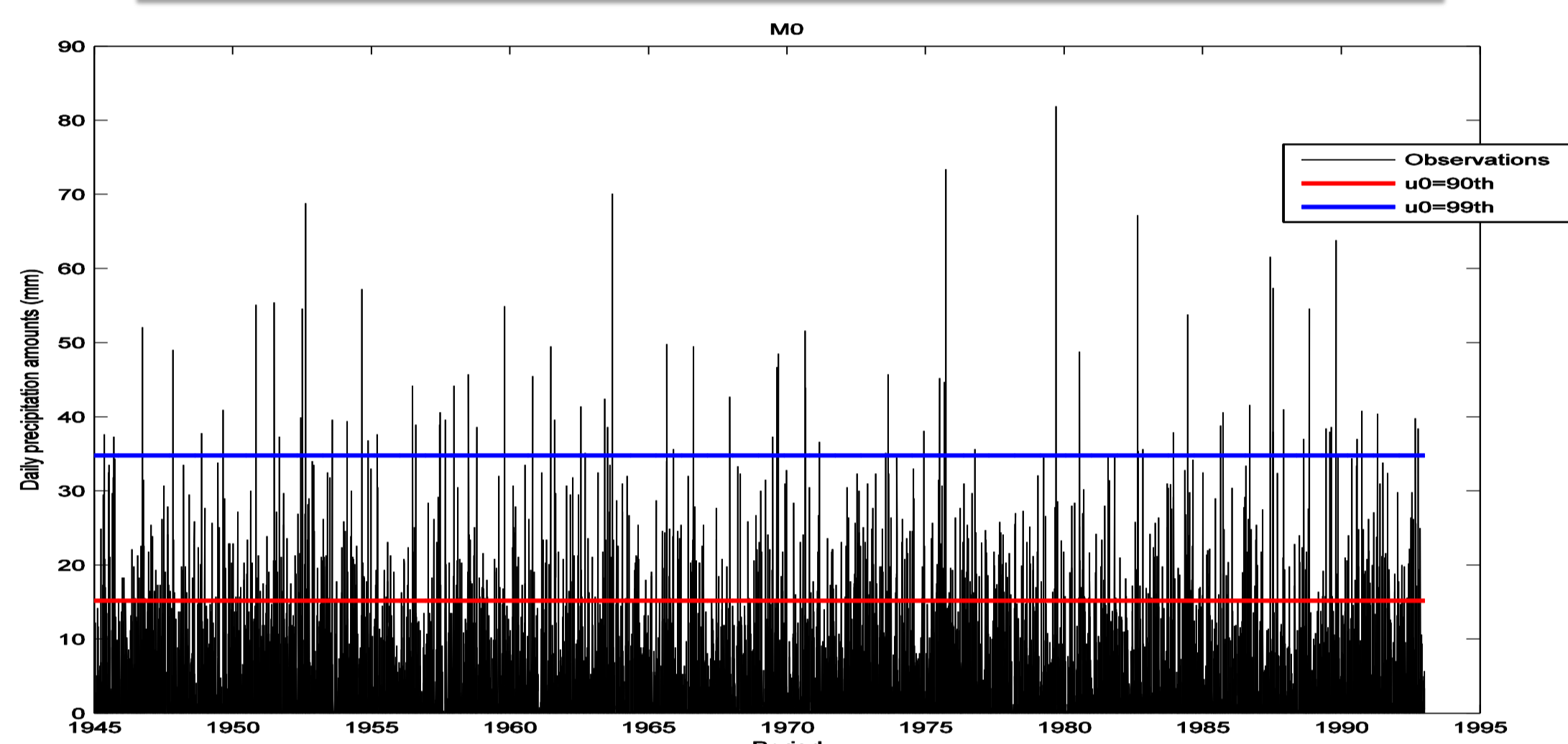
*Évaluation des quantiles de M0, M1 et M2 avec BR et REQMR.

RÉSULTATS DES SIMULATIONS

Erreurs Type I	M1 → M0	M1 → M1	M2 → M2
Scénario	p(u0)=0.8; Cs=3 (a0=0.44); b0=5; u0=3.58 k0=0.05; β0=3.42; u1=90ème percentile=8		
Estimations	u=u1 κ̂=-0.01 β̂=3.80 BR(Q10)=53% REQMR(Q10)=8%	â=0.5 b̂=8 û=3.61 κ̂=0.01 β̂=3.68 BR(Q10)=2% REQMR(Q10)=3%	â=0.5 b̂=7.33 û=2.23 κ̂=0.01 β̂=4.22 BR(Q10)=12% REQMR(Q10)=3%

