



Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du
Diplôme d'Actuaire EURIA
et de l'admission à l'Institut des Actuaire

le 21 Septembre 2021

Par : François BRÉHIN

Titre : Risque de crue de la Seine sur le bassin parisien

Confidentialité : Non

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

**Membres présents du jury de l'Institut
des Actuaire :**

Alexandre YOU

Ederm LE ROUX

Benjamin BALTESAR

Entreprise :

Sia Partners

Signature :

Membre présent du jury de l'EURIA :

Pierre AILLIOT

Directeur de mémoire en entreprise :

Clément BILLORE

Signature :

Invité :

Claire NICOLLE

Signature :

**Autorisation de publication et de mise en ligne sur un site de diffusion
de documents actuariels**

(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise :

Signature du candidat :

Résumé

L'évaluation des risques naturels sera de plus en plus présente au sein des métiers de l'assurance dans les années à venir. L'accélération du dérèglement climatique ces dernières années, comme le montre le dernier rapport du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) entraîne une explosion des catastrophes naturelles et menace la solvabilité du système financier. En France, dans ce contexte, le régulateur commence à inciter les institutions financières à mieux appréhender ces risques au travers d'exercices climatiques. Le risque inondation en est le parfait exemple. Premier risque naturel en France en termes de dommages provoqués et de communes exposées, il devient de ce fait primordial pour les assureurs de se prémunir face à cet aléa. De par sa situation géographique en aval immédiat de la confluence de la Seine et de la Marne, le bassin parisien n'échappe pas à cette menace et en fait l'une des métropoles européennes les plus exposées au risque d'inondation. Selon l'Établissement public territorial de bassin Seine grands lacs (EPTB), une crue semblable à celle observée en 1910 (ayant atteint 8,62 mètres) engendrerait des pertes d'*a minima* 20 milliards d'euros. Les pouvoirs publics se sont penchés depuis des années sur ce problème en imaginant de nombreux projets afin de se prémunir contre ce risque, sans jamais avoir trouvé de solution concluante. Il devient alors nécessaire pour les groupes assurantiels de se protéger contre ce risque en se munissant d'outils d'aide à la décision pour l'appréhender et le mesurer que ce soit dans des exercices d'évaluation de la solvabilité, de gestion de portefeuille ou encore de prévention. La complexité de l'évaluation du risque inondation vient du fait qu'il fait intervenir un grand nombre de paramètres. Ce mémoire vise à proposer une méthodologie en identifiant, cartographiant et probabilisant le risque sur la base des données d'un portefeuille MRH en y intégrant des variables externes issues de l'open data. La première étape consiste à identifier et cartographier les expositions sur différents niveaux de crues de la Seine. Dans un second temps, nous calibrerons des taux de destruction afin d'associer des pertes aux différentes crues de la Seine. Enfin, il conviendra d'évaluer le temps de retour des crues au travers de scénarios météorologiques probabilisés et d'en déduire la distribution des pertes brutes de réassurance.

Mots clés: Risque de crue, Risque inondation, Catastrophe Naturelle (CAT NAT), Risque bicentenaire, Période de retour, Modélisation climatique, Taux de destruction, Cartographie des risques, Assurance Dommage aux biens, Vulnérabilité, Séries temporelles, Copules, Réchauffement climatique, GIEC.

Abstract

The assessment of natural risks will be even more core to the insurer's strategy in the years to come. The acceleration of climate change in recent years, as shown in the latest Intergovernmental panel on climate change (IPCC) report, is leading to a large increase of natural disasters and threatens the solvency of the financial system. In France, in this context, the regulator is beginning to encourage financial institutions to better understand these risks through climate risk exercises. The flood risk is a perfect example. As the main natural risk in France in terms of damage caused and the number of towns exposed, it is becoming essential for insurers to take protective measures against this hazard. Due to its geographical location immediately downstream from the confluence of the Seine and Marne rivers, the Paris basin is exposed to this threat and is one of the European cities most exposed to the risk of flooding. According to the Public territorial basin establishment (PTBE), a flood similar to the one observed in 1910 (which reached 8.62 metres) would result in losses of at least 20 billion euros. The public authorities have been looking into this problem for years, devising numerous projects to protect people against this risk, without ever having found a conclusive solution. It is therefore necessary for insurance groups to protect themselves against this risk by developing decision making tools to apprehend and measure this risk, whether in solvency assessment, portfolio management or prevention exercises. The complexity of flood risk assessment stems from the fact that it involves a large number of parameters. This thesis aims at proposing a methodology to identify, map and probabilise the risk based on the data of a property insurance portfolio by integrating open data variables. The first step is to identify and map the exposures over different flood levels of the Seine. Secondly, we will calibrate the destruction rate in order to associate losses with the different Seine floods. Finally, we will evaluate the return time of floods through probabilised meteorological scenarios and then deduce the distribution of gross reinsurance losses.

Keywords: Flood risk, Natural disaster (NAT CAT), Bicentennial risk, Return period, Climate modelling, Destruction rate, Risk mapping, Property insurance, Vulnerability, Time series, Copula, Global warming, IPCC.

Note de synthèse

Contexte et problématique

Les catastrophes naturelles (CAT NAT) constituent aujourd’hui un enjeu de taille pour les assureurs. Leur faible fréquence d’occurrence donnant accès à un historique de sinistralité réduit implique une difficulté accrue dans la quantification de ce risque. De plus, leur étendue géographique les porte au grade d’évènements exceptionnels, touchant simultanément des centaines voire des milliers ou des dizaines de milliers de polices en même temps.

Pour toutes ces raisons, la quantification des catastrophes naturelles est à la fois primordiale et complexe. La directive Solvabilité II a obligé les assureurs à mieux appréhender les risques CAT, or, le développement de modèles internes pour l’évaluation de la solvabilité et le pilotage nécessite une compréhension et une quantification très fine de ces risques extrêmes, obligeant les assureurs à se munir d’outils et de modèles leur permettant d’améliorer la gestion de leurs risques.

De plus, la prise de conscience récente sur le changement climatique voit les régulateurs se pencher de plus en plus sérieusement sur la question. L’Autorité de contrôle prudentiel et de résolution (ACPR) a réalisé en 2020 un exercice pilote afin d’évaluer les impacts du réchauffement climatique sur le système bancaire et assurantiel en France. La Banque Centrale Européenne lancera des travaux similaires en 2022. Les régulateurs poussent ainsi les institutions financières à se doter d’une gouvernance pour évaluer et maîtriser les risques émergents en conséquence directe du réchauffement climatique. Le rapport du GIEC paru cet été accentue d’autant plus cette nécessité de mieux appréhender les risques qui découlent d’un tel changement.

La question de la gestion et de la modélisation des CAT NAT est donc un phénomène d’actualité. C’est dans ce cadre que s’inscrit ce mémoire, au travers d’une étude du risque inondation et plus précisément du risque de crue de la Seine sur le bassin parisien. Nous avons fait ce choix puisque les données concernant la crue de la Seine et le bassin parisien sont plus nombreuses, et également pour la complexité de la mise en place d’un tel modèle. Cependant, cette méthode reste généralisable à l’ensemble des bassins fluviaux du globe.

Démarche

La mise en place d'un modèle d'évaluation du risque CAT NAT suit une démarche modulaire bien précise. C'est la compilation de ces étapes qui permettra d'obtenir une vision globale de la quantification du risque inondation. Les 3 principaux modules sont :

- **Le module exposition** : C'est l'étude du portefeuille permettant d'apporter les informations nécessaires sur la localisation, les caractéristiques et les valeurs des biens assurés ;
- **Le module vulnérabilité** : Ce module sert à déterminer, en fonction de l'intensité de l'évènement, les pertes économiques. Pour cela, nous devons évaluer comment les biens sont touchés en fonction de leurs caractéristiques ;
- **Le module aléa** : Il consiste à modéliser le risque physique, c'est-à-dire à retranscrire statistiquement la probabilité d'apparition de la catastrophe. Afin de la décrire, il s'agit de modéliser sa localisation, sa fréquence de survenance et son intensité.

Module exposition

La première étape consiste à cartographier nos expositions. À partir d'une base de données de contrats Multirisque habitation (MRH), nous avons commencé par déterminer la localisation GPS (latitude/longitude) de chacune de nos expositions.

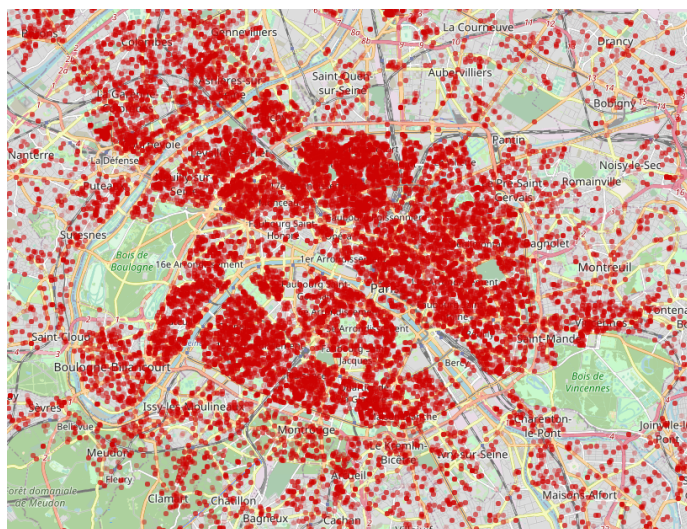


FIGURE 1 – Cartographie de nos expositions sur le bassin parisien.

Ensuite, nous avons cherché à évaluer l'altitude associée à chaque bien assuré, indicateur de vulnérabilité à une crue d'intensité donnée. Pour cela, nous avons utilisé des données topographiques provenant de la base de données en open source *Copernic*. À l'aide du logiciel **R**, nous avons donc dressé une cartographie de la topographie du bassin

parisien avec une maille d'environ 25 mètres par 25 mètres, constituant la résolution la plus fine à notre disposition. Cette résolution s'est avérée suffisante dans notre étude.

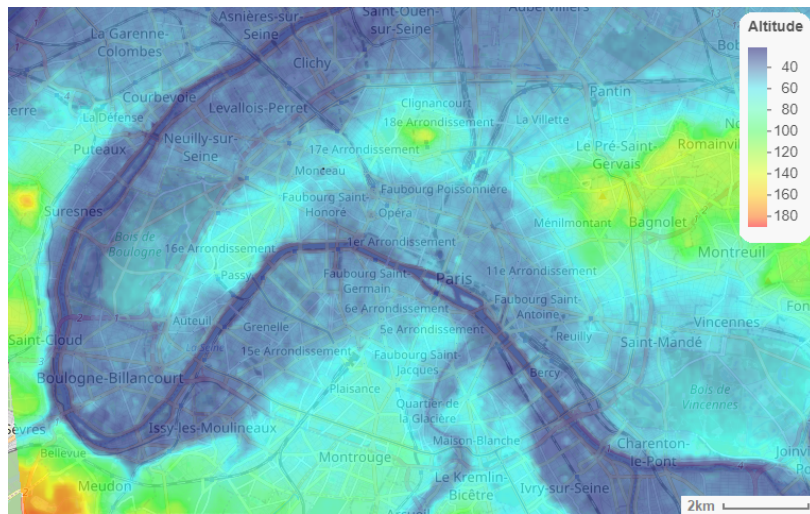


FIGURE 2 – Carte topographique du bassin parisien (base de données *Copernicus*).

Par la suite et au moyen de ces données nous avons cartographié la Seine traversant Paris. Il s'agissait globalement de récupérer les points les moins élevés de la cartographie à l'exception de certains bruits, et ce de façon dynamique, car la profondeur du fleuve varie en fonction de la distance avec son embouchure. La ville de Paris est globalement à une altitude variant entre 25 et 100 mètres au-dessus du niveau de la mer pour les points les plus hauts. Nous avons donc cartographié le fleuve et réalisé des cartes d'empreinte de crue, représentant les zones du bassin immergées lors d'une crue de la Seine. Voici un exemple pour une crue type 1910 (8,62 mètres à l'échelle du pont d'Austerlitz) :

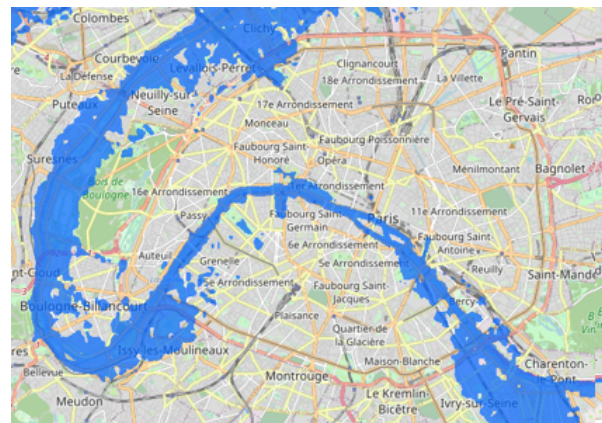


FIGURE 3 – Reproduction de la carte d'empreinte de la crue de 1910.

En combinant cartographie des expositions et cartographie des empreintes de crues, nous pouvons alors déterminer exactement quelles expositions assurées seront touchées par une crue de la Seine. Cela nous permet d’obtenir une vision agrégée de la part totale d’expositions touchées par une crue de la Seine d’une intensité donnée.

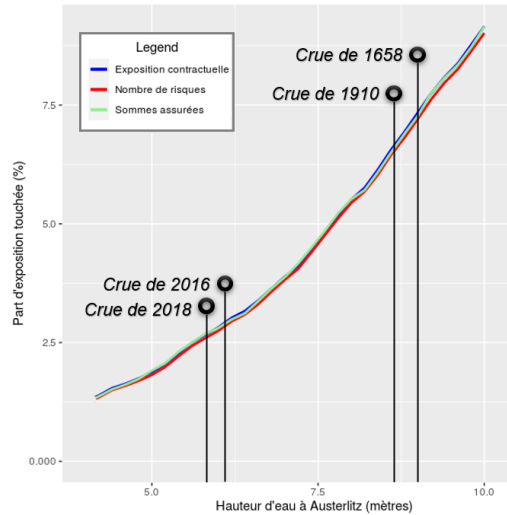


FIGURE 4 – Pourcentage d’exposition dans l’empreinte de crue selon la hauteur d’eau.

Maintenant que nous possédons une cartographie fine des expositions sous risque conditionnellement à l’intensité de la crue, il nous faut déterminer à quel point ces expositions seront affectées par une telle crue et les pertes qui en résultent.

