Projet ANR-08-RISK-03-01

Prédétermination des valeurs extrêmes de pluies et de crues (EXTRAFLO)

Programme RISKNAT 2008

Tâche III : Inter-comparaison des méthodes probabilistes

Rapport III.5 « *Comparaison des méthodes locales-régionales pour l'estimation des pluies extrêmes»*

Date : Juin 2010

Rapport réalisé par :

(1) HydroSciences Montpellier

Avec la participation de :

(2) Irstea, Centre d'Aix-en-Provence, OHAX
(3) Météo-France, Direction de la Climatologie
(4) EDF/DTG

Auteurs :

J. Carreau [1], L. Neppel [1], P. Arnaud [2], J.M. Soubeyroux [3], F. Garavaglia [4]

Sommaire

On rappelle les objectifs de l'action et les modèles concourants, puis le protocole de comparaison et les résultats.

1. Présentation de l'action 5

1.1. Objectifs

Elle vise à comparer les méthodes purement locales, purement régionales et locales-régionales pour la prédétermination des pluies.

Les méthodes locales-régionales possèdent les propriétés suivantes :

- Elles ont pour objectif l'estimation locale, en un site jaugé mais ne permettent aucune estimation en site non jaugé
- En plus des données locales, elles peuvent utiliser une information régionale.

On compare les performances des modèles suivants :

A. Approches locales (cf. action 1)
- Loi GP (Météo-France)
- Méthode MEWP (EDF)

B. Approches régionales (cf. action 3)
- Loi GEV_{reg} avec un index de pluie estimé à l'aide d'une régression (HSM)
- Méthode SIGEV, par interpolation spatiale des paramètres d'une loi GEV (HSM)

C. Approches locales-régionales
- Loi régionale LR : loi $GEV_{loc\text{-}reg}$ avec un index de pluie estimé à l'aide d'une série locale (HSM)
- Méthode $Shypre_{loc\text{-}reg}$, paramétrée à partir de la pluie journalière locale, mais en utilisant des relations régionales entre pluies journalières et paramètres

Mise à part la version locale régionale de la loi régionale (LR) que nous décrivons dans le paragraphe suivant, les autres modèles ont été présentés dans les actions 1 et 3.

1.2. Loi régionale (LR) en version locale régionale

La description de la LR en version purement régionale est détaillée dans le rapport de l'action 3. On rappelle ici les principes généraux et la différence avec la version locale-régionale.

On note Xi la variable aléatoire des pluies maximales annuelles au site i et $\hat{F}(x_i)$ une estimation de la fonction de répartition (fdr) de Xi. Le principe général des approches régionales consiste à estimer la fdr en un site i à partir des Nv observations situées dans un voisinage de Xi considéré comme homogène. On augmente ainsi l'échantillon permettant d'estimer $\hat{F}(x_i)$. La méthode choisie et celle des stations années. On transforme les Nv variables aléatoires initiales $(X_j)_{j=1 \text{ à } Nv}$:

$$X_j \rightarrow Y_j = \frac{X_j}{\overline{X_j}}$$

où $\overline{X_j}$ est la moyenne des pluies maximales annuelles en j, cette norme est appelée index value.

On constitue un échantillon unique des Nv observations normalisées Yj. Cet échantillon est utilisé pour l'estimation de la fdr de Yi au site i, $\hat{F}(y_i)$ appelée loi régionale. Pour estimer au site i un quantile $\widehat{x_i}(p)$ de probabilité de non-dépassement p, on estime d'abord $\widehat{y_i}(p)$ avec la loi régionale, puis la transformation inverse permet ensuite de revenir à $\widehat{x_i}(p)$, en multipliant $\widehat{y_i}(p)$ par l'index value.

La mise en œuvre de la méthode se décline en 4 étapes :

- Définition du voisinage associé au site i : Vi
- Estimation de la loi régionale : $\hat{F}(y_i)$
- Estimation de l'index value au site i : $\overline{X_i}^*$
- Estimation des intervalles de confiance des quantiles au site i

Dans la version purement régionale l'indexe value est interpolé à partir des stations jaugées, alors que dans la version locale-régionale il est calculé au site i (jaugé) où l'on veut estimer la fdr. La définition du voisinage, l'estimation de la loi régionale et des intervalles de confiance restent inchangés (Cf rapport action 3).

2. Méthode de comparaison

2.1. Les critères de comparaison

Comme pour les autres actions, la performance des méthodes est jugée sur :

- La justesse : critères NT10, NT100 et FF pour la queue de la distribution
- La robustesse : SPAN10, SPAN100 et SPAN1000
- La stabilité des intervalles de confiance : COVER10, COVER100 et COVER1000

Pour faciliter la comparaison ces critères sont normalisés entre 0 et 1, correspondant respectivement au moins bon et au meilleur score.

2.2. Les échantillons C/V

La constitution des échantillons de calage (C) et validation (V) doit permettre de juger le gain ou la perte de performance liée à l'information locale de celle liée à l'information régionale.

Pour les critères de justesse (Figure 1):

- 50% des stations sont utilisées pour le calage de l'information régionale, <u>en prenant toutes les années disponibles</u>.
- Sur le reste des stations, *n* années sont choisies aléatoirement, indépendamment pour chaque station. Elles sont utilisées pour le calage de l'information locale. On teste différentes valeurs de n : 5, 10 et 20 ans.

