

Un modèle de précipitations printanières en contexte d'inondation

Fateh Chebana * Hélène Cossette † Alexandre Lepage ‡‡ Étienne Marceau †

27 mars 2021

Résumé

Dans cet ouvrage, nous partons de [Jalbert et al., 2019] pour proposer une modèle alternatif de prédiction des pluies printanières totales dans un contexte de modélisation des inondations saisonnières. Notre approche bonifie les modèles DDF (*Dept, Duration, Frequency*) communément utilisés en pratique en ajoutant les notions de processus d'arrivée non stationnaire tel que présenté dans [Nelson and Gerhardt, 2011]. De plus, nous faisons une analyse exhaustive de la dépendance entre les différentes variables en jeux à l'aide de la théorie des copules telle que décrite dans [Joe, 1997] et [Nelsen, 2006].

Keywords— Processus d'arrivée, processus non stationnaire, dépendance, copules, valeurs extrêmes, précipitations, distribution Pareto généralisée, excès de seuil.

1 Introduction

Autant en assurance qu'en hydrologie, le sujet des inondations suscite l'intérêt des chercheurs qui tentent de modéliser ces événements afin de mieux se préparer à d'éventuelles catastrophes. Entre autres, le débordement du lac Champlain en 2011 a suscité l'intérêt de [Riboult and Brisette, 2016] qui a chercher à connaître les éléments déclencheurs d'une telle catastrophe. Lors de son étude, il arriva à la conclusion que, bien que la fonte des neiges soit une variable explicative importante, c'est l'accumulation de précipitations extrêmes dans la période de mars à juin qui a été la cause principale dans ce cas.

Parler de [Kao and Govindaraju, 2007]. Lire aussi [Vandenbergh et al., 2010]

Suite à la publication de [Riboult and Brisette, 2016], [Jalbert et al., 2019] a cherché à prédire l'accumulation des pluies printanières du lac Champlain afin d'estimer l'amplitude maximale qu'un tel événement aurait pu avoir et de calculer l'espérance du temps qui s'écoulera avant qu'un incident d'une telle ampleur survienne à nouveau.

Le modèle ainsi conçu sépare la modélisation des pluies en deux composantes. Une portion régulière qui représente la quantité de pluie totale tombée selon les normes saisonnières et une portion extrême où on considère les jours où la quantité de pluie tombée dépasse un certain seuil. Tandis que la première est bien modélisée avec une loi normale, la seconde nécessite plus de travail. En effet, afin de modéliser la quantité quotidienne de pluie

*Institut national de recherche scientifique, Québec, Qc, Canada

†École d'actuariat, université Laval, Québec, Qc, Canada

‡Tel.: 581-984-3704, E-mail: alexandre.lepage.3@ulaval.ca

tombée dans les cas extrêmes, [Jalbert et al., 2019] utilise la loi Pareto généralisée, aussi connue sous l'appellation *Peaks-Over-Threshold*(POT). Cependant, l'estimateur du paramètre de forme de la loi POT calculé sur les valeurs quotidiennes faisaient en sorte que la distribution avait un support fini et la probabilité qu'un événement de l'envergure de 2011 se produise était quasiment nulle. Conséquemment, [Jalbert et al., 2019] proposa une extension du modèle POT où il divise la quantité de pluie tombée lors d'une journée de pluie extrême par la proportion que cette quantité représente sur l'ensemble des précipitations tombées au cours d'une période de pluie continue. Se faisant, il n'y a plus de problème de borne supérieure à la distribution de la sévérité.