Module vulnérabilité

Il faut maintenant déterminer dans quelle mesure une exposition sera touchée. En effet, une maison plain-pied ou un appartement au cinquième étage ne seront pas touchés de façon similaire. De la même manière, avoir un faible niveau d’eau ou voir son appartement complètement immergé influera grandement sur les conséquences financières. La durée de submersion est également une variable explicative du montant de la perte. Aussi, il nous faut estimer les pertes pour chaque exposition touchée. S’il est impossible de les estimer une à une car trop de facteurs entrent en jeu, nous pouvons chercher à les estimer en moyenne. De façon évidente, les pertes potentielles sont liées à la somme assurée, car plus le bien est de valeur plus il coûtera cher lors de la survenance d’une catastrophe.

Ainsi, nous basons nos pertes brutes sur la mise en place de courbes de vulnérabilité, qui vont déterminer selon l’intensité de la catastrophe le taux de destruction du bien assuré, qui combiné au capital assuré nous donnera les pertes correspondantes par exposition. Afin d’être le plus précis possible, nous nous sommes appuyés sur une base de données de sinistres inondation et sur des hypothèses basées sur des études afin de mettre en place des courbes de destruction conditionnelles aux caractéristiques du bien assuré

et à la dynamique de crue. Hauteur d'eau maximale atteinte, altitude de l'exposition, durée d'immersion, type d'habitation, année de construction du bien assuré, tels sont les paramètres selon lesquels nous avons défini nos taux de destruction spécifiques pour chaque exposition, de manière à refléter au mieux les dégâts réels lors de l'apparition d'une catastrophe.

Au vu du nombre de paramètres entrant en jeu dans nos taux de destruction, nous ne pouvons pas représenter graphiquement l'intégralité de ces taux, mais si l'on regarde de plus près les taux de destruction obtenus selon la hauteur d'immersion et l'année de construction d'une maison immergée durant 3 jours, voici ce que nous obtenons :

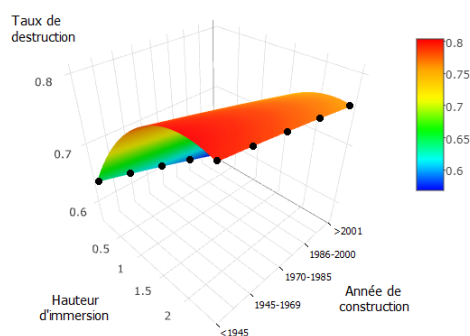


FIGURE 5 – Taux de destruction selon la hauteur d'immersion et l'année de construction du bien assuré (pour une maison et une durée d'immersion de 3 jours).

Afin d'avoir une seconde vision, nous pouvons effectuer des tables de destruction selon les caractéristiques du bien assuré et la dynamique de crue. Voici deux exemples, le premier pour une maison et le second pour un appartement au premier étage :

Durée d'immersion ↘ Année de construction ↓	Hauteur d'immersion											
	0.1 mètre			0.5 mètre			1 mètre			1,5 mètres		
	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j
≥ 2001	0.32	0.55	0.70	0.49	0.66	0.77	0.56	0.70	0.80	0.60	0.73	0.82
1986-2000	0.36	0.57	0.71	0.51	0.67	0.78	0.58	0.72	0.81	0.62	0.74	0.83
1970-1985	0.39	0.59	0.73	0.53	0.69	0.79	0.60	0.73	0.82	0.63	0.76	0.84
1945-1969	0.42	0.61	0.74	0.56	0.70	0.80	0.62	0.74	0.83	0.65	0.77	0.85
<1945	0.44	0.63	0.75	0.58	0.72	0.81	0.63	0.76	0.84	0.67	0.78	0.85

FIGURE 6 – Taux de destruction retenus adaptés aux caractéristiques de la maison.

Durée d'immersion ↘ Année de construction ↓	Appartement au 1er étage : Hauteur d'immersion (à partir du niveau 0 de l'immeuble)								
	≤ 2.5 mètres			3 mètres			3.5 mètres		
	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours
≥ 2001	0.05	0.05	0.05	0.58	0.72	0.81	0.66	0.78	0.85
1986-2000	0.05	0.05	0.05	0.60	0.74	0.82	0.68	0.79	0.86
1970-1985	0.05	0.05	0.05	0.62	0.75	0.83	0.70	0.80	0.86
1945-1969	0.05	0.05	0.05	0.64	0.76	0.84	0.71	0.81	0.87
<1945	0.05	0.05	0.05	0.66	0.77	0.85	0.72	0.82	0.88

FIGURE 7 – Taux de destruction retenus adaptés aux caractéristiques de l'appartement (1er étage).

Ce deuxième volet couplé avec le premier permet ainsi de déterminer lors de l'apparition d'une crue les expositions touchées et les pertes associées à l'aide de nos taux de destruction dynamiques. En sommant toutes ces pertes, nous obtenons les pertes brutes totales associées à la crue. Encore une fois, à une dynamique de crue correspond une courbe de pertes brutes pour le portefeuille il est donc impossible de toutes les représenter. Nous présentons alors une courbe de pertes brutes agrégées en fixant la durée d'immersion de toutes les expositions touchées à moins de deux jours, de deux à quatre jours ou plus de quatre jours (hypothèse non réaliste, mais réalisée pour l'exemple) :

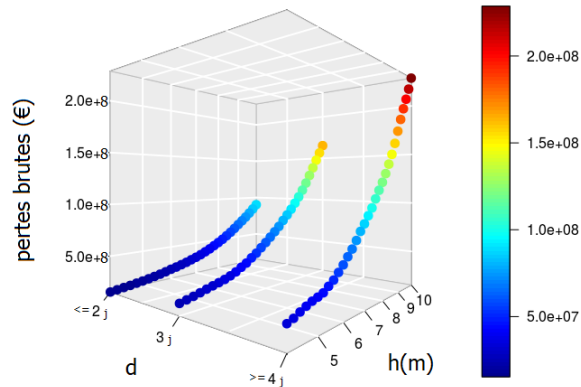


FIGURE 8 – Évaluation des pertes brutes selon la dynamique de crue (hauteur/durée).

Maintenant que nous savons déterminer précisément quelles expositions seront touchées et quelles sont les pertes associées selon la dynamique temporelle de la crue, nous devons modéliser l'aléa physique afin de probabiliser nos pertes avec les chances d'observer chacune des crues potentielles de la Seine.

Module aléa

Afin de modéliser l'aléa, nous avons collecté les données *HYDRO* qui regroupent les relevés de hauteur d'eau de différents bassins fluviaux en France depuis des dizaines d'années. Nous avons ainsi retenu plusieurs stations présentes sur le bassin versant de la Seine, pour former un réseau hydrométrique homogène dans le but de calculer les durées de propagation des ondes de crue.

Nous avons de cette façon pu estimer que la durée moyenne d'écoulement de l'eau entre Auxerre et Paris était d'environ 2,5 jours par exemple. Le but de cette démarche est de connaître la dynamique d'écoulement du fleuve, ce qui va nous servir dans la modélisation de nos crues. De fait, le phénomène principal expliquant la survenance de crues localisées est la présence de précipitations intenses et prolongées sur le bassin. Il faut cependant bien noter que ce n'est pas simplement la pluie qui tombe à Paris qui formera une crue dans la capitale, mais la pluie tombant sur l'ensemble du bassin versant. Ainsi, plusieurs crues historiques se sont formées du fait de très fortes précipitations sur la Marne, à quelques dizaines voire centaines de kilomètres de Paris. C'est pourquoi

connaître la dynamique du fleuve nous permettra de connaître approximativement le temps entrepris par une goutte de pluie tombant dans le bassin versant pour atteindre la capitale et donc influencer sur la hauteur de l'eau au pont d'Austerlitz. Nous avons par la suite étudié des relevés pluviométriques provenant de la base de données *ISD* et nous avons tracé les zones d'influence de ces stations en cartographiant chaque point de la carte à sa station *ISD* la plus proche.

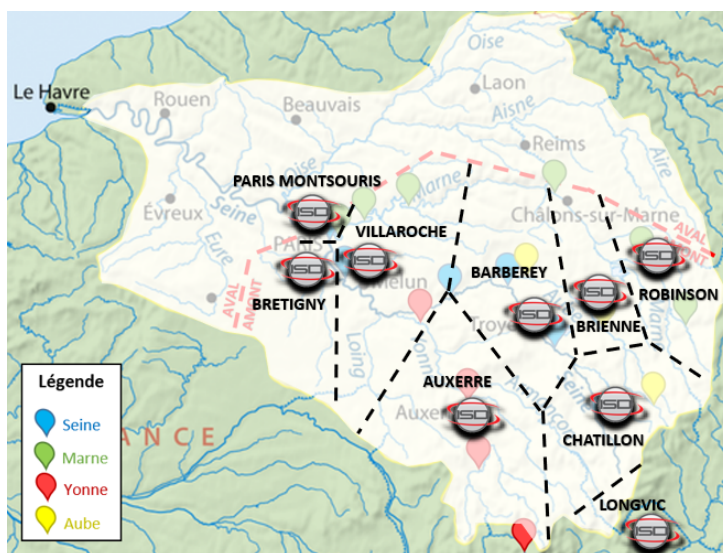


FIGURE 9 – Cartographie des stations *ISD* étudiées.

Nous connaissons alors approximativement le lag temporel existant entre la pluie tombant proche d'une station *ISD* et son écoulement jusqu'à Paris. Nous pouvons ainsi mettre en place, avec les données historiques, un modèle explicatif entre la pluie tombée dans le bassin versant de la Seine, corrigée des lags ainsi calculés et exprimés en demi-journées, et la hauteur d'eau relevée à Austerlitz.

Une fois notre modèle en place, nous allons pouvoir effectuer des simulations, afin d'obtenir des périodes de retour de crues historiques ainsi qu'une distribution des pertes annuelles brutes en lien avec les deux modules effectués plus tôt. Pour réaliser ces simulations, nous avons besoin de connaître la loi des précipitations du bassin versant. Pour cela, nous avons calibré une loi d'adéquation sur la pluviométrie historique conditionnellement à la station et au mois étudiés, car la distribution ne sera pas la même en fonction de la géolocalisation et selon la période de l'année considérée.

Cependant, les précipitations au même titre que tous les phénomènes météorologiques sont dynamiques spatialement et temporellement. Ainsi, afin de modéliser la structure de dépendance témoignant de la probabilité qu'il pleuve dans une zone donnée s'il pleut dans une autre a été modélisée au travers de l'utilisation de copules gaussiennes. Nous avons maintenant tous les outils pour effectuer nos simulations, au travers de trajectoires annuelles continues par pas de douze heures.

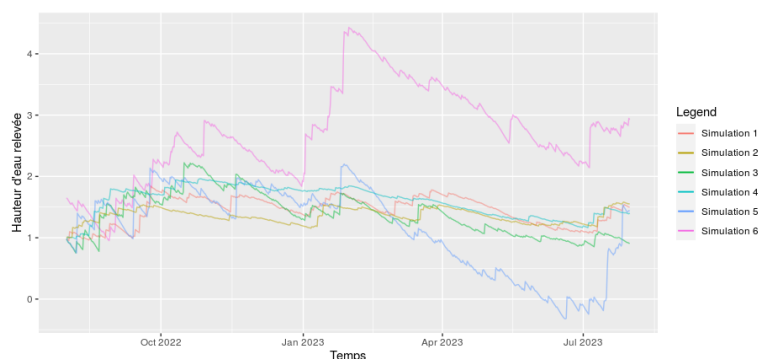


FIGURE 10 – Exemple de simulations annuelles de la hauteur d’eau à Austerlitz.

Résultats

Grâce à ces 3 modules, nous avons pu simuler 10 000 trajectoires et calculer nos périodes de retour selon l’intensité d’une crue donnée, ainsi que tracer la distribution de nos pertes brutes annuelles.

Hauteur d’eau (m)	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9
PdR estimée τ (a)	5	8	12	21	45	104	168	357	1250
Pertes estimées (M€)	31,5	35,2	39,7	47,8	63,4	74,5	84,3	101	126

TABLE 1 – Périodes de retour et pertes selon la hauteur d’eau.

Ces calculs nous permettent également d’obtenir la cartographie du bassin parisien associée aux différentes périodes de retour et donc de construire une vision géographique du risque de crue de la Seine.

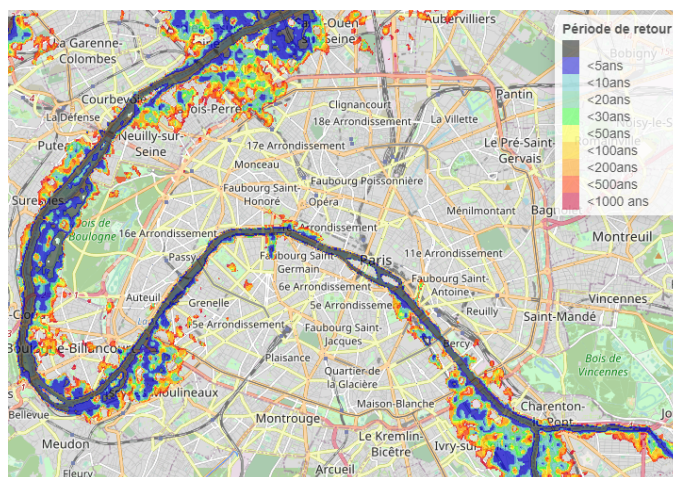


FIGURE 11 – Cartographie des zones sous risque pour les périodes de retour estimées.

Ouverture

Dans ce contexte inédit de changement climatique, nous avons cherché à mesurer celui-ci au travers de trois scénarios en jouant sur les paramètres de fréquence et d'intensité. Ces scénarios découlent d'une interprétation qualitative du rapport du GIEC et ne reflètent aucunement une réalité. Ils ont pour vocation à illustrer et poser un cadre d'application des travaux réalisés dans ce mémoire.