- La validation sera effectuée <u>sur les années complémentaires</u>, pour <u>50% des stations non utilisées</u> pour le calage de l'info régionale.

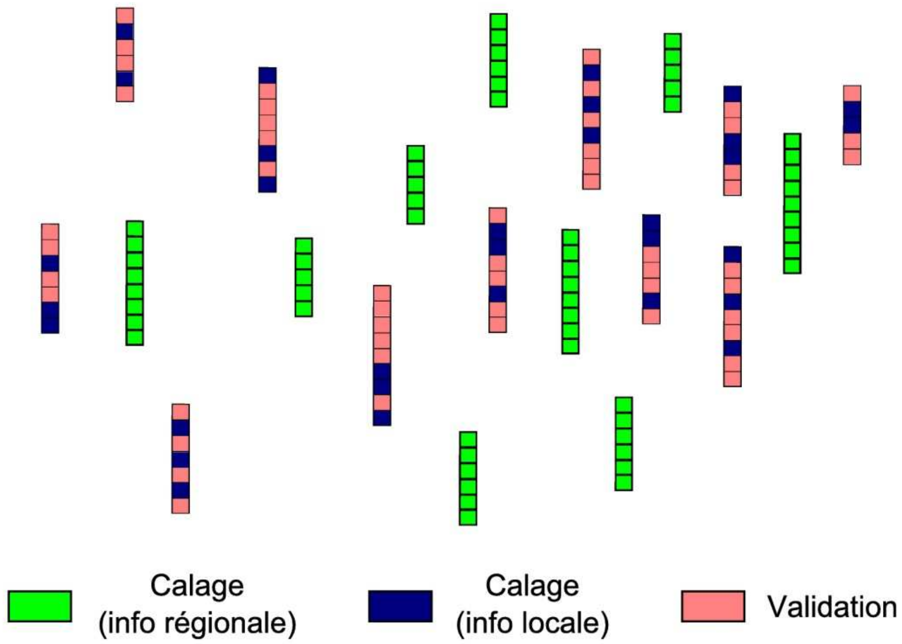


Figure 1 : décomposition C/V pour les critères de justesse

Pour la robustesse, on distingue :

- Le type 1 : robustesse par rapport à l'information locale
- Le type 2 : robustesse par rapport à l'information régionale

a) Robustesse par rapport à l'information régionale (Figure 2)

- On utilise 33% des stations pour constituer un premier jeu de calage de l'information régionale (R1) et 33% d'autres stations pour le second jeu (R2).
- Sur les 33% de stations restantes, on tire *n* années au hasard pour caler l'information locale.
- On construit deux jeux de calage C1 = R1 + L, et C2 = R2 + L. Ils diffèrent par l'information régionale, mais utilisent la même information locale.

b) Robustesse à l'information locale (Figure 3)

- L'information régionale est calée sur 33% des stations (R)
- On tire n années au hasard dans les 66% de stations restantes pour le premier jeu de calage de l'information locale (L1), et *n* autres années pour le second jeu (L2).
- On construit deux jeux de calage C1 = R + L1, et C2 = R + L2. Ils diffèrent par l'information locale mais utilisent la même information régionale.

Dans les deux cas a) et b) on teste différentes valeurs de n : 5, 10 et 20 ans.

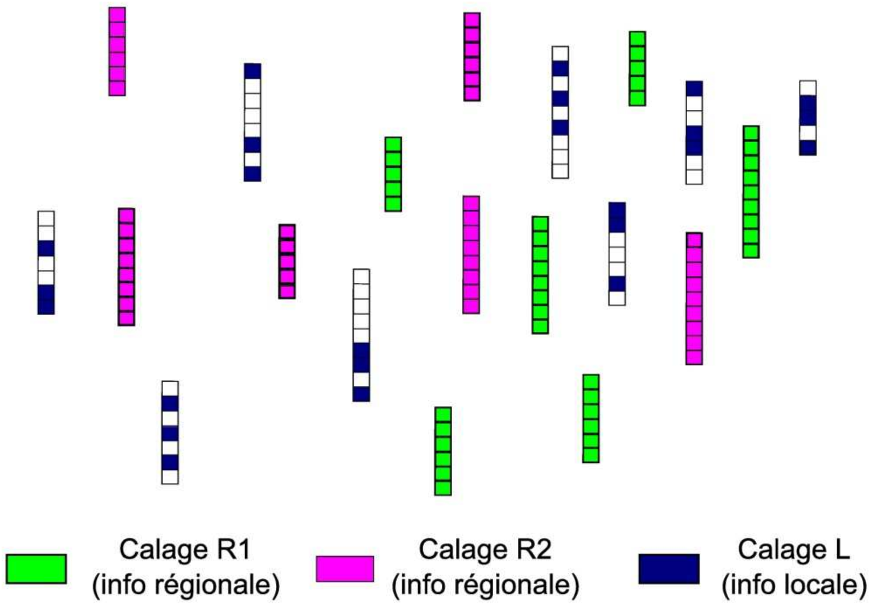


Figure 2 : Constitution des échantillons de calage pour tester la robustesse à l'information régionale