Le modèle que nous proposons est une alternative à celui de [Jalbert et al., 2019]. Il reprend le concept de regroupement des jours en périodes de pluie continue, mais il adopte une approche légèrement plus intuitive en s'inspirant grandement de la méthodologie proposée dans [Zhang and Singh, 2019] et [Shaw et al., 2010]. Comme il est expliqué, entre autre, au chapitre 9.6 de [Shaw et al., 2010], les modèles de précipitations utilisés dans la pratique, communément appelés modèles DDF (Dept, Duration, Frequency), sous-tendent les variables de sévérité de durée et de fréquence, lesquelles ont toutes leur importance. Or, [Jalbert et al., 2019] se concentre principalement sur la sévérité. La fréquence est modélisée avec une loi de Poisson conditionnelle-gamma, mais l'aspect de la durée est négligé. Notre approche, quant à elle, s'intéresse à toutes ces variables aléatoires et aux relations qui les unissent conformément à [Zhang and Singh, 2019]. Nous incorporons des notions des processus de renouvellement aux concepts communément utilisés en hydrologie afin d'améliorer la précision du modèle de fréquence et les liens de dépendance sont modélisés à l'aide de la théorie des copules. Également, une attention particulière est apportée à la non-stationnarité des distributions de probabilité afin de prendre en compte le contexte des changements climatiques.

L'article est découpé comme suit : la section 2 présente le modèle proposé...

2 Le modèle proposé

L'intuition derrière le modèle proposé est de considérer les périodes de précipitations continues comme des événements uniques en agrégant la quantité de pluie tombée lors de celles-ci. Se faisant, non seulement on considère les jours où le volume d'eau tombé sort de l'ordinaire, mais on considère également les séquences où le nombre de jours de pluie continue est anormalement élevé. De plus, on considère que le nombre d'événements extrêmes se produisant dans une année donnée est modélisé par un processus de renouvellement. Également, on s'intéressera à la présence de dépendance unissant les v.a. en jeu.

Comme les données étudiées dans la section 4 ne présente aucun signe d'auto-corrélation, on pose ici l'hypothèse d'indépendance séquentielle sur les distributions étudiées. En revanche, aucune hypothèse n'est faite sur la stationnarité des distributions puisque de fortes évidences pointent vers la non-stationnarité de celles-ci.

2.1 Description du modèle

On définit la suite de v.a. $\underline{Y} = \{Y_l\}_{l=1}^{91}$ comme la quantité en mm de pluie tombée lors des 91 jours du printemps, c.-à-d. du 1^{er} avril au 30 juin, pour une année donnée. Selon [Riboust and Brissette, 2016] et [Jalbert et al., 2019], cette période de temps est la plus critique lors des crues printanières.

Soit $\mathcal{C} = \{\mathcal{C}_j, j \in \mathbb{N}^+\}$, les périodes de pluies continues de cette même année. En concordance avec [Jalbert et al., 2019], on définit une période de précipitations continues comme une séquence de jours de pluie séparées par au moins une journée d'ensoleillement (0 mm de pluie). On définit $\underline{K} = \{K_j, j \in \mathbb{N}^+\}$ comme la suite de v.a. de la quantité de pluie totale tombée lors de la période \mathcal{C}_j telle que $K_j = \sum_{l \in \mathcal{C}_j} Y_l$.

Pour un seuil critique u , on sépare les événements en deux sous-ensembles $\mathcal{Z} = \{j : K_j \leq u\}$ et $\mathcal{X} = \{j : K_j > u\}$. En ce qui a trait aux événements extrêmes, afin de faciliter les prochaines étapes, posons $\underline{X} = \{X_i, i \in \mathbb{N}^+\} = \{K_j : j \in \mathcal{X}\}$.

Soient $\underline{T} = \{T_i, i \in \mathbb{N}\}$ la suite de v.a. représentant le temps d'arrivée des événements $\{\mathcal{C}_j, j \in \mathcal{X}\}$ et $\underline{D} = \{D_i, i \in \mathbb{N}\}$, la suite de v.a. représentant leur durée. On pose également la suite de v.a. $\underline{W} = \{W_i, i \in \mathbb{N}\}$, représentant les temps séparant les événements extrêmes telles que $W_i := T_i - T_{i-1} - D_{i-1} = T_i - T_{i-1}^*$, où $T_i^* := T_0 + \sum_{k=0}^i W_k + D_k$ et T_0 correspond au 1^{er} avril de l'année à l'étude, traduite en format numérique selon une date de référence choisie arbitrairement. Par convention, on a $W_0 = D_0 = 0$.