Les périodes de retour et les pertes associées à chaque hauteur d'eau ont été calculé pour chaque scénario :

- **Central**
- **SSP1-1.9** (Hypothèse d'augmentation de $+1,5^{\circ}\text{C}$) : Baisse de la fréquence de 5% / Augmentation de l'intensité de 5%
- **SSP2-4.5** (Hypothèse d'augmentation de $+2,5^{\circ}\text{C}$) : Hausse de la fréquence de 5% / Augmentation de l'intensité de 5%
- **SSP5-8.5** (Hypothèse d'augmentation de $+4,5^{\circ}\text{C}$) : Hausse de la fréquence de 10% / Augmentation de l'intensité de 10%

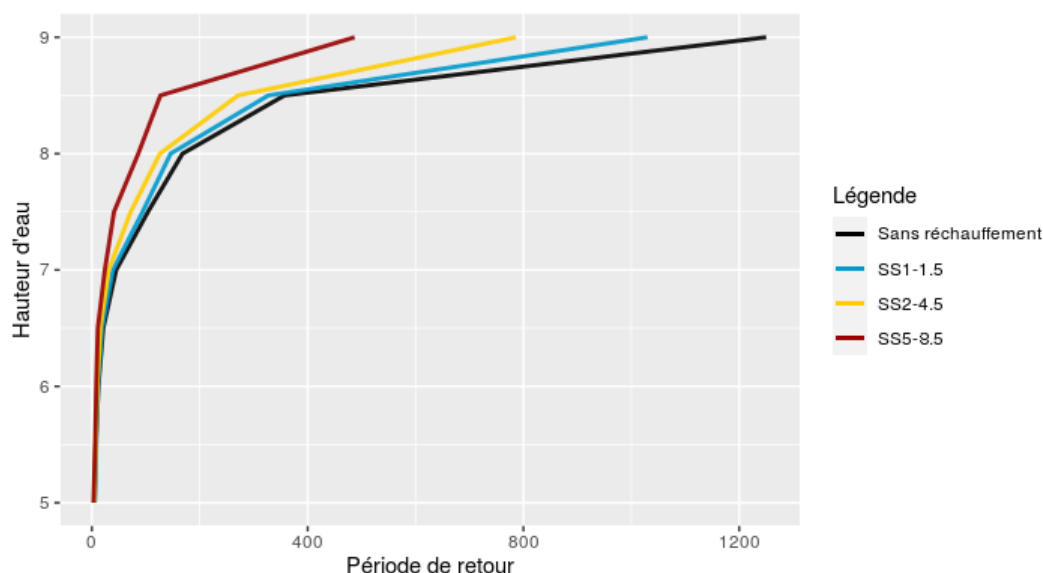


FIGURE 12 – Évolution des périodes de retour selon le scénario considéré.

On constate bien une augmentation du risque de crue de la Seine liée au réchauffement climatique. Nous pouvons ainsi tracer l'évolution de nos cartes de vulnérabilité par périodes de retour afin de visualiser les zones potentiellement à risque dans les années à venir.

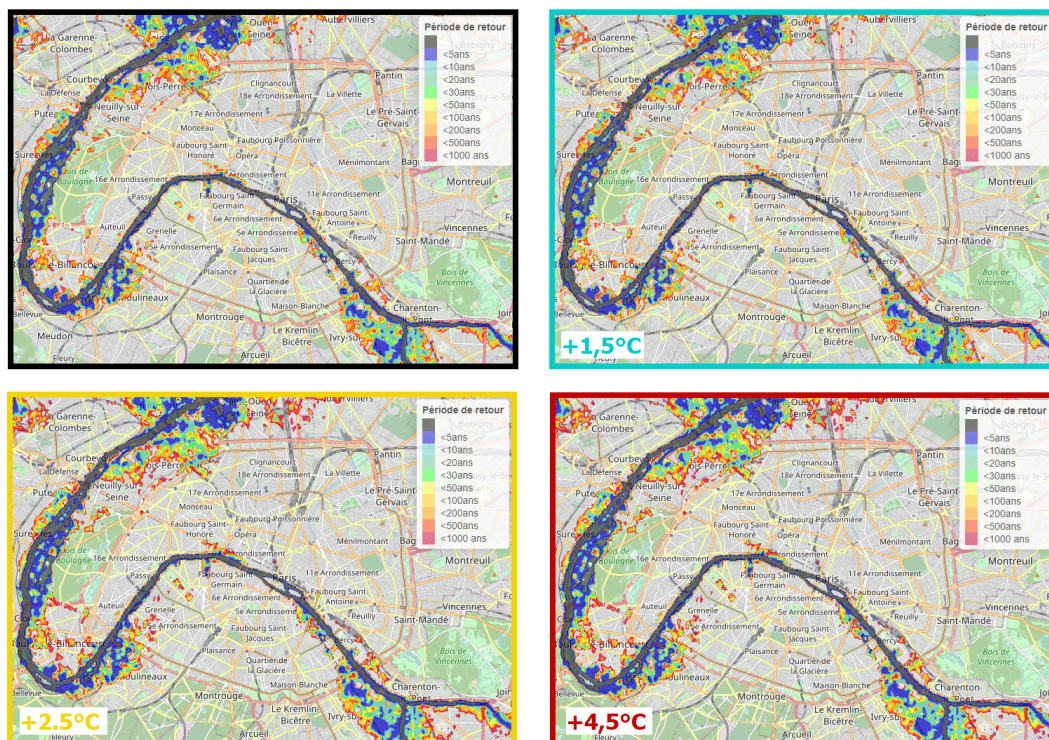


FIGURE 13 – Évolution des périodes de retour selon le scénario considéré.

Executive summary

Natural disasters (NAT CAT) are a major challenge for insurers today. Their low frequency of occurrence, which gives access to a reduced claims history, makes it more difficult to quantify the risk. Moreover, their geographical scope leads them to the rank of exceptional events, simultaneously affecting hundreds or even thousands or tens of thousands of policies at the same time.

It is for all these reasons that the quantification of natural catastrophes is both essential and complex. The Solvency II directive has obliged insurers to better understand CAT risks, and the development of internal models for solvency assessment and management requires a finer understanding and quantification of these extreme risks, forcing insurers to equip themselves with tools and models enabling them to improve their risk management.

In addition, the recent awareness of climate change has led regulators to take an increasingly serious look at the issue. The ACPR carried out a pilot climate exercise in 2020 to assess the impacts of global warming on the banking and insurance system in France. The European Central Bank will launch similar work in 2022. Regulators are thus pushing financial institutions to establish governance to assess and control the risks emerging as a direct consequence of global warming. The IPCC report published this summer further emphasises the need to better understand the risks arising from such change.

The issue of managing and modelling CATs is therefore a topical phenomenon. It is within this framework that this thesis is based on a study of the flood risk and more specifically the risk of flooding of the Seine in the Paris basin. This choice was made because of the data we had and because of the complexity of setting up such a model, but the method is also completely generalizable to all the river basins of the world.

Approach

The development of a CAT NAT risk assessment model follows a clear modular approach. It is the compilation of these steps that will provide an overall view of the quantification of flood risk. The 3 main modules are :

- **The exposure module** : This is the portfolio study that provides the necessary information on the location, characteristics and values of the insured assets ;
- **The vulnerability module** : This module is used to determine the economic

losses according to the intensity of the event. To do this, it is necessary to assess how the assets are affected according to their characteristics;

- **The hazard module** : This consists of modelling the physical risk, i.e. statistically transcribing the risk of the disaster occurring. In order to describe it, it is necessary to model its location, its frequency of occurrence and its intensity.

Exposure module

The first step is to map our exposures. Using a database of property insurance policies, we began by determining the GPS location (latitude/longitude) of each of our exposures.

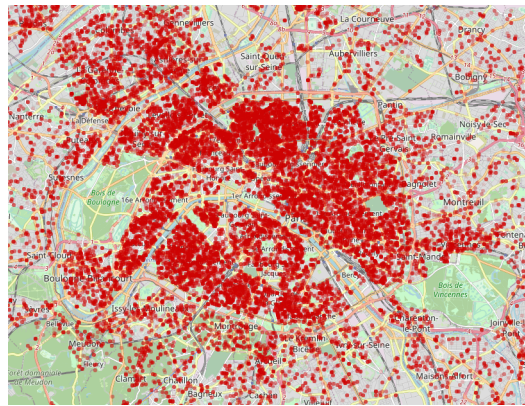


FIGURE 14 – Mapping of our exposures in the Paris basin.

Next, we sought to assess the elevation associated with each insured property, an indicator of vulnerability to a given flood intensity. For this purpose, we used topographic data from the open source database *Copernic*. Using **R** software, we mapped the topography of the Paris basin with a grid of approximately 25 metres by 25 metres, which is the finest resolution available to us. This resolution proved to be sufficient for our study.

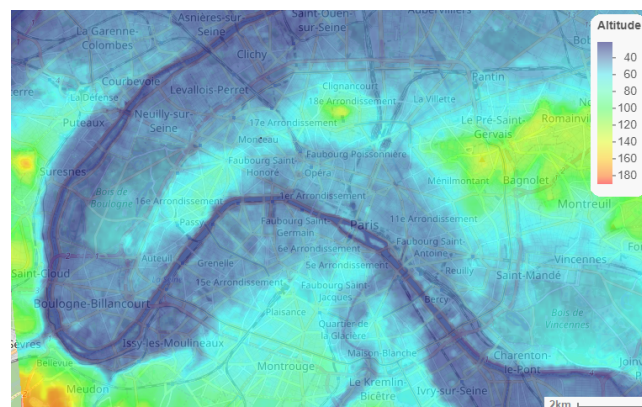


FIGURE 15 – Topographic map of the Paris basin (*Copernicus* data).

Then, using this data, we mapped the Seine River crossing Paris. This involved recovering the lowest points of the map, with the exception of certain noises, and this was done dynamically, as the depth of the river varies according to the distance from its mouth. The city of Paris is globally at an altitude varying between 25 and 100 metres above sea level for the highest points. We have therefore mapped the river and produced flood footprint maps, representing the areas of the basin submerged during a flood of the Seine. Here is an example for a 1910 flood (8.62 metres at the scale of the Austerlitz bridge) :

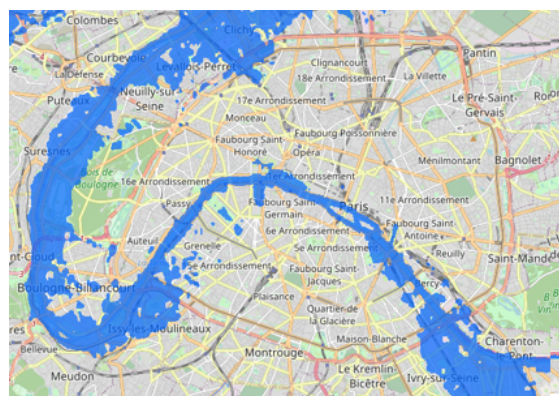


FIGURE 16 – Footprint map of the 1910 flood.

By combining exposure mapping and flood footprint mapping, we can then determine exactly which insured exposures will be affected by a Seine flood. This allows us to obtain an aggregated view of the total share of exposures affected by a Seine flood of a given intensity.

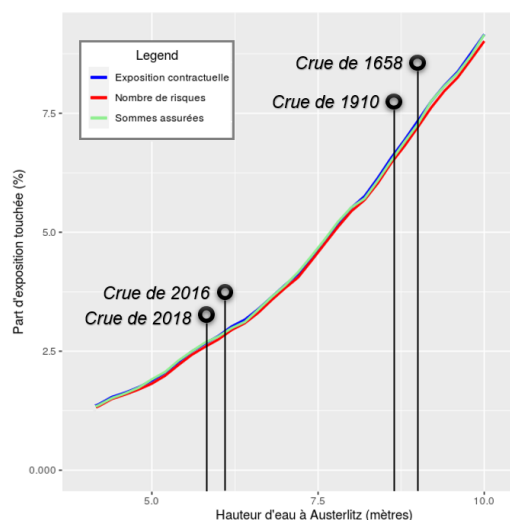


FIGURE 17 – Percentage of exposure in the flood footprint by water level.

Now that we have a fine map of the exposures at risk conditioned on the intensity of a flood, we need to determine the extent to which these exposures will be affected by a flood and the resulting losses.

Vulnerability module

It is now necessary to determine to what extent an exposure will be affected. A single-storey house or a fifth-floor flat will not be affected in the same way. Similarly, having a low water level or having your flat completely submerged will greatly influence the financial consequences. The duration of the flooding is also an explanatory variable for the amount of the loss. Also, we need to estimate the losses for each affected exhibition. While it is impossible to estimate them one by one, as too many factors come into play, we can try to estimate them on average. Obviously, potential losses are related to the sum insured, the more valuable the asset the more expensive it will be in the event of a catastrophe.

Thus, we base our gross losses by setting up vulnerability curves, which will determine, depending on the intensity of the catastrophe, the rate of destruction of the insured asset, which combined with the sum insured gives us the losses corresponding to the exposure in question. In order to be as accurate as possible, we have used a database of flood losses and assumptions based on studies to establish destruction rates and curves conditional on the characteristics of the insured property and the flood dynamics. Maximum water height, exposure altitude, duration of immersion, type of dwelling, year of construction of the insured property, these are the parameters according to which we have defined our specific destruction rates for each exposure, allowing us to best reflect the actual damage during a disaster.

Due to the number of parameters involved in our destruction rates, we cannot graphically represent all of these rates, but if we take a closer look at the destruction rates obtained according to the height of immersion and the year of construction of a house immersed for 3 days, here is what we obtain :

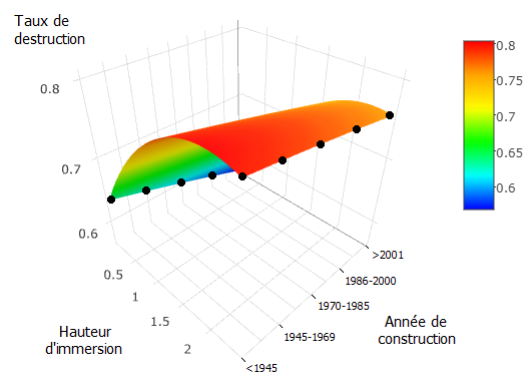


FIGURE 18 – Rate of destruction according to the height of immersion and the year of construction of the insured property (for a house and an immersion period of 3 days).

In order to have a second vision, we can make destruction tables according to the characteristics of the insured property and the flood dynamics. Here are two examples, the first for a house and the second for a first floor flat :

Durée d'immersion ↘ Année de construction ↓	Hauteur d'immersion											
	0,1 mètre			0,5 mètre			1 mètre			1,5 mètres		
	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j
>=2001	0.32	0.55	0.70	0.49	0.66	0.77	0.56	0.70	0.80	0.60	0.73	0.82
1986-2000	0.36	0.57	0.71	0.51	0.67	0.78	0.58	0.72	0.81	0.62	0.74	0.83
1970-1985	0.39	0.59	0.73	0.53	0.69	0.79	0.60	0.73	0.82	0.63	0.76	0.84
1945-1969	0.42	0.61	0.74	0.56	0.70	0.80	0.62	0.74	0.83	0.65	0.77	0.85
<1945	0.44	0.63	0.75	0.58	0.72	0.81	0.63	0.76	0.84	0.67	0.78	0.85

FIGURE 19 – Selected destruction rates adapted to the characteristics of the house.