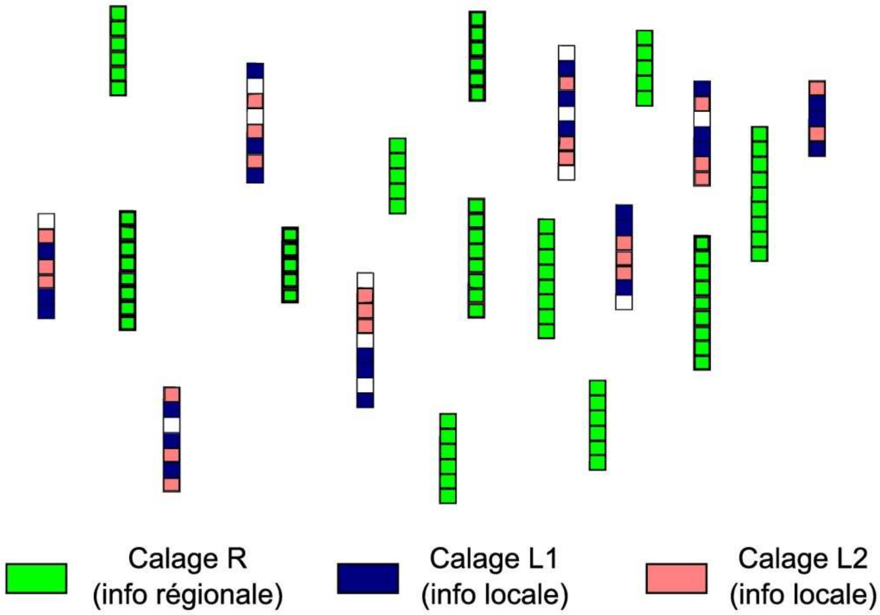


Figure 3 : Constitution des échantillons de calage pour tester la robustesse à l'information locale

3. Résultats de la comparaison

On présente dans un premier temps une comparaison de la loi régionale en version purement régionale et en version locale régionale.

Pour plus de clarté, nous comparerons ensuite d'abord la justesse de tous les modèles, puis la robustesse à l'information locale des modèles ayant une composante locale et enfin la robustesse à l'information régionale des modèles concernés.

3.1. Comparaison entre la LR locale-régionale et la LR purement régionale

3.1.1. Justesse

L'allure générale des critères NT10 montre un manque de flexibilité de la LR, prédictive ou centrale. On observe une amélioration de la justesse en augmentant le nombre d'années utilisés pour le calage dans le cas local-régional, ce qui était attendu. Pour le NT10 et le FF la justesse du modèle purement régional est du même ordre que l'approche local-régionale : on ne dégrade pas trop la justesse de la LR en mode purement régional, en tout cas par rapport à la méthode locale régionale basée sur moins 20 années.

On remarquera que les critères de justesse du NT100 sont systématiquement meilleurs que pour le NT10, ce qui peut paraître paradoxal. Comme indiqué dans le rapport II.1 *« Méthodologie de comparaison d'approches probabilistes d'estimation des valeurs extrêmes »,* le test sur le critère NT doit être utilisé à période de retour *T* fixée pour comparer plusieurs méthodes. Le fait que les courbes N_{100} soient plus proches de la bissectrice que celles de N_{10} reflète qu'il est plus difficile de détecter un manque de justesse pour la valeur centrale, car on ne dispose que peu de dépassements d'une valeur centennale dans la série des observations.

Entre prédictive et distribution centrale, les performances de justesse sont comparables.

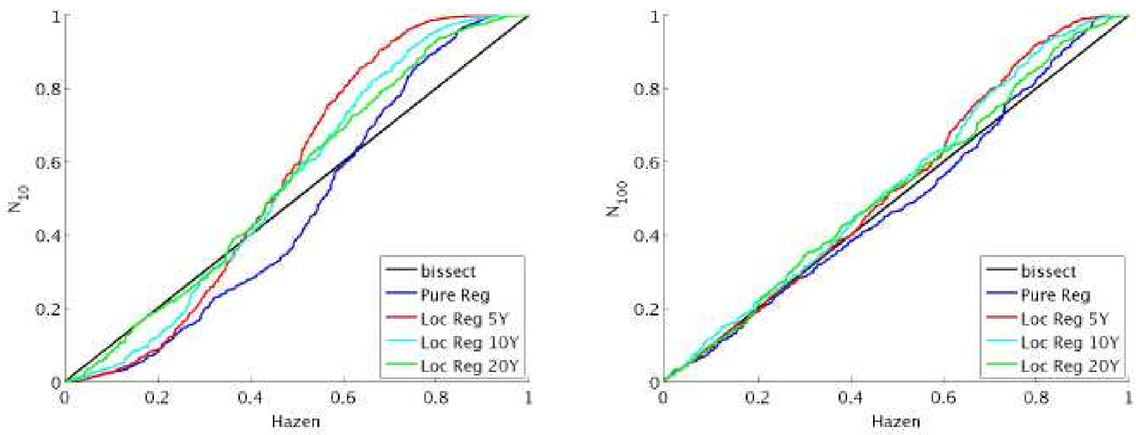


Figure 4 : NT10 et NT100 pour les LR version purement régionale et version locales régionales pour n=5, 10 et 20 – distributions centrales.