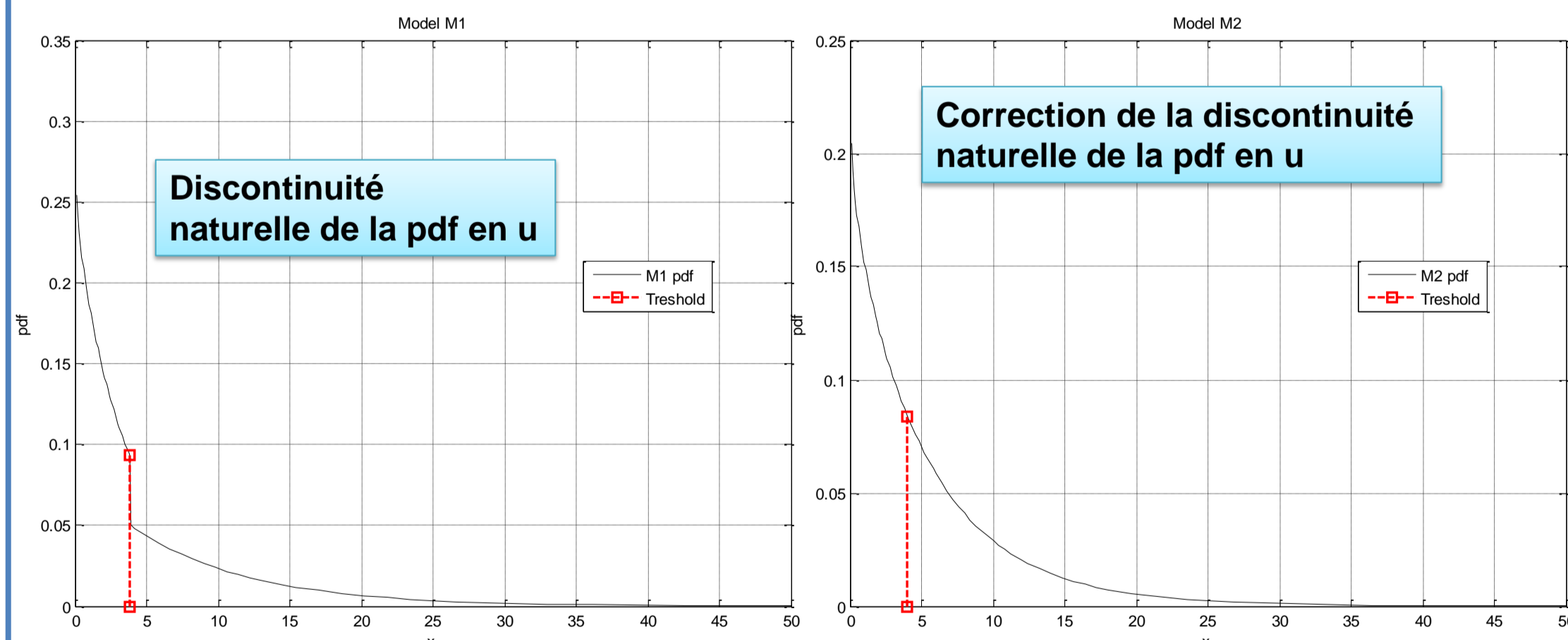
Erreurs Type II	GEV → M0	GEV → M1	GEV → M2
Scénario	ξ0=0.1; σ0=100; μ0=1000; u1=90ème percentile=1243		
Estimations	u=u1 κ̂=0.03 β̂=131.75 BR(Q10)=-12% REQMR(Q10)=2%	â=123.83 b̂=8.37 û=1017.7 κ̂=0.01 β̂=135.24 BR(Q10)=-7% REQMR(Q10)=1%	â=123.48 b̂=8.39 û=1016.3 κ̂=0.01 β̂=134.01 BR(Q10)=-7% REQMR(Q10)=1%

APPLICATION À DES DONNÉES RÉELLES



M0	M1	M2
n=687	â=0.95	â=0.94
u=u1=15mm	b̂=4.47	b̂=5.99
κ̂=-0.02	û=3.80	û=4.03
β̂=9.64	κ̂=0.01	κ̂=0.05
Q10=64 [62; 65]	β̂=7.98	β̂=5.70
	Q10=56 [55; 58]	Q10=48 [47; 49]

Station: 702525
Location: Aéroport Trudeau
Coord: (45.47; -37.75; 36)
Période: 1945-1992 (48ans)
n=7869/17532 obs. Jr.
(M1 & M2)
→ $P_{jad} = (P_j - p_0)/(1 - p_0)$
u1=90th percentile=15mm



CONCLUSION

*Simulations

***M1: meilleur modèle que M2 puis M0

***grande valeur de u avec M0 / u_0 et \hat{u} de M1 et M2

*** \hat{u} de M1 comparable à u_0

***Légère sous-estimation de \hat{u} avec M2 / u_0

***↓ Erreurs d'estimations des quantiles avec GEV → M0 / M1 → M0

***↓ Erreurs d'estimations des quantiles avec GEV → M2 / M2 → M2

*Application → Sous-est. (surest.) des quantiles avec M2 (M0) / M1

QUELQUES REFERENCES

- *Behrens CN, Lopes Hedibert F, Gerner D (2004) Bayesian analysis of extreme events with threshold estimation. *Statistical modelling* 4:227-244.
- *Carreau J, Bengio Y (2009) A hybrid Pareto model for asymmetric fat-tailed data: the univariate case. *Extremes* 12:53-76.
- *de Zea Bermudez P, Turkman MAA (2003) Bayesian approach to parameter estimation of the generalized Pareto distribution. *Test* 12: 259-277.
- *El Adlouni S, Ezzahid E, Moutassim Y (2011) Mixed distribution for loss severity modelling with zeros in the operational risk loss. *International Journal of Applied Mathematics & Statistics* 21(11). ISSN 0973-7545 (Online).
- *Liseo B (1993) Elimination of nuisance parameters with reference non-informative priors. *Biometrika* 80: 295-304.
- *Martins ES, Stedinger JR (2001) Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. *Water Resources Research* 37(10): 2551-2557.
- *Scarrott C, MacDonald A (2012) A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT-Statistical Journal* 10(1):33-60.