Durée d'immersion ↘ Année de construction ↓	Appartement au 1er étage : Hauteur d'immersion (à partir du niveau 0 de l'immeuble)								
	≤ 2,5 mètres			3 mètres			3,5 mètres		
	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours
≥ 2001	0.05	0.05	0.05	0.58	0.72	0.81	0.66	0.78	0.85
1986-2000	0.05	0.05	0.05	0.60	0.74	0.82	0.68	0.79	0.86
1970-1985	0.05	0.05	0.05	0.62	0.75	0.83	0.70	0.80	0.86
1945-1969	0.05	0.05	0.05	0.64	0.76	0.84	0.71	0.81	0.87
<1945	0.05	0.05	0.05	0.66	0.77	0.85	0.72	0.82	0.88

FIGURE 20 – Selected destruction rates adapted to the characteristics of the flat (1st floor).

This second view, coupled with the first, allows us to determine the exposures affected by a flood, and the associated losses, using our dynamic destruction rates. By summing all these losses, we obtain the total gross losses associated with the flood. Again, each flood dynamic has a corresponding gross loss curve for the portfolio, so it is impossible to represent all of them. We then present an aggregated gross loss curve by setting the duration of immersion of all affected exposures to less than two days, two to four days or more than four days (not a realistic assumption, but achieved for the example) :

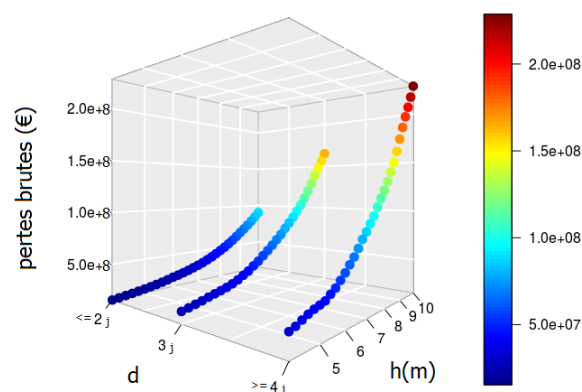


FIGURE 21 – Assessment of gross losses according to flood dynamics (height/duration).

Now that we know precisely which exposures will be affected and what the associated losses are according to the temporal dynamics of the flood, we need to model the physical hazard in order to probabilise our losses with the chances of observing each of the potential floods of the Seine.

Hazard module

In order to model the hazard, we have collected *HYDRO* data which constitute water height records, notably of the Seine over the years. We thus selected several stations present on the Seine catchment area, to form a homogeneous hydrometric network with the aim of calculating the propagation times of the flood waves.

In this way, we were able to estimate that the average duration of water flow between Auxerre and Paris was about 2.5 days, for example. The aim of this approach is to find out the flow dynamics of the river, which will be used to model our floods. Indeed, the main phenomenon explaining the appearance of floods is the presence of intense and prolonged rainfall. It should be noted, however, that it is not simply the rain that falls in Paris that will cause a flood in the capital, but the rain that falls over the entire catchment area. Several historical floods were formed because of very heavy rainfall on the Marne, for example, a few dozen or even hundreds of kilometres from Paris. Thus, knowing the dynamics of the river will allow us to know approximately the time taken for a drop of rain falling in the catchment area to reach the capital. We then studied rainfall records from the *ISD* database and plotted the zones of influence of these stations by mapping each point on the map to its nearest *ISD* station.

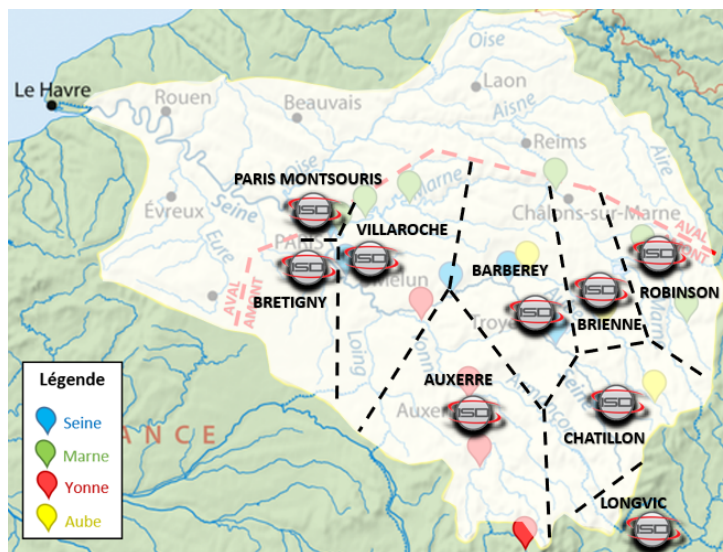


FIGURE 22 – Mapping of considered ISD stations.

We then know approximately the temporal lag between the rain falling near an *ISD* station and its runoff to Paris. We can thus set up, with the historical data, an explana-

tory model between the rainfall in the Seine catchment area, corrected by the lags thus calculated and expressed in half days, and the water level recorded at Austerlitz.

Once our model is in place, we will be able to carry out simulations, in order to obtain return periods of historical floods as well as a distribution of the gross annual losses in relation to the two modules carried out earlier. In order to carry out these simulations, we need to know the precipitation law of the catchment. For this purpose, we calibrated a matching law for rainfall conditional on the station and month studied, as the distribution will not be the same according to the geolocation and the period of the year considered.

However, rainfall, like all meteorological phenomena, is spatially and temporally dynamic. Thus, in order to model the dependence structure showing the probability of it raining in a given area if it is raining in another area, a Gaussian copula has been modeled. We now have all the tools to carry out our simulations, through continuous annual trajectories in twelve-hour steps.

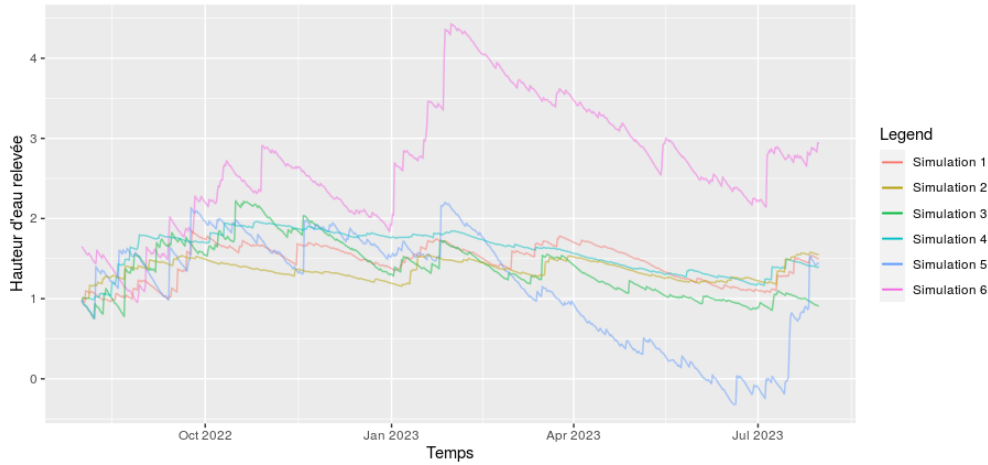


FIGURE 23 – Examples of annual simulations of the water level at Austerlitz.

Results

Thanks to these 3 modules, we were able to simulate 10,000 trajectories and calculate our return periods according to the intensity of a given flood, as well as to plot the distribution of our annual gross losses.

Water level (m)	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9
Return Period (y)	5	8	12	21	45	104	168	357	1250
Losses (M€)	31.5	35.2	39.7	47.8	63.4	74.5	84.3	101	126

TABLE 2 – Return periods and losses by water level.

These calculations also allow us to obtain the cartography of the Parisian basin associated with the different return periods and thus to build a geographical vision of the risk of flooding of the Seine.

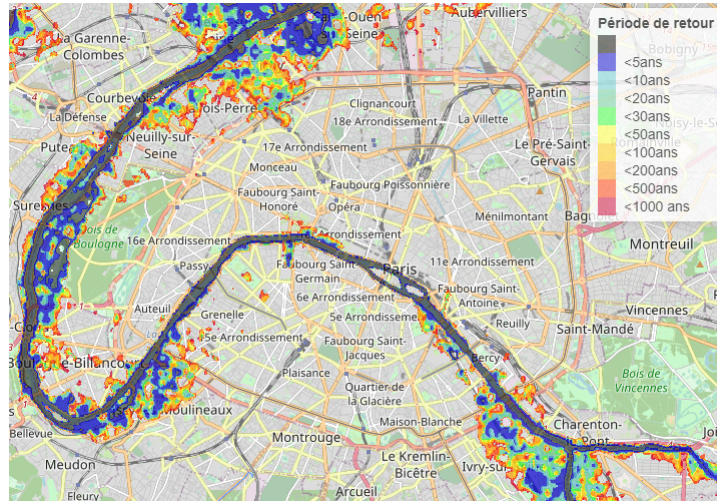


FIGURE 24 – Mapping of areas at risk with estimated return periods.

Externalization

In this unprecedented context of climate change, we sought to study the impact of an increase in precipitation, whether in frequency, intensity or both combined. Thus, through three scenarios that we have constructed based on the latest IPCC report, we have estimated the return periods of floods by modifying our precipitation laws.

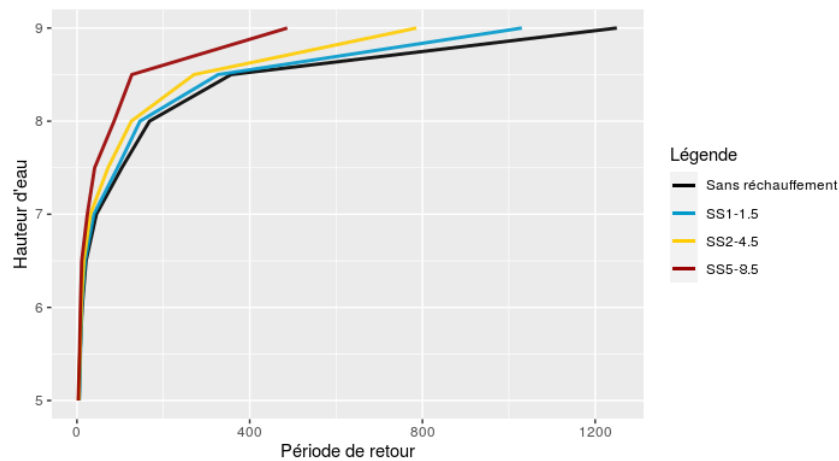


FIGURE 25 – Evolution of return periods according to the scenario considered.

We can clearly see an increase in the risk of flooding of the Seine due to global warming. We can thus plot the evolution of our vulnerability maps by return period in order to visualise the areas potentially at risk in the years to come.

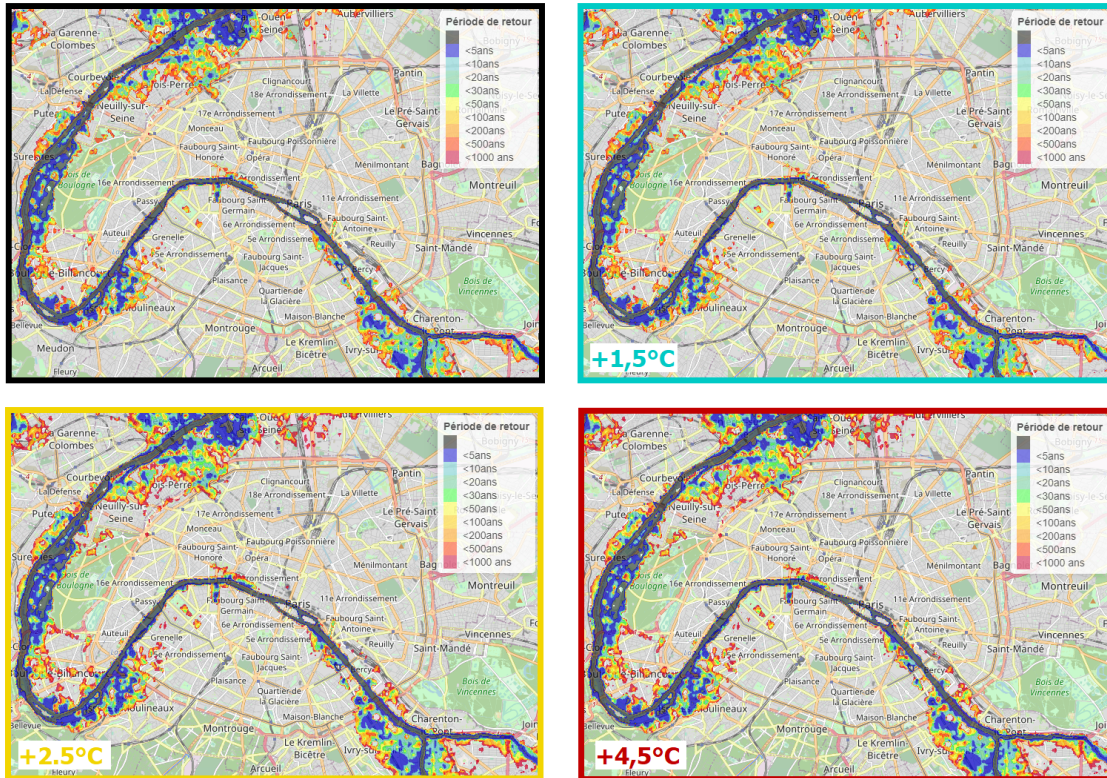


FIGURE 26 – Evolution of return periods according to the scenario considered.

Remerciements

Je tiens à remercier toutes les personnes ayant contribué de près de ou de loin à la réalisation de ce mémoire, représentant pour moi l'aboutissement de mes sept années d'études supérieures.

Je voudrais dans un premier temps remercier particulièrement mes responsables de mémoire en entreprise, Claire NICOLLE et Clément BILLORE pour leur extrême disponibilité et la pertinence de leur encadrement tout au long de mon stage.

Je tiens à exprimer toute ma reconnaissance envers mes directeurs en entreprise, Michael DONIO et Ronan DAVIT, ayant veillé au bon déroulement de mon stage.

J'adresse de la même manière mes sincères remerciements à mon ancien directeur INSA, James LEDOUX, qui m'a permis d'être là où je suis aujourd'hui et en qui j'ai toujours ressenti une confiance absolue. J'adresse également ma reconnaissance à Franck VERMET, directeur de l'EURIA, ayant également permis la réalisation de mon projet de double cursus.

Enfin, j'exprime toute ma gratitude à ma famille et à mes amis qui m'ont aidé dans la correction de ce mémoire et pour les nombreux échanges qui m'ont aidé à avancer.