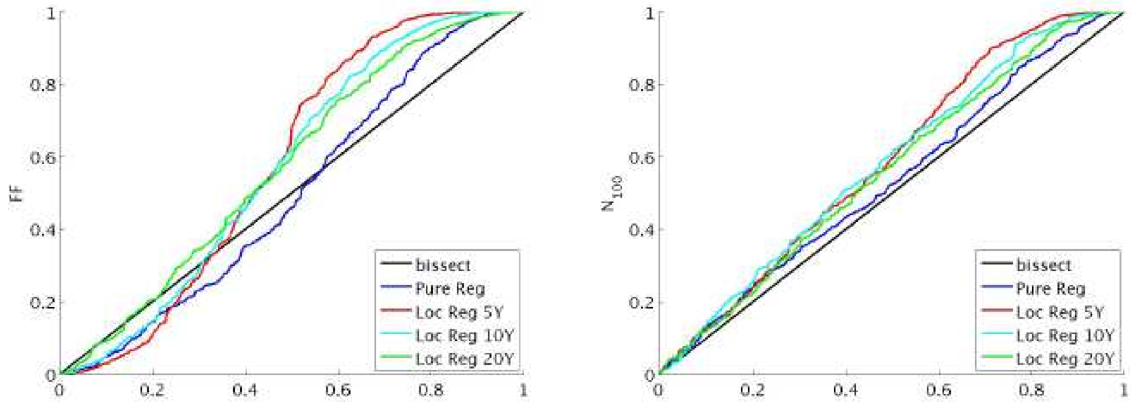


Figure 5 : NT10 et NT100 pour les LR version purement régionale et version locales régionales pour n=5, 10 et 20 – distributions prédictives.

	N_{10}	N_{100}	FF	N_{10}	N_{100}	FF
Pure-Reg	0.8384	0.9529	0.8838	0.8880	0.9310	0.9400
Loc-Reg 5Y	0.7817	0.9229	0.9200	0.7627	0.8225	0.8471
Loc-Reg 10Y	0.8461	0.9197	0.9170	0.8099	0.8349	0.8561
Loc-Reg 20Y	0.8893	0.9441	0.9609	0.8457	0.8765	0.8910

Tableau 1 : Résumé des critères de justesse, distribution centrale à gauche et prédictive à droite.

La justesse dans les extrêmes montre pour la distribution centrale une légère amélioration de la performance pour la version locale-régionale (estimation centrale) par rapport à la version purement régionale (Figure 6). Par contre le nombre d'années utilisées pour caler l'index value semble sans influence notoire. Ces résultats sont différents pour la loi prédictive où la méthode purement régionale semble meilleure.

Pour les deux versions l'asymptote horizontale du FF montre que les modèles tendent à surestimer les forts quantiles (> 0.98).

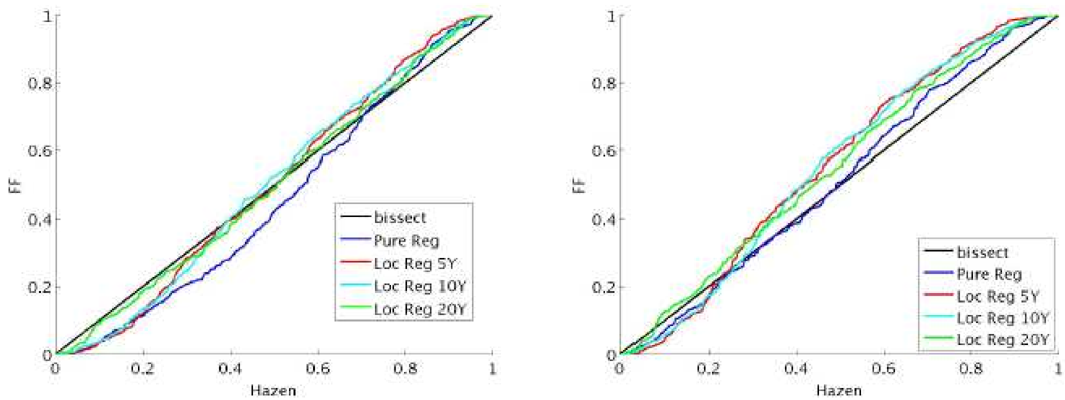


Figure 6 : Critère FF pour les LR, centrale à gauche et prédictive à droite.

3.1.2. Robustesse

a) Par rapport à l'information locale

On constate que le SPAN dépend uniquement du nombre d'années disponibles localement, mais est indépendant du niveau de retour (Figure 7). Ceci est cohérent dans la mesure où la variabilité de robustesse en fonction de la période de retour est liée à l'indice de forme de la LR, lequel ne change pas ici car il est fonction de l'information régionale. Ce critère n'est pas représenté pour la loi purement régionale car elle ne dépend pas de l'information locale.

