Supplementary material for "Generalized Pareto Regression Trees for extreme events analysis"

Sébastien Farkas[1], Antoine Heranval[1,2], Olivier Lopez[1] and Maud Thomas[1*]

[1]Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, CNRS, 4 place Jussieu, Paris, F-75005, France.
[2]Mission Risques Naturels, 1 rue Jules Lefebvre, Paris, 75009, France.

*Corresponding author(s). E-mail(s): maud.thomas@sorbonne-universite.fr;
Contributing authors: sebastien.farkas@sorbonne-universite.fr; antoine.heranval@sorbonne-universite.fr; olivier.lopez@sorbonne-universite.fr;

A. Boxplots of the quadratic errors of the simulation study

In this section, the boxplots of the quadratic errors $\int_0^1(\widehat{\gamma}(x)-\gamma_0(x))^2\mathrm{d}x$ are shown for both cases (i) step-wise function (Figure 1) and (ii) smooth function (Figure 2).

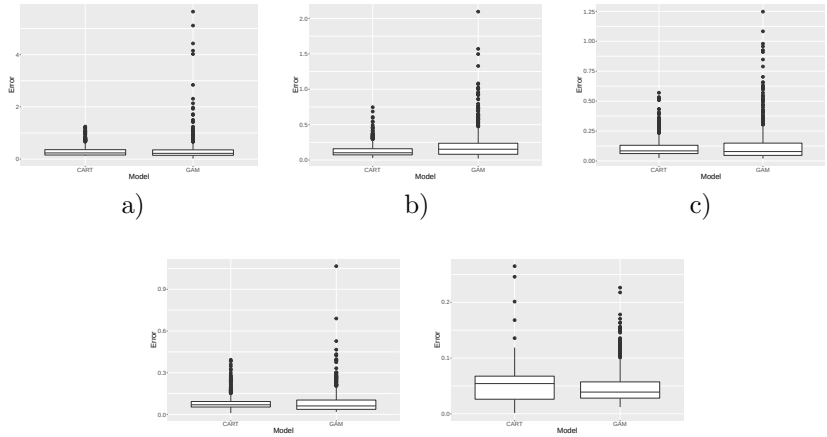


Fig. 1 Boxplots of the quadratic errors for each model in the step-wise case for a) 100 b) 250 c) 500 d) 1 000 and e) 2 500 excesses.

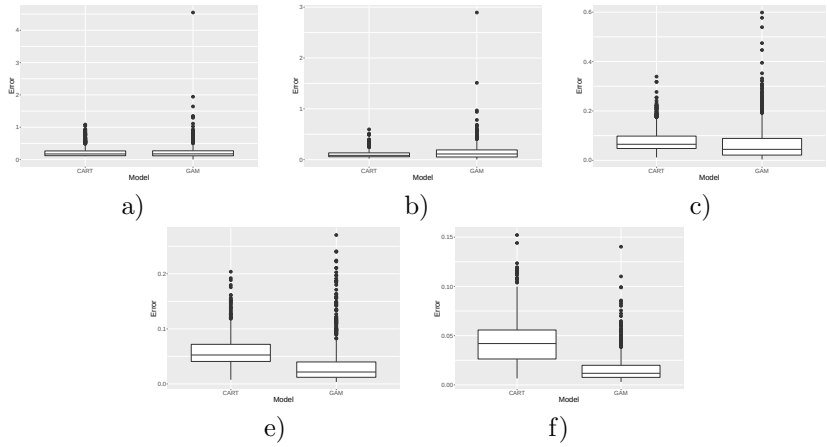
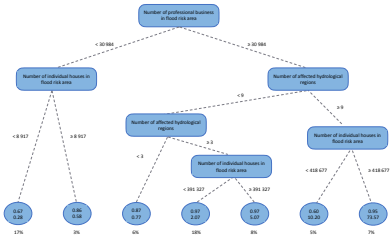


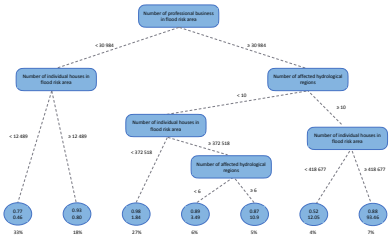
Fig. 2 Boxplots of the quadratic errors for each model in the smooth case for a) 100 b) 250 c) 500 d) 1 000 and e) 2 500 excesses.