Soit N , le processus de renouvellement représentant le nombre de périodes de pluies continues où la quantité cumulée d'eau par événement dépasse un seuil critique u . Par définition des processus de renouvellement, pour un accroissement correspondant à une saison printanière $([T_0, T_0 + 91])$, on a

$$N := \sum_{i=1}^{\infty} \mathbf{1}\{T_i^* \leq T_0 + 91\} = \sum_{i=1}^{\infty} \mathbf{1}\{T_i + D_i \leq T_0 + 91\}. \quad (1)$$

Afin d'étudier la quantité de pluie totale tombée lors d'une saison, on définit la v.a. des précipitations non-extrêmes totales comme $Z := \sum_{j \in \mathcal{Z}} K_j$, de même que la v.a. des précipitations extrêmes totales

$$V := \sum_{i=1}^{\infty} X_i \mathbf{1}\{T_i + D_i \leq T_0 + 91\}. \quad (2)$$

On peut donc modéliser la quantité de pluie totale tombée pendant une saison printanière avec $S = Z + V$. La dépendance unissant les variables en jeu est modélisée avec la théorie des copules décrite dans [Joe, 1997] et [Nelsen, 2006].

2.2 Choix des distributions marginales

Avec la théorie des valeurs extrêmes, comme on cherche à modéliser l'ensemble des précipitations qui sortent des normales saisonnières, la méthode POT est tout indiquée. Alors on pose $(X_i - u | X_i > u) \sim \text{GPD}(\xi, \sigma)$ telle que la fonction de répartition s'exprime comme

$$F_u(x) := \begin{cases} 1 - (1 + \xi x / \sigma)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-x / \sigma), & \xi = 0, \end{cases} \quad (3)$$

où $x \geq 0$, $\sigma > 0$, $\xi \in \mathbb{R}$ et $1 + \xi x / \sigma > 0$, pour un seuil $u > 0$. À noter que la paramétrisation de la loi est grandement influencée par la valeur attribuée au paramètre u , lequel peut être obtenu par optimisation tel que présenté dans [Bader et al., 2016] et [Bader et al., 2018] ou encore avec [Northrop et al., 2015]. Bien que certains auteurs ([Citer certains auteurs](#)) utilisent un seuil non stationnaire, nous avons décidé d'utiliser un seuil fixe. Les raisons étant que ce n'est pas parce qu'il existe une tendance dans la quantité d'eau qui tombe dans une année qu'un événement est nécessairement plus extrême qu'avant (boîteux...), ainsi que par souci de simplicité ([Appuyer avec la littérature](#))...à revoir...

En ce qui a trait à la modélisation de la durée des périodes de pluie, l'approche proposée s'appuie sur... [Trouver des références...](#)

Puis, pour la v.a. des temps inter-occurrences, les distributions considérées sont l'exponentielle, la Weibull et la gamma puisque... [Trouver des références...](#)

3 La méthodologie

Sujets :

1. Clustering
2. Vérification de l'indépendance séquentielle
3. Sélection de seuil automatique
4. Vérification de la stationnarité
5. Distribution de l'excédent de seuil
6. Distribution des normales de saison
7. Distribution des temps inter-occurrences
8. Sélection de la copule
9. Études de cas

Relativement au traitement des distributions non stationnaires, les approches proposées dans ([À lire : \[Nelson and Gerhardt, 2011\]](#)), [\[Khaliq et al., 2006\]](#) et dans [\[Chebana and Ouarda, 2021\]](#) sont considérées.

Du point de vue des distributions marginales, certaines v.a. présentent de fortes évidences de non-stationnarité. Afin d'en tenir compte dans la paramétrisation des lois de probabilité, on détermine que la moyenne est une fonction linéaire du temps d'arrivée $\{T_i\}_{i \geq 0}$. Puis le paramètre d'échelle de la loi devient une fonction de cette moyenne. Pour ce qui est du paramètre de forme de ces lois, celui-ci est présumé comme étant stationnaire.