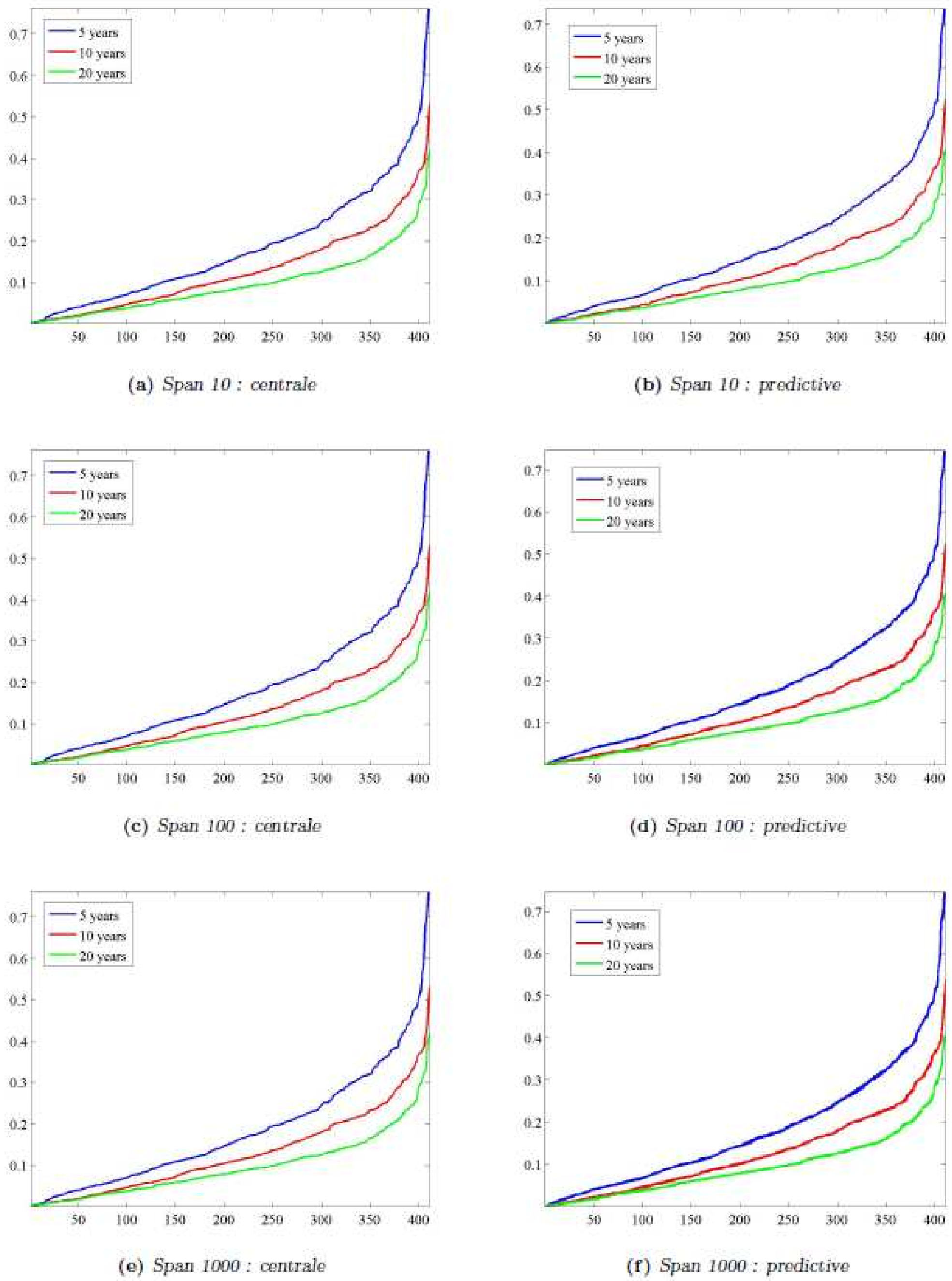


Figure 7 : robustesse de la LR locale-régionale par rapport à l'information locale, à gauche estimation centrale, à droite estimation prédictive.

b) Par rapport à l'information régionale

La robustesse à l'information régionale de dépend pas de l'information locale. La Figure 8 montre le critère SPAN pour les trois périodes de retour 10, 100 et 1000 ans et pour les versions purement régionale et locale régionale. On observe classiquement une dégradation de la robustesse quand la période de retour augmente. A fréquence fixée, la version locale régionale est toujours plus robuste que la version purement régionale. C'est cohérent car il y a deux sources de variabilité dans la version purement régionale, l'index value et la loi régionale, alors dans la version locale régionale l'index value n'est pas interpolé mais calculé localement. On remarque aussi que les distributions prédictives semblent gagner sensiblement en robustesse par rapport aux distributions centrales pour la LR local-régionale (Tableau 2). De manière générale le gain de la distribution prédictive par rapport à la distribution centrale n'est pas flagrant pour la LR.