B. Impact of threshold choice on the GP regression tree in the flooding event study

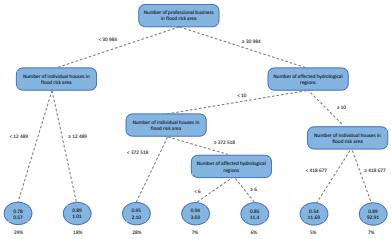
Figure 3 shows the sensitivity of our method to the choice of threshold for the GP distribution fit.



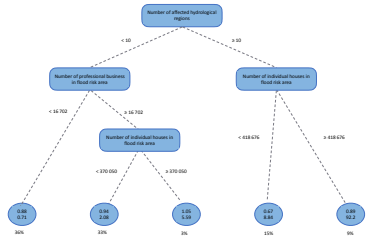
a) 2 000 largest observations
$u = 21\ 540$



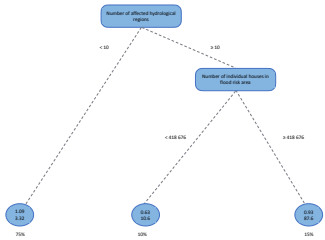
b) 1 750 largest observations
$u = 32\ 500$



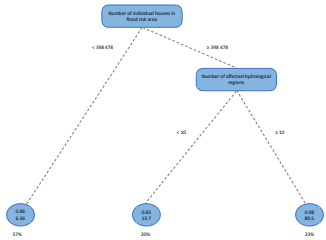
c) 1 500 largest observations
$u = 47900$



d) 1 250 largest observations
$u = 74150$



e) 750 largest observations
$u = 218000$



f) 500 largest observations
$u = 462\ 700$

Fig. 3 Trees obtained by the GP regression procedure fitted on the observations exceeding different thresholds making the number of excesses ranging from 2 000 to 500 by steps of 250. For each leaf, the estimates of γ and σ_u and the proportion of observations are given.

C. Quantile-quantile plots of the GP regression tree of the flooding events

The quantile-quantile plots for the fit of the GP distribution in each leaf of the tree (Figure 2 of the main document) are shown.

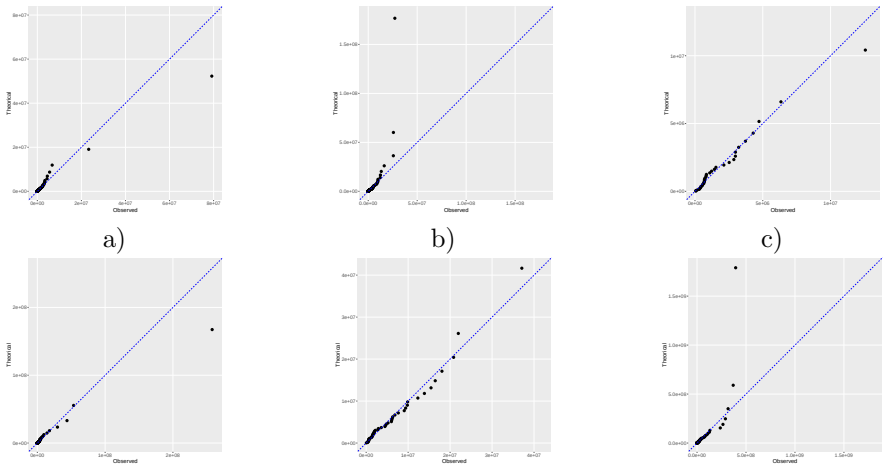


Fig. 4 Quantile-quantile plots for each leaf of the GP regression tree of the flooding events a) Leaf 1 b) Leaf 2 c) Leaf 3) d) Leaf 4 e) Leaf 5 and f) Leaf 6.