Table des matières

Introduction	1
1 Le risque de crue de la Seine	3
1.1 Les catastrophes naturelles	3
1.1.1 Définition des CAT NAT	3
1.1.2 Le risque CAT NAT en France	5
1.1.3 Le régime de couvertures CAT NAT en France	7
1.2 Le risque inondation	8
1.2.1 Un risque majeur	8
1.2.2 Un risque d'actualité	11
1.2.3 Un risque mal appréhendé	12
1.3 Le bassin de la Seine	13
1.3.1 Géographie de la Seine	13
1.3.2 Hydrologie de la Seine	15
1.3.3 Historique des crues de la Seine	18
1.3.4 Aménagements modernes	21
1.4 Le bassin parisien	23
1.4.1 La Seine à Paris	23
1.4.2 Impacts sociaux-économiques d'une grande crue de la Seine	24
1.5 Modélisation d'une crue et de ses impacts	25
1.5.1 Démarche générale d'un modèle CAT NAT	25
1.5.2 Module exposition : Cartographie des risques	26
1.5.3 Module aléa : Calibrage du péril	28
1.5.4 Module vulnérabilité : Évaluation de l'impact	31
1.5.5 Module financier : Calcul des pertes	32
2 Cartographie du risque inondation en région parisienne	33
2.1 Topographie générale	36
2.1.1 Problématique	36
2.1.2 Base de données COPERNIC	37
2.2 Cartographie de la topographie parisienne	39
2.2.1 Méthodologie	39
2.2.2 Résultats	39

2.3	Cartographie de la Seine	41
2.3.1	Méthodologie	41
2.3.2	Résultats	41
2.3.3	Empreintes de crues	44
2.4	Cartographie des expositions	45
2.4.1	Scrapping de données	45
2.4.2	Analyse de la base de données contrats MRH	45
2.4.3	Mailles d'exposition	50
2.4.4	Cartographie du risque et taux de sinistralité	53
2.5	Évaluation de la Vulnérabilité	61
2.5.1	Méthodologie	61
2.5.2	Analyse de la base de données sinistres MRH	62
2.5.3	Calibrage des taux de destruction	64
2.6	Évaluation des pertes	71
2.6.1	Méthodologie	71
2.6.2	Représentation graphique de la fonction de perte	72
3	Modélisation de l'aléa et probabilisation des pertes brutes annuelles	75
3.1	Hydrométrie	77
3.1.1	Base de données HYDRO	77
3.1.2	Étude de la dynamique de la Seine	80
3.2	Pluviométrie	85
3.2.1	Base de données ISD	85
3.2.2	Analyse des données pluviométriques	87
3.2.3	Calibrage paramétrique	91
3.2.4	Mesures de dépendances	97
3.2.5	Copules multivariées	98
3.3	Mise en place d'un modèle explicatif	100
3.3.1	Méthodologie	100
3.3.2	Modèle linéaire (LM)	101
3.4	Mise en place des simulations de la hauteur d'eau	104
3.4.1	Simulation de la pluviométrie annuelle	104
3.4.2	Étude de l'hydrométrie d'Austerlitz	106
3.5	Étude des pertes brutes et des périodes de retour	107
3.5.1	Calcul des pertes brutes associées à chaque simulation	107
3.5.2	Évaluation de la distribution des pertes brutes annuelles	108
3.5.3	Évaluation des périodes de retour	110
3.5.4	Cartographie des périodes de retour	111
3.6	Scénarios d'application du changement climatique	112
3.6.1	Contexte	112
3.6.2	Intégration d'hypothèses GIEC	113
3.6.3	Résultats	115
	Conclusion	121

A Analyses sur le nombre de pièces du bien assuré	131
B Résumé schématique des étapes du module aléa	133
C Interpolation et agrégation des données ISD.	135
D Lois testées pour la simulation des précipitations	137
E Méthode du maximum de vraisemblance pour la loi Gamma	139
F Résultats d'adéquation des données ISD	141
G Test de Kolmogorov-Smirnov	153
H Théorie sur les copules	155
I Dépendances entre les stations	161
J Régression linéaire - Méthode des moindres carrés	169
Bibliographie	172

Introduction

Les aléas climatiques représentent un des risques les plus difficiles à appréhender pour les assureurs. Ils sont à la fois extrêmement coûteux car ils touchent un ensemble conséquent de polices en même temps mais aussi très compliqués à modéliser du fait de leur faible fréquence d'occurrence. Cependant, leur nombre ne cesse d'augmenter au fil du temps. Si l'on en comptait moins d'une centaine par année avant les années 1980, on en recense plusieurs centaines aujourd'hui, et leur coût est en constante progression. Il est donc capital pour les assureurs de se prémunir au mieux contre ces risques.

Aujourd'hui les assureurs utilisent des modèles de place dit "boîte noire" qui ne leur permettent pas de comprendre et maîtriser le risque catastrophe naturelle. Ces modèles proviennent pour la plupart de trois sociétés externes expertes en modélisation CAT NAT : AIR WorldWide (AIR), Risk Management Solutions (RMS) et EQECAT (EQE). La directive Solvabilité II a obligé les assureurs à quantifier les risques CAT NAT et a poussé les compagnies d'assurance à mieux appréhender ces risques. Le développement de modèles internes pour l'évaluation de la solvabilité et le pilotage nécessite d'aller encore plus loin dans la compréhension et la quantification de ces risques extrêmes et poussent les assureurs à se munir d'outils et de modèles leur permettant d'améliorer la compréhension de ces risques.

De plus, le changement climatique, avec la prise de conscience générale depuis notamment la COP21 à Paris (accord auquel 195 pays se sont engagés à réduire leurs émissions de CO₂ et limiter le réchauffement climatique à +1,5°C d'ici à 2050) voient les régulateurs nationaux dans le monde se pencher sur la question. L'ACPR a réalisé en 2020 un exercice pilote afin d'évaluer les impacts du réchauffement climatique sur le système bancaire et assurantiel en France [ACPR, 2021]. La Banque Centrale Européenne a également entrepris le lancement de travaux similaires pour l'année 2022. Les régulateurs poussent ainsi les institutions financières à se doter d'une gouvernance pour évaluer et maîtriser les risques émergents en conséquence directe du réchauffement climatique. Le rapport du GIEC paru cet été accentue d'autant plus cette nécessité de mieux appréhender les risques qui découlent d'un tel changement. Il nous servira d'appui également pour proposer et développer des scénarios d'applications dans ce mémoire.

Si beaucoup d'aléas climatiques sont sensibles à l'évolution du climat, c'est notam-

ment le cas de l'inondation. Premier risque CAT NAT en France, le risque inondation constitue plus de 50% de la charge totale des dommages CAT NAT depuis des années sur le sol français. Le ministère fait état de pas moins de 17 millions d'habitants et 22 500 communes situées en zones inondables et faisant face à ce risque permanent. L'association Missions Risques se projette même sur une hausse de 114% des charges totales liées aux inondations en France à horizon 2040.

Une crue de la Seine équivalente à celle de 1910 engendrerait des pertes colossales, estimées à 20 milliards d'euros au minimum. L'objectif de ce mémoire est de répondre à la problématique de la modélisation du risque inondation en développant un modèle permettant d'appréhender le risque de crue de la Seine sur un portefeuille parisien. Pour ce faire, il est nécessaire de réaliser une étude des mécanismes régissant la hauteur d'eau d'un fleuve et de s'approprier les phénomènes naturels inhérents à l'apparition de crues.

Nous allons donc dans un premier temps étudier toutes les caractéristiques de la Seine et de son bassin versant. Ensuite, nous procéderons à une cartographie du cours d'eau ainsi que de la topographie du bassin parisien qui nous permettra de relier le phénomène de crue avec nos expositions sous risque. Par la suite, nous analyserons une base de sinistres inondations pour calibrer des taux de destruction moyens. Ensuite, nous nous attacherons à étudier et à modéliser les crues de la Seine au travers de la pluviométrie du bassin, principal facteur explicatif de l'apparition de crues. Ceci nous permettra de calculer des périodes de retour, et de vérifier des assertions pas toujours scientifiquement fondées. La réunification de tous ces travaux nous permettra finalement de calculer les pertes brutes associées à notre portefeuille initial. Enfin nous proposerons trois scénarios issus des prévisions du rapport du GIEC et mesurerons les impacts de ces évolutions sur les périodes de retour des crues de la Seine.

Chapitre 1

Le risque de crue de la Seine

1.1 Les catastrophes naturelles

1.1.1 Définition des CAT NAT

Une catastrophe naturelle est un évènement intense d'origine naturelle générant de nombreux dégâts, aussi bien matériels qu'humains. On en recense plusieurs aux origines et aux fréquences d'apparition diverses. Le tableau ci-dessous propose un aperçu de ces différentes catastrophes naturelles, classées par typologie :

Type de catastrophe	Catastrophes
Biologique	Epidémie Invasion d'insectes
Climatique	Températures extrêmes Feux de végétation Sécheresse
Géologique	Séismes Glissement de terrain sec Eruption volcanique
Hydrologique	Glissement de terrain humide Inondation
Météorologique	Tempête

TABLE 1.1 – Répertoire des catastrophes naturelles

Certains de ces évènements sont plutôt rares comme les tsunamis par exemple, lorsque d'autres sont bien plus fréquents. Parmi les catastrophes les plus communes, on peut retrouver les inondations et les tempêtes, qui à elles seules, représentent près de 60% des catastrophes mondiales entre 1990 et 2007. A noter que les chiffres présentés ci-dessous sont établis en termes de fréquence de survenance.

Le bilan d'une catastrophe naturelle est bien souvent désastreux et possède de nom-

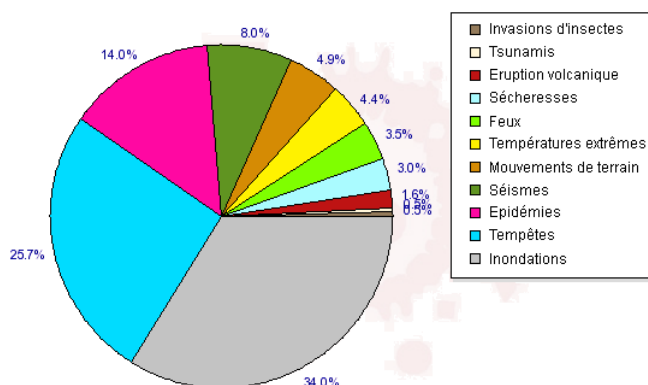


FIGURE 1.1 – Répartition des catastrophes naturelles par typologie (1990 à 2007). *Source : EM-DAT - The Emergency Events Database*

breux enjeux : humains, sociaux, économiques, environnementaux... A titre d'exemple, l'année 2011 qui a connu notamment le tsunami de Honshu au Japon a engendré plus de 360 milliards de dollars de pertes économiques. Cependant, le plus alarmant n'est pas tant les sommes astronomiques causées par ces événements que l'évolution de la fréquence de celles-ci au cours du temps. Depuis presque un siècle, elles ne cessent d'augmenter et coûtent ainsi de plus en plus cher.

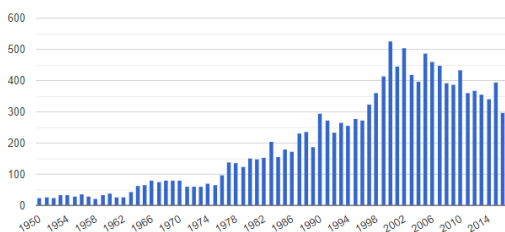


FIGURE 1.2 – Evolution du nombre de catastrophes naturelles depuis 1950. *Source : EM-DAT - The Emergency Events Database*

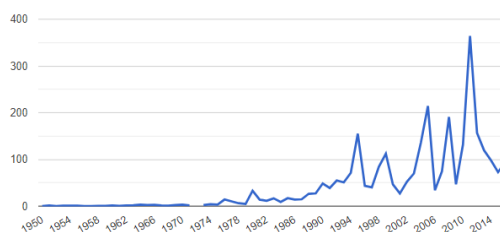


FIGURE 1.3 – Evolution du coût des catastrophes naturelles depuis 1950. *Source : EM-DAT - The Emergency Events Database*

Ainsi, de nombreux experts estiment que le réchauffement climatique a déjà contribué à augmenter la fréquence d'apparition et l'intensité des catastrophes naturelles et que cela n'ira pas en s'arrangeant dans les prochaines décennies. C'est pour cela que les catastrophes naturelles ainsi que le réchauffement climatique sont étroitement liés et conduisent bon nombre de régulateurs à s'y pencher de plus en plus sérieusement au fil des années.

1.1.2 Le risque CAT NAT en France

La France est un pays qui n'échappe pas à la problématique des catastrophes naturelles, bien au contraire. Elle est soumise à l'ensemble des aléas climatiques existants, ce qui en fait l'un des pays européens les plus exposés au risque CAT NAT. Si l'on se réfère à l'observatoire francophone des risques naturels *Catnat.net*, la France possède un indice global de risque "fort" à court terme quant aux risques naturels. Ce classement est réalisé sous l'analyse de cinq indicateurs de risque : Exposition, Gravité humaine, Gravité matérielle, Vulnérabilité, Résilience. Leur agrégation permet alors de définir un seuil de risque global mettant en relation l'exposition générale d'un pays à sa capacité de réponse en cas de catastrophe naturelle.

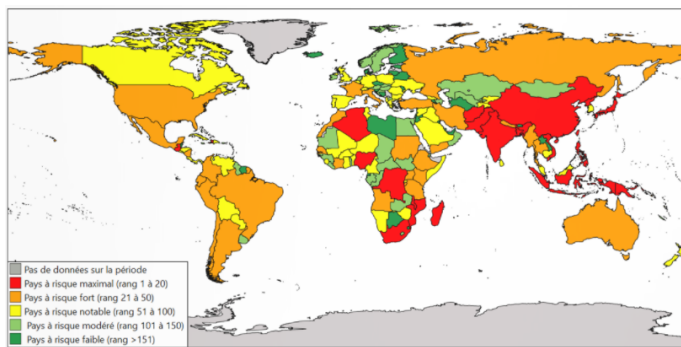


FIGURE 1.4 – Indice de risque CAT NAT par pays. *Source : catnat.net*

On observe bien ici les risques encourus par la France. Cependant, à la différence de la vue mondiale, la France possède un profil de risque propre à sa géographie. Ainsi, elle est moins exposée aux tempêtes mais plus exposée aux inondations ou aux sécheresses que ce que l'on peut constater sur le reste du globe. La Caisse centrale de réassurance (CCR) fournit des chiffres quant à la sinistralité CAT NAT sur le sol français entre 1982 et 2020 :

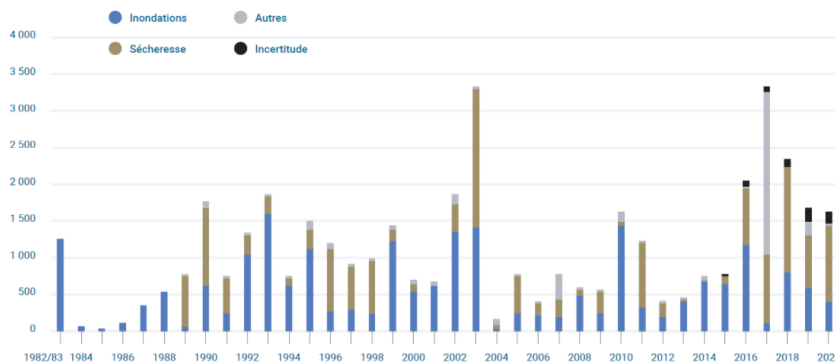


FIGURE 1.5 – Evolution du coût des CAT NAT en France depuis 1982. *Source : CCR*

Cela nous donne alors un aperçu plus précis des risques encourus dans l'hexagone et les DOM TOM¹. A noter que les données présentées par la CCR sont évaluées en pourcentage de la charge totale, ce qui n'est donc pas immédiatement comparable avec les chiffres d'EM-DAT. On peut observer ci-dessous avec la vue agrégée que depuis les années 80, la France s'acquitte en majorité de dégâts causés par les inondations sur son territoire.