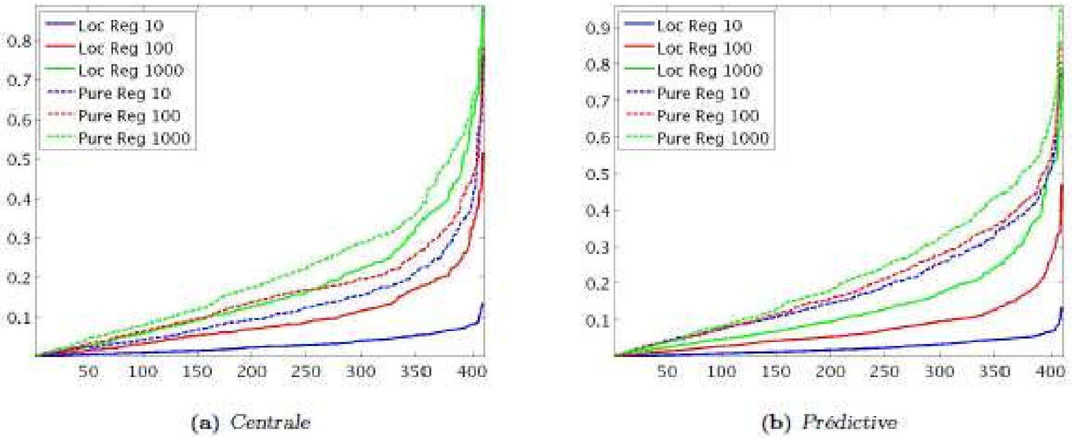


Figure 8 : Robustesse par rapport à l'information régionale.

	$SPAN_{10}$	$SPAN_{100}$	$SPAN_{1000}$	$SPAN_{10}$	$SPAN_{100}$	$SPAN_{1000}$
Pure-Reg	0.9394	0.9234	0.8942	0.9088	0.9017	0.8857
Loc-Reg	0.9861	0.9546	0.9149	0.9883	0.9631	0.9322
	-	-	-	$COVER_{10}$	$COVER_{100}$	$COVER_{1000}$
Pure-Reg				0.9732	0.9412	0.9023
Loc-Reg 5				0.9799	0.9506	0.9153
Loc-Reg 10				0.9868	0.9551	0.9155
Loc-Reg 20				0.9848	0.9490	0.9100

Tableau 2 : critère de robustesse à l'information régionale (la robustesse augmente quand on tend vers 1), à gauche distribution centrale, à droite distribution prédictive

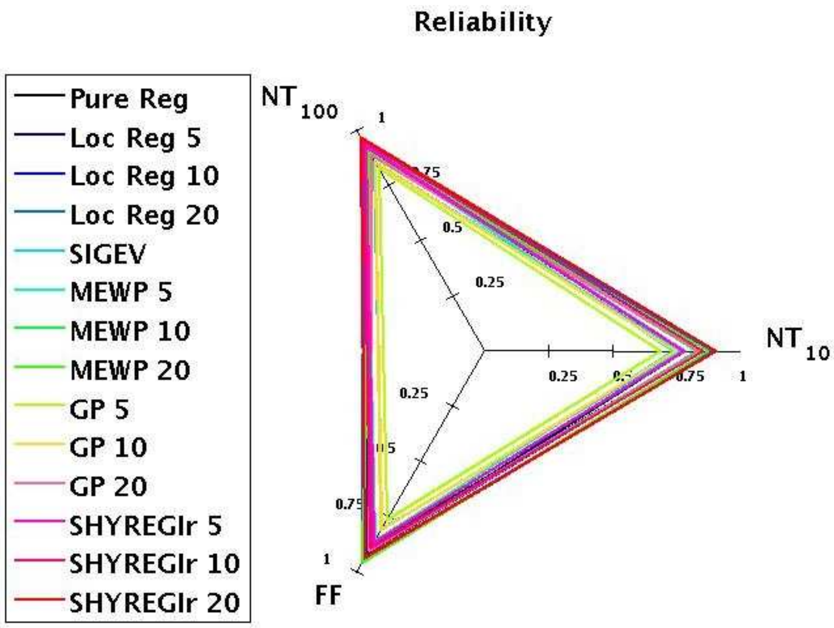
3.2.Critères de Justesse entre les différents modèles candidats

Influence du nombre d'années de calage :

Le modèle GP est le plus sensible à n, puis vient l'autre modèle purement local MWEP et les deux modèles local-régional, avec une faible influence de n sur la justesse dans les extrêmes pour la LR locale-régionale.

Globalement sur la justesse dans le corps de distribution à partir de 10 ans de données locales il est difficile de départager Shypre, MEWP et loi LR. L'approche GP est moins performante.

La justesse dans les extrêmes met aussi globalement ces trois modèles au même niveau dès que n est au moins de 10 ans pour MEWP, on note toutefois des performances un peu plus basse pour Shypre. Le modèle GP affichant les moins bonnes performances, mais encore une fois, à partir de n=20 ans, les écarts entre méthodes se réduisent.