Du point de vue de la stationnarité de la distribution des copules, l'approche proposée par [\[Chebana and Ouarda, 2021\]](#) a été considérée. Cependant, étant donnée la faible quantité d'observations extrêmes dans le jeu de données, la méthode est très instable. Conséquemment, par souci de simplicité, la corrélation entre les variables aléatoires est présumée stable dans le temps.

Afin de vérifier l'hypothèse d'indépendance séquentielle, l'approche proposée par [\[Genest and Rémillard, 2004\]](#) est utilisée. Cependant, il est important de mentionner que cette approche suppose que la séquence d'observations sur laquelle le test est effectué est continue. Or, dans le cas présent, on ne considère que les saisons printanières de chacune des années. Conséquemment, on se retrouve avec des intervalles de temps disjoints. De plus, comme le nombre d'observations par saison est généralement faible, le test statistique peut être non fiable. Donc, pour tester l'hypothèse d'indépendance séquentielle, l'approche employée fût d'effectuer le test de [\[Genest and Rémillard, 2004\]](#) pour chacune des années indépendamment, puis d'observer les quantiles empiriques obtenus sur les *p-values*. À noter que la méthode utilisée n'est pas appuyé par la littérature, mais que la création d'un test formel pour ce genre de situation pourrait devenir un sujet de recherche intéressant que nous laissons à d'autres.

Pour ce qui est de la modélisation de la dépendance, celle-ci est effectuée selon la théorie des copules décrite dans [\[Joe, 1997\]](#) et dans [\[Nelsen, 2006\]](#). Dans le contexte particulier des précipitations, on s'intéressera particulièrement à [\[Zhang and Singh, 2019\]](#) [élaborer davantage...est-ce que je considère les copules en vigne ou je continue avec des copules bivariées ? Lire aussi \[Salvadori and De Michele, 2007\]](#). Également, dans le contexte des valeurs extrêmes, on considère [\[Gudendorf and Segers, 2010\]](#) qui recommande, entre-autre, les copules de Gumbel, de Tawn et de

Galambos. En théorie, les copules elliptiques sont appréciées dans la littérature en hydrologie étant donnée leur flexibilité et la facilité avec laquelle on peut les paramétriser. Cependant, dans le contexte de la modélisation de valeurs extrêmes, nous les écartons, conformément aux conclusions de [Renard and Lang, 2007], puisque cette famille de copule tend à sous-estimer la dépendance lorsque les marginales sont constituées à la fois de v.a. extrêmes et de v.a. non-extrêmes.

4 Études de cas

Copule BB1 : Voir [Joe, 1997]

4.1 Lac Champlain

La station USW00014742 située à Burlington, au Vermont comporte quelques données manquantes. Afin d'interpoler sur celles-ci, l'approche recommandée par [Shaw et al., 2010] est de moyenner les observations obtenues sur les stations avoisinantes. De plus, comme cette station a fermé le 3 juin 1943, nous avons compensé, comme indiqué par [Jalbert et al., 2019], avec la station de l'aéroport de Burlington, soit la station USW00014742, laquelle a commencé ses opérations le 1^{er} décembre 1940.

Pour le test d'indépendance séquentielle, les données du Lac Champlain, les résultats statistiques descriptives liées aux *p-values* obtenues avec le test de [Genest and Rémillard, 2004] sur chacune des années sont résumées dans le tableau 1.

Nombre d'événements extrêmes par année						
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	9.00	17.00	18.00	18.23	20.00	24.00
<i>P-values</i> pour le test d'indépendance séquentielle						
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
<i>K</i>	0.006	0.208	0.466	0.480	0.745	0.994

TABLEAU 1 – Statistiques descriptives des *p-values* obtenues sur chacune des années lors du test d'indépendance séquentielle proposé par [Genest and Rémillard, 2004] pour le Lac Champlain.

Afin de vérifier la stationnarité des distributions, le test de Mann-Kendall révèle une statistique de test égale à 0.0401 et une *p-value* pour l'hypothèse alternative bi-latérale de 0.0028164. Conséquemment, on rejette l'hypothèse nulle du test et on conclu qu'il existe une tendance ascendante dans la distribution de la suite de v.a. *K*.