FIGURE 1.6 – Répartition des CAT NAT en France. *Source : CCR*

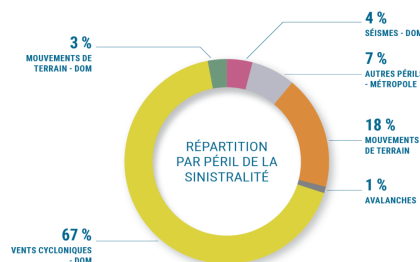


FIGURE 1.7 – Répartition des CAT NAT en France (AUTRES). *Source : CCR*

Si la France est un pays à risque, elle l'est sur l'ensemble de son territoire. Aujourd'hui, bon nombre de communes sont exposées à un ou plusieurs risques naturels. D'après le rapport public de janvier 2020 du ministère de la transition écologique sur la vulnérabilité de la France face aux risques climatiques, on recense par exemple pas moins de :

- 64% des communes touchées par le risque inondation ;
- 55% des communes touchées par le risque de mouvements de terrain ;
- 19% des communes touchées par le risque de feux de forêt.

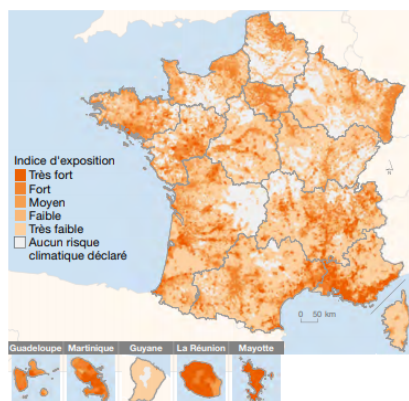


FIGURE 1.8 – Exposition de la population française aux aléas climatiques (2016). *Sources : MTES, 2017 ; Insee, 2014 ; IGN, 2016. Traitements : SDES, 2019.*

1. L'analyse porte sur l'ensemble du marché français (Métropole et Outre-mer). La sinistralité des exercices récents n'est pas totalement consolidée, des estimations CCR sont donc fournis. Avant 1989, la sinistralité par péril n'est pas ventilée mais est à l'époque principalement causée par les inondations.

Pire encore, certaines communes connaissent un agrégat de risques simultanés. On comptabilise 36% de ces communes qui cumulent deux aléas, 10% qui en cumulent trois et même 1% cumulant les quatre. La carte 2.8 montre ainsi la ventilation géographique de ces risques. Le rapport finit par conclure que pas moins de six français sur dix seraient déjà concernés par un ou plusieurs aléas climatiques en 2016 [MTES, 2020]. Et ce chiffre pourrait augmenter dans les années à venir.

1.1.3 Le régime de couvertures CAT NAT en France

Il existe en France un régime d'indemnisation couvrant la quasi-totalité des catastrophes naturelles. Il a été créé à la suite de l'inondation de la Saône en décembre 1981, causant des coupures de chauffage pendant plus de huit jours à Mâcon notamment, ce qui poussa le parlement à légiférer pour pallier l'absence criante de couverture contre ces événements exceptionnels. Six mois plus tard était mis en place le régime CAT NAT français. Désormais, tout souscripteur d'un contrat d'assurance de dommages aux biens bénéficie alors d'une couverture automatique contre les risques de catastrophes naturelles. Selon l'article L125-1 du Code des Assurances, cette garantie prend en charge les "dommages matériels directs non assurables ayant eu pour cause déterminante l'intensité anormale d'un agent naturel, lorsque les mesures habituelles à prendre pour prévenir ces dommages n'ont pu empêcher leur survenance ou n'ont pu être prises.". C'est l'État qui garantit la viabilité de ce système en reconnaissant un événement comme catastrophe naturelle. Sur la base de rapports scientifiques, des arrêtés ministériels sont publiés sur le journal officiel à la suite de requêtes déposées par le préfet sur demande des communes touchées. Une commission interministérielle est alors chargée de définir le cadre de l'évènement : période de la catastrophe, nature des dommages couverts par la garantie et communes concernées par l'évènement (article L.125-1 du Code des Assurances). Le processus d'indemnisation des assurés peut finalement être assez long.

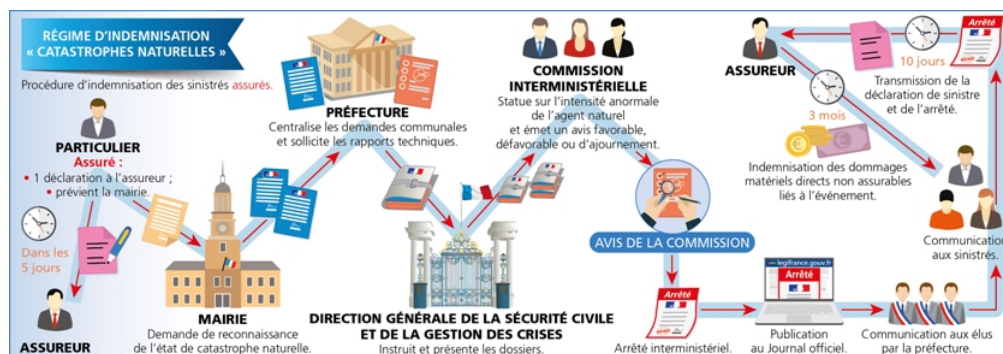


FIGURE 1.9 – Dispositif d'indemnisation des assurés. *Source : Ministère de l'intérieur*

La garantie CAT NAT couvre des dommages matériels directs causés uniquement aux biens assurés lorsque ceux-ci sont couverts dans le contrat d'assurance type MRH (Multirisques Habitation). Les dommages corporels ne sont pas couverts, tout comme les dommages indirects. En contrepartie de cette couverture, les assurés paient une surprime

définie par l'État et géographiquement indifférenciée (quel que soit l'exposition au risque et le type de péril contre lequel l'assuré se prémunit). Elle est fixée depuis 2009 à 12% de la prime afférente aux garanties dommages pour les biens hors véhicules à moteur, et 6% des primes vol et incendie (ou à défaut, 0,50% de la prime dommage) pour les véhicules terrestres à moteur.

L'ampleur des dégâts provoqués par de tels événements oblige les sociétés d'assurance à partager leurs risques. C'est ainsi qu'interviennent les réassureurs. Dans le cas du régime CAT NAT, l'État garantit une réassurance illimitée à la CCR (Caisse Centrale de Réassurance), SA détenue par l'État, qui la propose elle-même aux sociétés d'assurance. Chaque compagnie d'assurance prend alors en charge les indemnités jusqu'à un certain seuil au-delà duquel la CCR intervient et si nécessaire, la garantie de l'État. Une franchise est définie assurant un principe de responsabilité de l'assuré. Elles sont définies depuis 2001 tel que suit (franchise générale) :

- 380 € pour les biens à usage d'habitation et les véhicules automobiles ;
- 10% des dommages avec un minimum de 1 140 € pour les biens à usage professionnel ;
- 3 jours ouvrés avec un minimum de 1 140 € pour les pertes d'exploitation.

L'articulation générale de l'indemnisation des dégâts CAT NAT est résumée ci-dessous :

Chronologie	Indemnisation
1.	Franchise retenue par l'assuré
2.	Part à la charge de l'assureur
3.	Partie endossée par la CCR
4.	Intervention de la garantie de l'État

TABLE 1.2 – Articulation du régime d'indemnisation CAT NAT en France.

1.2 Le risque inondation

1.2.1 Un risque majeur

Le risque inondation est comme beaucoup de CAT NAT un risque difficile à dompter. Les pertes liées à ces événements n'ont cessé d'augmenter au fil du temps. La France, premier réseau fluvial d'Europe n'échappe pas à ce risque bien au contraire. Comme on a pu le voir plus tôt, les inondations représentent la principale source de risque de catastrophe naturelle pour le pays. En effet, il concerne une partie importante de la population et demeure le risque naturel qui provoque le plus de dégâts matériels en métropole. Ainsi, on recense selon le Ministère en charge de l'écologie 17,1 millions d'habitants et 9 millions d'emplois qui se trouvent dans des zones inondables par débordement de cours d'eau, et 1,4 millions d'habitants et 850 000 emplois qui sont exposés aux submersions marines [MTES, 2017a]. Car en effet, il existe des inondations de différentes sortes et représentant des amplitudes de risque variées.

Les inondations par ruissellement

Définition : L'inondation par ruissellement se produit lorsque les eaux de pluie ne peuvent pas ou plus s'infiltrer dans le sol.

Causes : A l'origine de ce phénomène, ce sont de fortes pluies s'abattant sur une courte période empêchant l'eau de s'infiltrer assez vite ou bien un important cumul de pluie sur plusieurs jours.

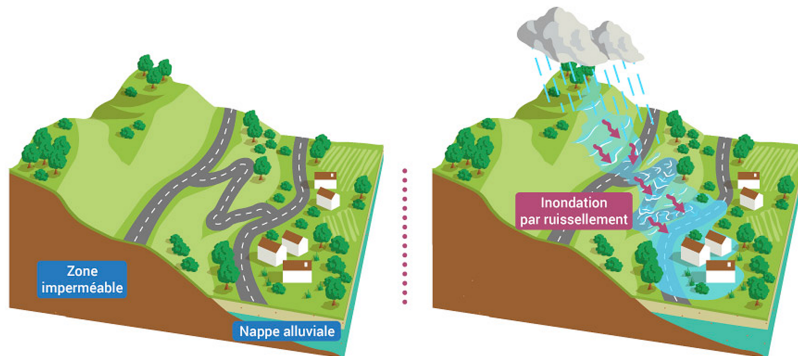


FIGURE 1.10 – Inondation par ruissellement. *Source : Hydro Eau France*

Les inondations par submersion marine

Définition : Les submersions marines sont des inondations de la zone côtière par la mer lors de conditions météorologiques et océaniques défavorables. Elles menacent alors tout le littoral.

Causes : La montée des eaux ainsi que la houle causées notamment par les marées et tempêtes peuvent conduire la mer à submerger un territoire côtier.

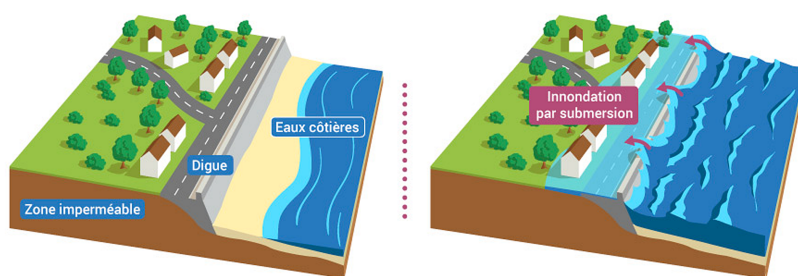


FIGURE 1.11 – Inondation par submersion marine. *Source : Hydro Eau France*

Les inondations par débordement de cours d'eau

Définition : L'inondation par débordement de cours d'eau se produit lorsqu'un cours d'eau déborde de son lit habituel. On peut distinguer dans ses débordements les crues lentes de plaine des crues rapides et torrentielles.

Causes : Les causes sont là encore les fortes pluies intenses par leur durée ou par leur débit.

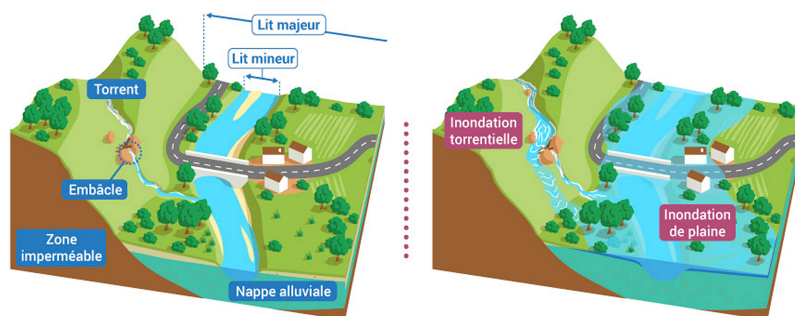


FIGURE 1.12 – Inondation par débordement de cours d'eau. *Source : Hydro Eau France*

Les inondations par remontée de nappe

Définition : L'inondation par remontée de nappes est provoquée lorsque le niveau de la nappe phréatique atteint la surface du sol.

Causes : Lors d'intenses épisodes pluvieux, les nappes phréatiques se remplissent. Toute pluie tombant sur une zone alors saturée va provoquer la stagnation ou l'écoulement de l'eau vers la vallée la plus proche. Cela peut alors même inonder des zones éloignées d'un cours d'eau.

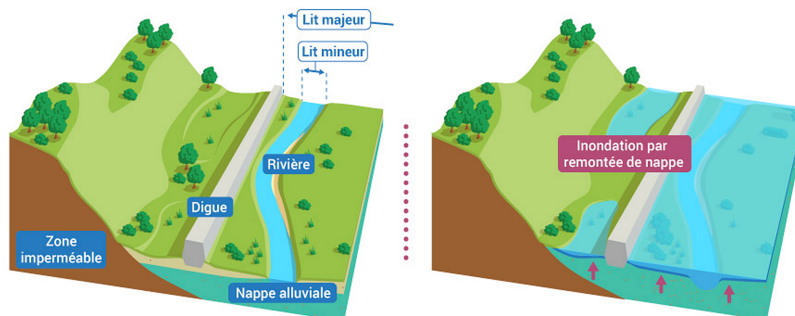


FIGURE 1.13 – Inondation par remontée de nappes. *Source : Hydro Eau France*

Ainsi, pour caractériser une inondation en plus de sa typologie, on définit trois critères qui vont alors témoigner de l'intensité de la catastrophe :

- **Intensité** : Pour une crue, elle se mesure par la hauteur d'eau relevée sur une zone donnée. L'intensité est évolutive, elle sera à des degrés différents au fil de l'épisode de crue ;
- **Durée** : Elle représente la persistance dans la durée de l'épisode de crue. Cela peut aller de quelques heures à plusieurs semaines et va fortement modifier l'ampleur des dégâts ;
- **Périodicité** : On la mesure au travers de la période de retour. A une hauteur d'eau donnée, ou une perte donnée on associe la probabilité annuelle que l'évènement se produise. Ainsi, une période de retour de 100 ans signifie qu'on a 1 chance sur 100 chaque année d'observer un évènement de ce type.