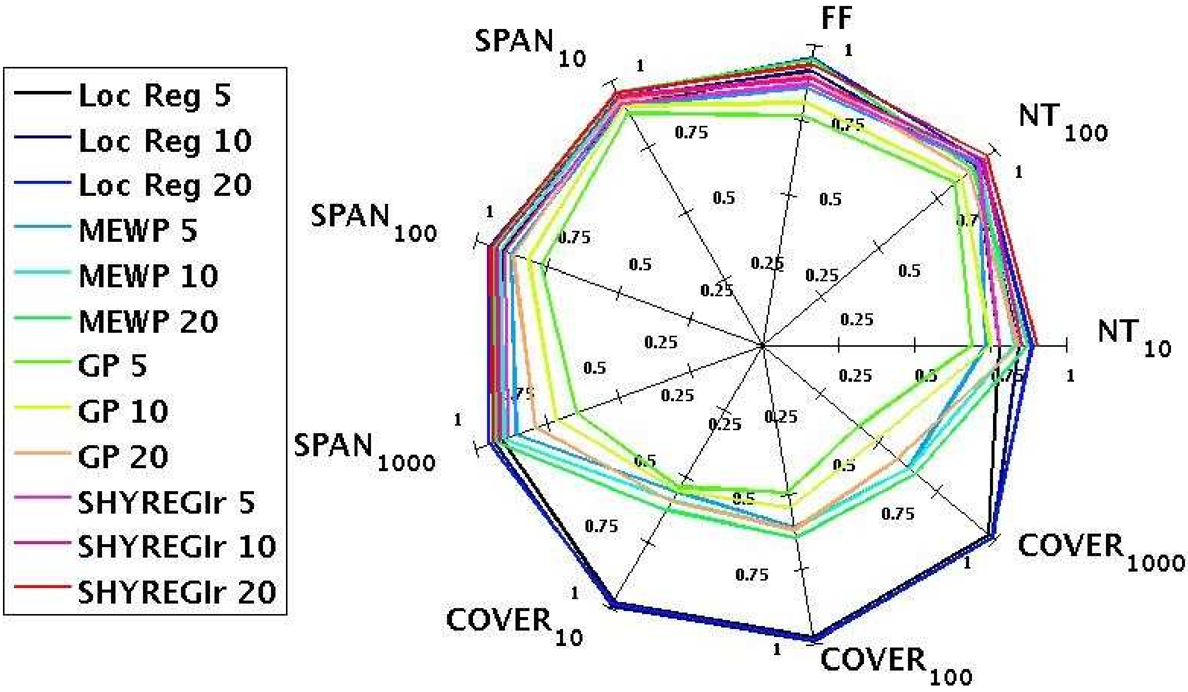
	NT_{10}	NT_{100}	FF
Pure Reg	0.8384	0.9529	0.8838
Loc Reg 5	0.7817	0.9229	0.9200
Loc Reg 10	0.8461	0.9197	0.9170
Loc Reg 20	0.8893	0.9441	0.9609
SIGEV	0.7207	0.6536	0.8077
MEWP 5	0.7380	0.9273	0.8583
MEWP 10	0.8328	0.9330	0.9014
MEWP 20	0.8774	0.9061	0.9566
GP 5	0.6915	0.8295	0.7677
GP 10	0.7530	0.8563	0.8115
GP 20	0.8431	0.8883	0.9028
SHYREGlr 5	0.7813	0.9325	0.8730
SHYREGlr 10	0.8567	0.9517	0.8928
SHYREGlr 20	0.9060	0.9688	0.9387

Figure 9 : Synthèse des critères de justesse pour les différentes méthodes.

3.3.Comparaison des modèles ayant une composante locale

On ne revient pas sur la justesse décrite précédemment. La loi LR est globalement plus robuste, pour les quantiles et surtout les intervalles de confiances, par rapport aux modèles purement locaux. Elle se démarque d'autant plus que les périodes de retour sont grandes. On remarquera toutefois que le modèle Shypre puis MEWP ont des performances de robustesse similaires à la loi LR pour l'estimation des quantiles.

La robustesse est peu dépendante du nombre d'années utilisées pour l'information locale avec la LR, Shypre et avec le modèle MEWP, en tout cas pour la décennale, les SPAN se dégradent avec ce modèle un peu plus entre n=20 et 5 ans pour des fréquences rares. La robustesse de la GP est plus sensible à n que les autres modèles. A noter que les critères de robustesse sont indépendants de la fréquence pour la LR locale-régionale, contrairement aux modèles purement locaux : on teste ici la robustesse à l'information locale (type 1), laquelle a peu d'influence sur les extrêmes.



	$SPAN_{10}$	$SPAN_{100}$	$SPAN_{1000}$	$COVER_{10}$	$COVER_{100}$	$COVER_{1000}$
Loc Reg 5	0.9097	0.9097	0.9097	0.9771	0.9767	0.9756
Loc Reg 10	0.9363	0.9363	0.9363	0.9924	0.9921	0.9919
Loc Reg 20	0.9525	0.9525	0.9525	0.9840	0.9852	0.9884
MEWP 5	0.9100	0.8804	0.8610	0.5578	0.6103	0.6320
MEWP 10	0.9367	0.9166	0.9035	0.5917	0.6162	0.6317
MEWP 20	0.9570	0.9418	0.9299	0.6276	0.6458	0.6585
GP 5	0.8820	0.7737	0.6518	0.5422	0.4936	0.4171
GP 10	0.9096	0.8191	0.7183	0.5563	0.5455	0.4886
GP 20	0.9350	0.8702	0.7947	0.5956	0.6185	0.5834
SHYREGlr 5	0.9190	0.9003	0.8887	-	-	-
SHYREGlr 10	0.9441	0.9302	0.9196	-	-	-
SHYREGlr 20	0.9595	0.9502	0.9428	-	-	-