Pour ce qui est de la détermination du seuil *u* servant à séparer les événements extrêmes des saisonniers, on applique l'approche proposée par [Bader et al., 2018] pour trouver un seuil optimal de 26.67mm de pluie. Cette quantité correspond au 86.39^e percentile des observations empiriques de *K* et permet de conserver 340 observations en excédent de seuil. Conformément à [Choulakian and Stephens, 2001], les tests d'Anderson-Darling et de Cramervon Mises pour vérifier l'adéquation des excès de seuil avec la famille de loi GPD offrent des *p-value* de 0.132 et 0.130. Conséquemment, on ne peut rejeter l'hypothèse nulle que les excès de seuil sont issues de la loi de Pareto généralisée au seuil de 5%.

Pour modéliser Z , comme la p-value du test de Mann-Kendall est de 0.512, on ne peut rejeter l'hypothèse nulle qu'il n'y a pas de tendance. Conséquemment, une simple loi Normale est entraînée pour modéliser les précipitations saisonnières non-extrêmes. L'illustration 1 atteste de l'adéquation de la loi avec les données. Ainsi, on y voit que la loi normale stationnaire est adéquate, mais que cette adéquation n'est pas parfaite. D'ailleurs, le test de Shapiro-Wilk donne une p-value de 0.030, laissant présager que les données ne sont pas réellement normales... à retravailler... trouver une loi alternative à la loi normale... Faute d'avoir des test d'adéquation tenant compte de la

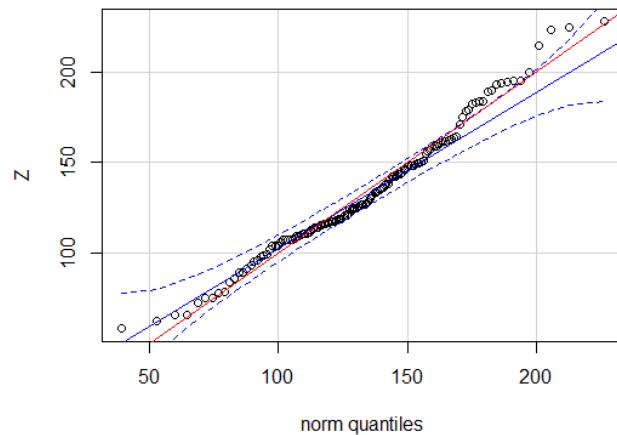


ILLUSTRATION 1 – Diagramme quantile-quantile de Z : La ligne bleue présente la tendance des points de rencontre entre les quantiles. Les pointillés présentent l'intervalle de confiance au seuil de 5% pour la droite de tendance et la ligne rouge est la diagonale d'adéquation parfaite. L'objectif est que la diagonale rouge se situe dans l'intervalle de confiance.

non-stationnarité, on utilise les statistiques d'Anderson-Darling et de Kolmogorov-Smirnov pour vérifier l'adéquation de manière quantitative. Les seuils observés de ces tests sont de

Du point de vue des excédents de seuil, le test de Mann-Kendall offre une statistique de test de 0.1705 et de 0.5681. Conséquemment, au seuil de 5%, on ne peut rejeter l'hypothèse nulle que les données proviennent de la loi normale non stationnaire ainsi paramétrée.

4.2 Lac Claire

5 Discussion

— How to write a Discussion — (À rédiger tout de suite après les résultats)

Ajouter les points suivants :

1. **donnez du poids à votre travail** (Mettez toujours en valeur vos résultats par rapport à ceux des autres, et non l'inverse)
2. **Interpréter les résultats**