L'intensité et la durée de la crue permettront de calculer les dégâts causés par une crue typique. Dans le cadre de la réglementation Solvabilité II, il s'agira d'évaluer l'impact d'une crue bicentennale (1 chance sur 200) afin d'assurer la couverture des assureurs contre cet évènement rare.

1.2.2 Un risque d'actualité

De façon contraire à la logique et malgré la conscienciosité générale par rapport à l'évolution croissante du risque inondation, l'Homme continue de construire et de s'installer près des rivières et des fleuves. En effet, depuis des décennies la valeur des biens construits dans des zones inondables augmentent, ce qui rend les crues plus coûteuses qu'auparavant. Effectivement, selon le SOeS (Service de l'Observation et des Statistiques), près de 100 000 logements supplémentaires ont été construits sur des zones potentiellement inondables entre 1999 et 2006. En outre, le rapport du ministère de la transition écologique et solidaire datant d'octobre 2017 chiffre à 25% l'habitat existant se situant déjà en zone inondable. Et malgré la conscience collective et l'inflation du risque inondation, il n'est pourtant pas envisagé de délocaliser ces zones sur-urbanisées. La métropolisation et la pression foncière conduit déjà et conduira encore plus à l'avenir à densifier l'habitat le long des voies d'eau, là où sont implantés les grandes métropoles. Le bassin parisien en est le parfait exemple. Une crue du type de celle de 1910 aujourd'hui coûterait bien plus qu'à l'époque du fait de l'augmentation de cette vulnérabilité. On peut prendre comme exemple le secteur du transport. En 1910, les locomotives au charbon continuaient à rouler même avec plusieurs centimètres d'eau, là où aujourd'hui le réseau RATP se retrouverait paralysé du fait de son fonctionnement à l'électronique. Tout ceci augmente la vulnérabilité de ces territoires face aux inondations.

Cependant, ce n'est pas l'unique module qui connaît une évolution relativement défavorable. En effet, la probabilité d'apparition d'épisodes de crues intenses est aussi sensiblement en train d'augmenter. Bien qu'à ce jour les scientifiques ne soient pas tous d'accord quant à la progression de la quantité de pluie tombant sur le pays au fur et à mesure que le réchauffement climatique s'accélère, bon nombre s'accordent à dire que la fréquence et la puissance d'épisodes pluvieux courts mais intenses tendent à augmenter de façon certaine avec le temps. Et c'est typiquement ce genre de phénomène qui peut entraîner des crues dévastatrices. L'élévation du niveau de la mer induit par ce changement climatique ne vient pas aider dans l'équation, les fleuves ayant ainsi plus de mal à

écouler leurs eaux dans la mer.

En résumé, la combinaison de tous ces facteurs aggravants tels que cités précédemment menace le globe :

- **Constructions en zones inondables** : augmentation de la vulnérabilité et des dégâts ;
- **Artificialisation des sols** : diminution de la retenue d'eau par la végétation et les sols causant une aggravation des crues ;
- **Changement climatique** : épisodes pluvieux plus fréquents et plus intenses.

Ainsi, la MRN (Mission Risques Naturels) estime à +114% la hausse du coût cumulé des inondations pour la période 2014-2039. Pire encore, selon le centre scientifique de l'Union Européenne, le coût annuel des dommages provoqués dans l'UE par les inondations uniquement côtières pourrait déjà atteindre 1 000 000 000 000 \$ en 2100. L'appréhension de ce risque est alors capital afin de se prémunir contre un ou plusieurs événements extrêmes qui pourraient survenir plus vite que prévu.



FIGURE 1.14 – Carte des zones inondables. *Source : Ministère de l'écologie*

1.2.3 Un risque mal appréhendé

Depuis des décennies, des responsables de différents secteurs se sont penchés sur la problématique des crues. Qu'ils soient gestionnaires, politiques ou météorologues, d'innombrables études et rapports ont été rédigés par leurs soins dans le but de mieux comprendre et de mieux appréhender ce risque. Différentes issues ont été dégagées sans

pour autant en faire des solutions miracles. Recours à la nature, création de bassins anti-crues, réduction de la vulnérabilité des constructions situées en zone inondable (BIGOT et GALIBERT, 2017) sont des pistes qui ont été évoquées.

Outre le secteur de la gestion publique, le monde de l'assurance est un des premiers impacté par cette inquiétante croissance du risque de catastrophe naturelle et notamment du risque inondation. À ce jour, le marché de la modélisation des CAT NAT est très réduit. En effet, il est aujourd'hui composé d'un quasi-monopole (environ 90% des parts du marché) des trois sociétés phares en la matière : AIR, RMS et EQE. Cependant, ces dernières années plusieurs acteurs ont souligné leur volonté de s'affranchir de leur dépendance à ces structures. A cela s'ajoute la remise en cause grandissante de la fiabilité des modèles "boîte noire" de ces sociétés impossible donc à contrôler. C'est pourquoi de nombreux acteurs du marché de l'assurance tentent de développer leur propre modèle interne. Ils cherchent alors à créer leur outil afin d'assurer la mise en place et la compréhension des paramètres sous-jacent. C'est ainsi qu'est né le système PERILS, créé en 2009 et regroupant les principaux organismes d'assurance européens Allianz, Axa, Groupama, Generali ainsi que les réassureurs Munich Re, Swiss Re et Partners Re. Le but était de constituer une base de données d'exposition au risque afin d'estimer les pertes engendrées par les tempêtes en Europe et les inondations au Royaume-Uni. Plus récemment, le projet européen OASIS Loss Modelling Software (LMF) regroupant une gamme plus large d'acteurs (Assureurs, Banques, Universités, ...) a été mis en place en 2014. L'objectif de ce système est plus ambitieux qu'un simple outil d'estimation des pertes comme celui déployé par PERILS. Plateforme communautaire fondée sur la logique de l'Open Source, elle a été développée dans le but de tester un grand nombre de modèles sous le contexte Solvabilité II. En parallèle de ces projets collaboratifs et dans une logique de course à l'innovation, nombreux sont les assureurs ayant lancé le développement de leur propre modèle interne de catastrophes naturelles adapté à leur portefeuille.

C'est précisément dans ce cadre que s'inscrit ce mémoire. L'objectif est d'appréhender au mieux le risque de crue de la Seine sur le bassin parisien qui constitue probablement le risque majeur d'inondation en France étant donné la densité de population et d'infrastructures sous risque dans la capitale. Pour ce faire, il est d'abord crucial de comprendre les mécanismes qui régissent les crues du bassin de la Seine, et de s'appropriier les différentes données caractéristiques du fleuve.

1.3 Le bassin de la Seine

1.3.1 Géographie de la Seine

Troisième plus grand bassin versant de France après ceux de la Loire et du Rhône, le bassin de la Seine s'étend du massif du Morvan aux côtes havraises où il se jette, dans le département de la Manche. Au sein des régions Ile-de-France et Normandie, le faible dénivelé de la Seine a participé à la formation de nombreux et profonds cours d'eaux sinueux. Ces méandres sont parties prenantes du fleuve car leur déversement dans la Seine participent grandement à son écoulement.



FIGURE 1.15 – Bassin de la Seine. *Source : GIP Seine-Aval.*

Le fleuve est alors composé d'une quarantaine de cours d'eau de plus de 50km. Les plus importants sont la Marne, l'Oise et l'Yonne. Le fleuve s'étend sur 775 km avec une source au plateau de Langres d'altitude 446 m, ce qui lui donne une déclivité relativement faible d'à peine plus de 50 cm par kilomètre. La conséquence est que certaines marées peuvent se faire ressentir sur une centaine de kilomètres en amont de son embouchure.

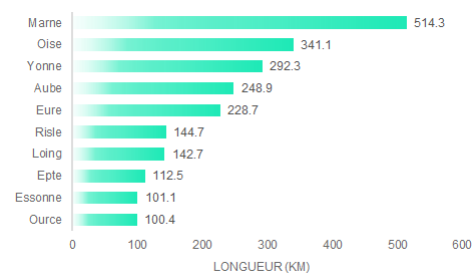


FIGURE 1.16 – Principaux affluents de la Seine.

Presque la totalité du bassin versant de la Seine se situe en France excepté les sources et les premiers kilomètres de l'Oise, de l'Artoise ainsi que le ruisseau des Anorelles en Belgique. Les métropoles les plus connues traversées par le fleuve sont Paris, Reims, Troyes, Auxerre, Le Havre et Rouen.

En résumé, le bassin de la Seine c'est :

- **76 238 km²** de superficie ;
- **17 millions** d’habitants ;
- **225 h/km²** de densité moyenne ;
- **6 700** communes traversées.

1.3.2 Hydrologie de la Seine

Débit

Le faible dénivelé de la Seine lui donne un écoulement que l’on peut qualifier de lent en comparaison à d’autres cours d’eau. En effet, il est selon *Eau France* le cinquième fleuve au débit le plus faible sur les six principaux fleuves français. Son débit moyen sur l’ensemble du bassin est d’environ 480m³/s toujours selon *Eau France*, avec des valeurs variables selon la région (563m³/s au Havre, 485m³/s à Poses, 310m³/s à Paris, 173m³/s à Montereau-Fault-Yonne, 77m³/s à Pont-sur-Seine). Ces chiffres sont cependant à prendre avec précaution car ce sont des moyennes annuelles ne prenant pas en compte les variations saisonnières. En effet, l’été le débit se situe généralement aux alentours de 100m³/s à Paris lorsque l’hiver il atteint généralement les 600m³/s.

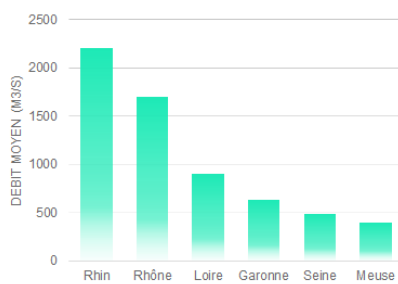


FIGURE 1.17 – Débits moyens des fleuves français.

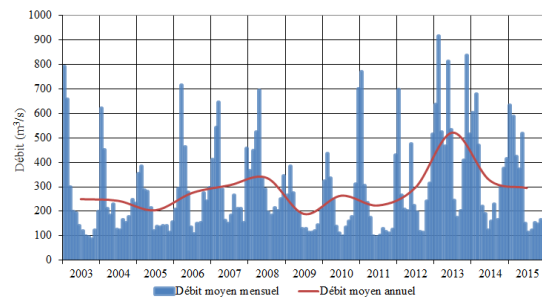


FIGURE 1.18 – Débits moyens annuels à Austerlitz. *Source : DRIEE*

Hauteur

La vitesse de montée ou descente du niveau de la Seine est elle aussi plus lente que certains cours d’eau. En effet, depuis les 130 dernières années, les augmentations du niveau de la Seine de plus de 50 cm en 24 heures se font rares, ce qui fait des crues de la Seine des crues principalement lente et prolongées dans le temps, à la différence des crues rapides connaissant une montée des eaux soudaine en l’espace de quelques heures seulement. A titre d’exemple, la Vienne (affluent de la Loire) a connu une montée des eaux de quatre mètres en 24 heures ce 21 juillet 2021, se rapprochant ainsi de son plus haut niveau historique.

La hauteur d’eau est évaluée par la banque *HYDRO*, directement dépendante du ministère de l’écologie, au travers de stations de mesure disposées sur le réseau hydrologique

français. Comme précisé plus tôt, les crues de la Seine sont définies par cette hauteur d'eau pouvant atteindre des sommets comme en 1910. A la station d'Austerlitz, l'échelle affichait à l'époque 8,10 mètres soit près de 7 mètres au-dessus de son niveau habituel (qui est à 1,5 mètres environ. Cette hauteur est calculée à partir du zéro de l'échelle de la station, se situant 25,92 mètres au-dessus du niveau de la mer.

A l'inverse, en période de sécheresse extrême, la hauteur d'eau peut atteindre un niveau dangereusement bas comme ce fut le cas en 1719, la Seine atteignant son plus bas niveau historique : 26,25 mètres au-dessus du niveau de la mer correspondant au zéro de l'échelle hydrométrique du pont de la Tournelle autrefois utilisé pour mesurer la crue de la Seine. Cela correspond à une hauteur de 33 centimètres sur l'échelle de son homologue Austerlitz.

Pluviométrie

Le climat du bassin de la Seine est océanique. L'humidité du bassin est véhiculée de façon constante par les vents dominants de l'ouest offrant une pluviométrie bien répartie au cours de l'année. La Seine connaît cependant des crues sévères majoritairement l'hiver. Ces crues font suite à des pluies sévères et continues présentes dans les parties amont du bassin. Bien que combiné à la piézométrie des sols, la végétation, l'urbanisation des sols et quelques autres facteurs, il est aujourd'hui connu que la pluie est le paramètre préalable à l'apparition de crues.

La pluviométrie moyenne sur l'ensemble du bassin de la Seine est de 800 mm/an. C'est l'équivalent d'environ quatre baignoires qui se déversent chaque année sur un mètre carré de la région. Cependant, tous les territoires du bassin ne sont pas soumis aux mêmes quantités de pluies annuelles. Le centre du bassin ne reçoit que 550 mm/an en moyenne là où la quantité d'eau grimpe à 1 200 mm/an pour la partie est du bassin. On peut également apercevoir cette différence sur la carte ci-dessous au travers d'une vision alternative : le nombre de jours de pluie sur l'année.