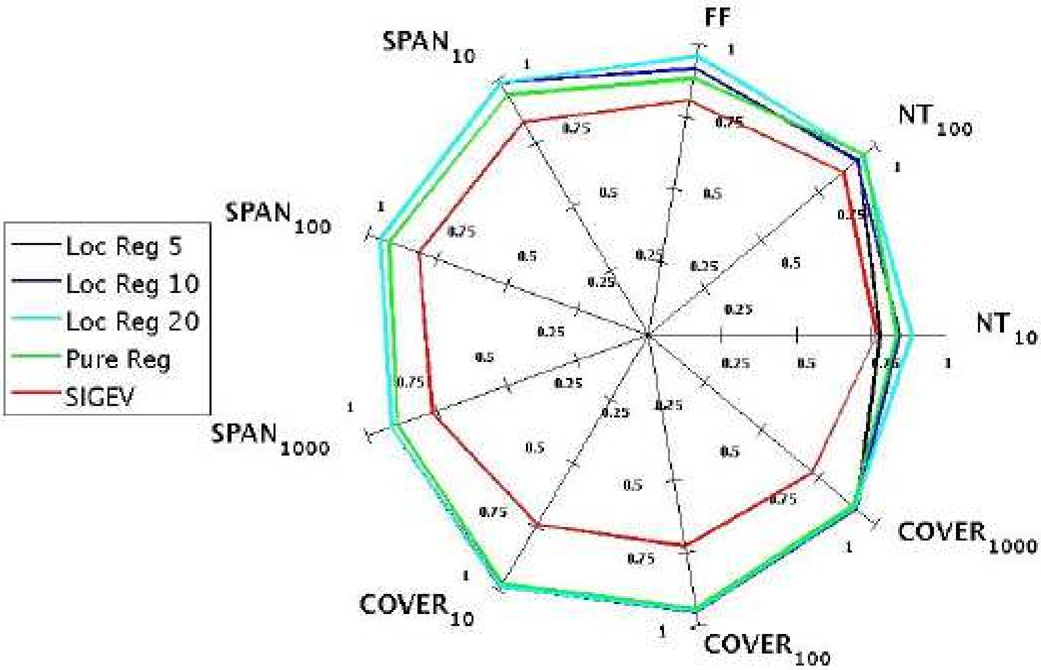
Figure 10 : Synthèse des critères pour les modèles ayant une composante locale

3.4.Comparaison des modèles ayant une composante régionale

On teste ici la robustesse de type 2 par rapport à l'information régionale.

On constate que le modèle local régional et la LR purement régionale ont des performances similaires en robustesse, quel que soit le quantile. Le modèle SIGEV l'est clairement moins. De même la stabilité des intervalles de confiances est très peu variable suivant la version locale-régionale ou purement régionale et elles se démarquent nettement du modèle SIGEV.

Comme attendu, la robustesse de la loi LR locale régionale ne dépend pas du nombre d'années de calage de l'information locale.



	$SPAN_{10}$	$SPAN_{100}$	$SPAN_{1000}$	$COVER_{10}$	$COVER_{100}$	$COVER_{1000}$
Loc Reg 5	0.9861	0.9546	0.9149	0.9799	0.9506	0.9153
Loc Reg 10	0.9861	0.9546	0.9149	0.9868	0.9551	0.9155
Loc Reg 20	0.9861	0.9546	0.9149	0.9848	0.9490	0.9100
Pure Reg	0.9394	0.9234	0.8942	0.9732	0.9412	0.9023
SIGEV	0.8348	0.8162	0.7697	0.7413	0.7273	0.7247

Figure 11 : Synthèse des critères de robustesse par rapport à l'information régionale (type II)

4. Conclusions

Les principales conclusions :

- Les modèles avec une composante régionale sont peu influencés par le nombre d'années de calage par rapport aux modèles purement locaux, excepté MWEP qui parmi les modèles locaux semble le moins sensible à l'effectif local.
- Il est difficile de départager les modèles par rapport à la justesse dans le corps de distribution dès que n>=10 ans, sauf l'approche GPD qui semble un peu moins performante.
- Pour la justesse des extrêmes la LR présente les meilleures performances, suivi de près par MEWP et Shypre local régional.
- La hiérarchisation des modèles par rapport à la robustesse des quantiles est la même que précédemment, mais par rapport à la robustesse des intervalles de confiance la LR est plus performante.

Si l'on s'en tient à l'objectif de cette action à savoir comparer les modèles type « local-régional », seul $\text{Shyreg}_{\text{loc-reg}}$ et la loi LR rentrent dans cette catégorie. Au vu des écarts faibles entre critère il nous paraît difficile de les départager. La LR présente deux avantages : les intervalles de confiances des estimations non fournis par Shypre et une mise en œuvre plus simple de la méthode. La méthode Shypre a l'avantage de ne pas se limiter à la pluie journalière et de fournir des distributions des pluies multidurées de 1h à 72h

La comparaison avec les méthodes purement locales montre que MEWP a des performances de robustesse et de justesse des extrêmes comparables aux méthodes avec une composante régionale, en particulier quand on dispose de plus de 10 ans de données en calage. L'approche GP par contre a des moins bonnes performances d'autant plus que le nombre d'années disponibles localement est faible.

La comparaison entre modèle local régional et purement régional ne montre qu'une dégradation légère des performances en n'exploitant aucune information locale (site non jaugé).