3. Conséquences des résultats

4. Placez les arguments les plus importants au début de la discussion

Limites : Dans la méthodologie utilisée, mentionnons que le test statistique pour valider l'indépendance séquentielle des données utilisée est expérimentale. Également, la quantité de données ne permet pas d'avoir une appréciation fiable de la stationnarité de la dépendance entre les variables conformément à [Chebana and Ouarda, 2021]. Ceci étant dit, les hypothèses de stationnarité de la dépendance et de l'indépendance séquentielle des variables aléatoires réduit considérablement la complexité du modèle. En ce qui a trait à la principale faiblesses du modèle par rapport à ceux n'agrégant pas les jours de pluie est que cette agrégation réduit considérablement le nombre d'observations disponibles pour entraîner le modèle. Malgré tout, les résultats obtenus sont excellents et le fait de considérer des périodes de plusieurs jours fait en sorte que la loi GP ainsi entraînée obtient un estimateur de forme lui permettant d'avoir un support non fini.

conclusion : Un paragraphe concluant la discussion

1. redonnez le point fort, c'est-à-dire le résultat majeur et son apport original dans le champ scientifique concerné : Nos résultats ont montré que... ;
2. ne terminez jamais sur les travaux des autres, car cela met en doute la validité du travail présenté et cela en diminue la portée. En conséquence, la conclusion ne comporte aucune référence bibliographique ;
3. évitez les formules du type : il se pourrait que... Il serait possible de... Il pourrait être suggéré... Éventuellement... ;
4. évitez de suggérer des études à plus grande échelle! Cela sous-entend qu'elles doivent être conduites pour vérifier ce que vous avez fait ;
5. évitez de terminer par : Nous sommes en train d'étudier... Le lecteur répliquera : Plutôt que de publier cet article, attendons les résultats !

5.1 Pour aller plus loin (**à inclure ?**)

Pour aller plus loin dans la construction du modèle, il resterait à considérer, non seulement le temps, mais aussi l'espace dans le modèle. En effet, [Baigorria et al., 2007] démontre qu'il existe une forte corrélation entre les stations avoisinantes. À cette fin, certains auteurs utilisent des méthodes de régression telles que [Hession and Moore, 2011] et [Beguería and Vicente-Serrano, 2006] afin de réaliser un lissage spatial des paramètres de leur modèle de précipitations. Cependant, une telle approche sous-tend l'hypothèse que la dépendance unissant les différentes stations est linéaire. Or, comme l'a démontré [Zhang and Singh, 2019] cette relation n'est pas linéaire puisque la dépendance tend à être plus forte dans les ailes de distributions conjointes. C'est pourquoi ce dernier propose l'utilisation de copules en vignes. Cependant, cette dernière approche ajoute une grande complexité au modèle et ne permet pas de faire un lissage à grande échelle. Ainsi, pour aller plus loin, on pourrait tester une approche utilisant les réseaux de neurones convolutifs récurrents telle que présentée dans [Liu et al., 2016].

Finalement, pour revenir à la motivation initiale de [Jalbert et al., 2019] et de notre modèle, l'objectif ultime serait de prédire les inondations de manière exhaustive. Pour se faire, en s'inspirant de [Riboust and Brissette, 2016], quelqu'un pourrait proposer un modèle tri-varié où les variables en jeu seraient les précipitations hivernales ainsi que la température et les précipitations printanières. Ainsi, on aurait un modèle complet pour prédire la quantité totale d'eau accumulée lors du printemps.

Remerciements

L'auteur aimeraient remercier l'INRS ainsi que la Chaire d'actuariat de l'Université Laval pour avoir financé ce projet. Il aimeraient également remercier les professeurs Étienne Marceau et Hélène Cossette pour leurs conseils et commentaires, de même que Fateh Chebana pour avoir supervisé le projet.

Références

- [Bader et al., 2016] Bader, B., Yan, J., and Zhang, X. (2016). Automated threshold selection for extreme value analysis via goodness-of-fit tests with application to batched return level mapping. *arXiv preprint arXiv :1604.02024*.
- [Bader et al., 2018] Bader, B., Yan, J., Zhang, X., et al. (2018). Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate. *The Annals of Applied Statistics*, 12(1) :310–329.
- [Baigorria et al., 2007] Baigorria, G. A., Jones, J. W., and O'Brien, J. J. (2007). Understanding rainfall spatial variability in southeast usa at different timescales. *International Journal of Climatology : A Journal of the Royal Meteorological Society*, 27(6) :749–760.
- [Beguería and Vicente-Serrano, 2006] Beguería, S. and Vicente-Serrano, S. M. (2006). Mapping the hazard of extreme rainfall by peaks over threshold extreme value analysis and spatial regression techniques. *Journal of applied meteorology and climatology*, 45(1) :108–124.
- [Chebana and Ouarda, 2021] Chebana, F. and Ouarda, T. B. (2021). Multivariate non-stationary hydrological frequency analysis. *Journal of Hydrology*, 593 :125907.
- [Choulakian and Stephens, 2001] Choulakian, V. and Stephens, M. A. (2001). Goodness-of-fit tests for the generalized pareto distribution. *Technometrics*, 43(4) :478–484.
- [Genest and Rémillard, 2004] Genest, C. and Rémillard, B. (2004). Test of independence and randomness based on the empirical copula process. *Test*, 13(2) :335–369.
- [Gudendorf and Segers, 2010] Gudendorf, G. and Segers, J. (2010). Extreme-value copulas. In *Copula theory and its applications*, pages 127–145. Springer.
- [Hession and Moore, 2011] Hession, S. L. and Moore, N. (2011). A spatial regression analysis of the influence of topography on monthly rainfall in east africa. *International journal of climatology*, 31(10) :1440–1456.
- [Jalbert et al., 2019] Jalbert, J., Murphy, O. A., Genest, C., and Nešlehová, J. G. (2019). Modelling extreme rain accumulation with an application to the 2011 lake champlain flood. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 68(4) :831–858.
- [Joe, 1997] Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.
- [Kao and Govindaraju, 2007] Kao, S.-C. and Govindaraju, R. S. (2007). A bivariate frequency analysis of extreme rainfall with implications for design. *Journal of Geophysical Research : Atmospheres*, 112(D13).
- [Khaliq et al., 2006] Khaliq, M. N., Ouarda, T. B., Ondo, J.-C., Gachon, P., and Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations : A review. *Journal of hydrology*, 329(3-4) :534–552.
- [Liu et al., 2016] Liu, Y., Racah, E., Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., Wehner, M., Collins, W., et al. (2016). Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv preprint arXiv :1605.01156*.

- [Nelsen, 2006] Nelsen, R. B. (2006). An introduction to copulas. Springer, New York. *MR2197664*.
- [Nelson and Gerhardt, 2011] Nelson, B. L. and Gerhardt, I. (2011). Modelling and simulating non-stationary arrival processes to facilitate analysis. *Journal of Simulation*, 5(1) :3–8.
- [Northrop et al., 2015] Northrop, P., Attalides, N., and Jonathan, P. (2015). Cross-validatory extreme value threshold selection and uncertainty with application to ocean storm severity. *arXiv preprint arXiv:1504.06653*.
- [Renard and Lang, 2007] Renard, B. and Lang, M. (2007). Use of a gaussian copula for multivariate extreme value analysis : some case studies in hydrology. *Advances in Water Resources*, 30(4) :897–912.
- [Riboult and Brissette, 2016] Riboult, P. and Brissette, F. (2016). Analysis of lake Champlain/Richelieu river's historical 2011 flood. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques*, 41(1–2) :174–185.
- [Salvadori and De Michele, 2007] Salvadori, G. and De Michele, C. (2007). On the use of copulas in hydrology : theory and practice. *Journal of Hydrologic Engineering*, 12(4) :369–380.
- [Shaw et al., 2010] Shaw, E. M., Beven, K. J., Chappell, N. A., and Lamb, R. (2010). *Hydrology in practice*. CRC press.
- [Vandenberghhe et al., 2010] Vandenberghhe, S., Verhoest, N., and De Baets, B. (2010). Fitting bivariate copulas to the dependence structure between storm characteristics : A detailed analysis based on 105 year 10 min rainfall. *Water resources research*, 46(1).
- [Zhang and Singh, 2019] Zhang, L. and Singh, V. P. (2019). *Copulas and their applications in water resources engineering*. Cambridge University Press.