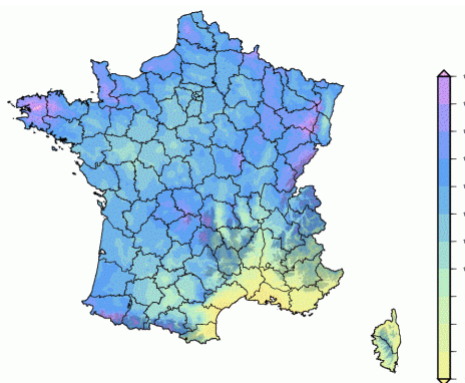


FIGURE 1.19 – Pluviométrie en France.² Source : Météo France

2. Moyenne annuelle de jours avec au moins 1 mm de précipitation.

Si l'on s'intéresse de plus près à la pluviométrie mensuelle observée sur le bassin parisien, on remarquera qu'elle est assez variable et qu'on peut dégager une composante saisonnière. Cependant, si l'on fait varier la profondeur d'historique des relevés, on obtient des résultats partiellement différents. On pourrait alors vouloir prendre la profondeur d'historique la plus grande possible afin d'éviter le problème de stabilité des données, mais la question du changement climatique pose problème. On peut se demander si considérer une pluviométrie vieille d'un siècle demeure pertinent ou non.

En observant la pluviométrie mensuelle historique de Paris sur les 20 dernières années, on constate que les mois les plus intensément pluvieux sont ceux de décembre et de mai. Mais alors comment expliquer que les plus grandes crues historiques de la Seine soient régulièrement survenues durant la période hivernale et pas en mai ou même au printemps ? Cela n'est pas dû au hasard, une des raisons est que la période hivernale est une période de pluies prolongées (novembre, décembre puis janvier) là où les pluies de mai sont plus ponctuelles. De plus, il faut aussi considérer la pluviométrie de l'ensemble du bassin. Car les crues de la Seine ne sont pas expliquées par la seule pluviométrie de Paris. Au regard des historiques des principales villes affluentes, novembre, décembre et janvier se classent alors presque systématiquement parmi les six mois les plus pluvieux de l'année sur l'ensemble du bassin.

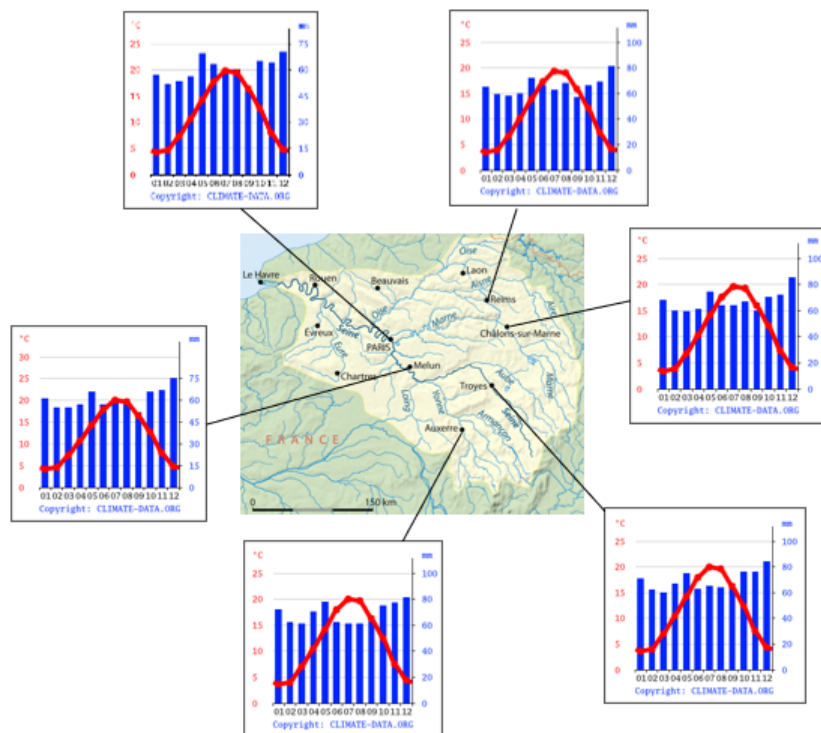


FIGURE 1.20 – Pluviométries mensuelles sur le bassin de la Seine.⁴ *Source : Climate data*

4. Moyenne mensuelles calculées entre 1999 et 2019. Base de données ECMWF.

1.3.3 Historique des crues de la Seine

Des siècles de crues

Les plus anciennes crues de la Seine identifiées furent en 358 et en 582. Depuis la fin du XIX^{ème} siècle, une échelle hydrométrique installée au pont d'Austerlitz à Paris mesure les crues de la Seine dans la capitale. Le record fut atteint en 1910. En effet, en janvier la Seine atteint une hauteur de 8,68 mètres. Cette crue a été qualifiée à l'époque de crue "centennale". Par ailleurs, au XX^{ème} siècle on recense cinq grandes crues d'une hauteur supérieure à 6,50 mètres : 1910, 1920, 1924, 1945 et 1955. En 2016, la Seine connaît une nouvelle crue importante puisque le niveau atteint 6,10 mètres. Cette crue fut la plus forte à Paris depuis plus de 30 ans. En 2018, une nouvelle crue importante fait son apparition atteignant la barre des 5,84 mètres de hauteur. Voici un condensé des crues les plus importantes recensées depuis le XVII^{ème} siècle :

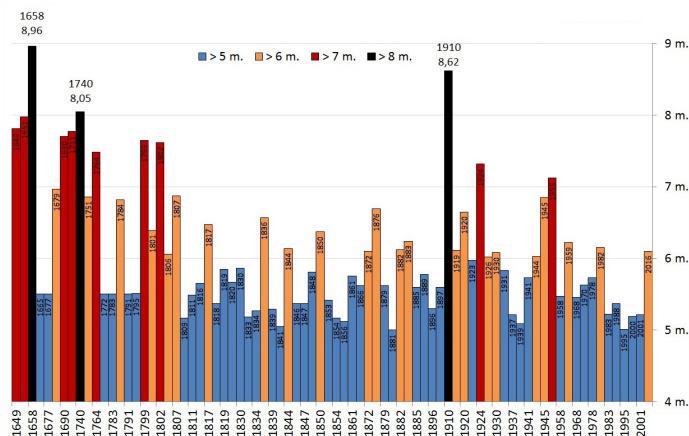


FIGURE 1.21 – Historique des crues de la Seine (1649-2016).⁶ Source : Oliver BER-RUYER.

Comme expliqué ci-dessus, une échelle hydrométrique n'a été installée qu'à la fin du XIX^{ème} siècle, c'est pourquoi les relevés antérieurs à 1885 sont considérés seulement comme "relativement fiables" jusqu'en 1800, "assez fiables" jusqu'en 1719 et seulement "indicatifs" depuis 1649. En dehors des trois plus grandes crues de la période, on recense neuf crues de 7 à 8 mètres de hauteur qu'on pourrait qualifier de "bi-centennales" et 23 crues de six à 7 mètres au caractère plutôt "vingtennal". Les crues de cinq à 6 mètres sont plus fréquentes. Nous allons retracer l'histoire de deux crues particulières : celle de 1910 et celle de 2016. La première car elle fait office de référence dans tous les papiers et les journaux. C'est en effet la plus importante de ces trois derniers siècles et la seule de ce genre qui est apparue assez tard pour nous permettre de disposer de documentation et de connaissances sur son apparition. La deuxième est quant à elle la plus grande crue de

6. Crues de la Seine supérieures à 5 mètres à l'échelle d'Austerlitz. Données lacunaires pour les crues de 5 mètres à 6 mètres avant 1800. Données provenant de la banque HYDRO.

l'histoire récente et aussi la plus importante depuis 1982. On aurait également pu citer celle de janvier 2018, mais qui fut légèrement moins violente que sa prédécesseure.

Crue de 1910

La crue de 1910 fut un cauchemar pour tous les parisiens de l'époque. Elle est restée dans les mémoires et demeure encore aujourd'hui la crue de référence que l'on appréhende revoir dans la capitale. Cependant, son caractère centennal et toutes les études s'accordent à dire que nous risquons d'observer une telle catastrophe dans les années à venir. Néanmoins, son caractère centennal a été prodigué à l'époque sans réels fondements scientifiques, et nombreux sont ceux à douter de cette théorie. Cette étude permettra notamment de challenger cette assertion.



FIGURE 1.22 – Photos d'époque lors de la crue de 1910.

Janvier 1910, Paris est sous l'eau. Mais comment est apparue cette crue dévastatrice dans la capitale ? Selon Météo France, cette crue s'inscrit dans un contexte météorologique exceptionnel. Condition sinequanone à la montée des eaux, la pluie est évidemment la première responsable. Cependant, si on regarde de plus près on constate que cette crue exceptionnelle a été engendrée par la concomitance de plusieurs facteurs simultanés. La fin de l'année 1909 a été très humide, avec des précipitations de l'ordre de 40% au-delà des normales de saison.

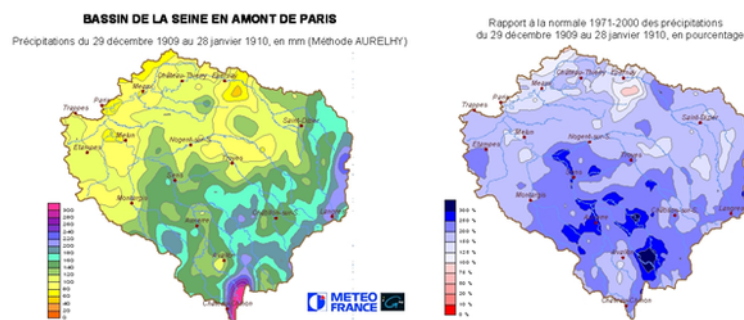


FIGURE 1.23 – Précipitations précédant la crue de 1910. *Source : Météo France.*

Ces pluies à répétition ont alors gorgé d'eau de jour en jour les nappes phréatiques. Durant le mois de janvier, la pluie continue de tomber et des épisodes neigeux font leur apparition, du fait d'une chute de la température. Les sols sont alors gelés, et l'eau ruisselle plus facilement. Le 17 janvier, les précipitations s'intensifient et viennent aggraver

la situation qui commençait déjà à devenir critique. Au fil des jours et des perturbations, la Seine monte.

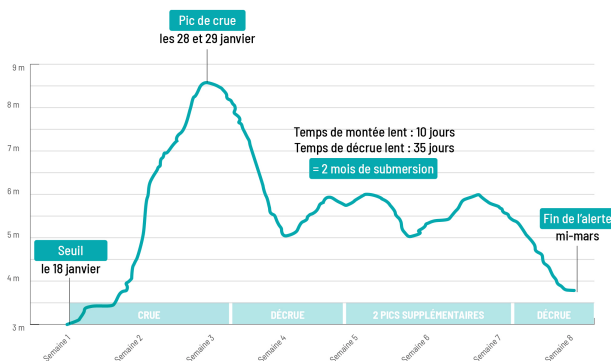


FIGURE 1.24 – Hydrogramme de la crue de 1910. *Source : Episeine.*

Le 20 janvier, la navigation est interdite en raison d'une hauteur trop faible sous les ponts parisiens. Les gares de Lyon et de Saint Lazare sont inondées, le réseau parisien est perturbé, des centaines de rues sont immergées, plus de 50 000 habitations sont touchées. Paris restera inondé pendant plus de deux mois du fait du caractère pluriel de cette crue. En effet, on peut remarquer trois maxima distincts sur son hydrogramme. Ceci est lié au fait que deux fois, alors que la Seine était en décrue pour retrouver son niveau habituel, des épisodes de pluies sont réapparues venant recréer une nouvelle sous-crue. Ce n'est que mi-mars que la Seine redescend sous les 3 mètres et que Paris respire à nouveau. Les pertes économiques engendrées par cette crue seront finalement évaluées à plus de 30 milliards de Francs.

Crue de 2016

Fin mai 2016, des pluies diluviennes s'abattent sur l'Europe. Des masses d'air humides en provenance de l'Atlantique causent des précipitations historiques. Jamais un mois de mai n'avait été aussi pluvieux, établissant même à Paris un record tous mois confondus depuis 1960 [Berruyer, 2016].

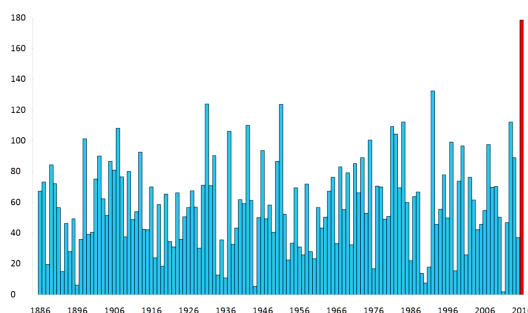


FIGURE 1.25 – Pluviométrie moyenne historique en mai à Paris. *Source : ECA.*

Du 28 au 31 mai, les précipitations sont exceptionnelles, avec un pic atteint le 30 mai. Dans les régions les plus touchées, on relève l'équivalent d'un mois de précipitations sur cette seule journée comme à Trappes ou à Orléans. Les cours d'eau montent alors au fur et à mesure que l'eau de pluies s'accumule, et provoque des crues très importantes. Les 4,63 mètres sont atteints à Nemours en Seine-et-Marne, dans la nuit du 1er juin, battant le record de la crue de 1910 (4,25 m) à cet endroit. L'onde de crue se propage ensuite pour atteindre la Seine via ses affluents, où le maximum est relevé dans la nuit du vendredi 3 au samedi 4 juin, avec 6,10 m. Le graphique ci-dessous retrace l'évolution de la hauteur d'eau de la Seine lors de cette crue et représente le lag entre les précipitations tombées dans le bassin parisien et la montée du niveau du fleuve au niveau du pont d'Austerlitz.

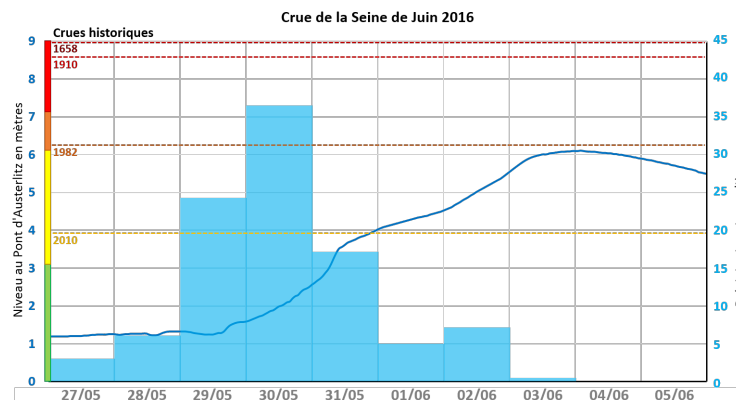


FIGURE 1.26 – Crue de le Seine de mai 2016.⁷ *Source : Kilohn limahn.*

Cette catastrophe a obligé les autorités à faire évacuer plus de 20 000 habitants et a coûté la vie à quatre personnes. Le trafic routier, ferroviaire et fluvial a été entièrement perturbé. L'autoroute A10 a notamment été fermée à la circulation pendant plusieurs jours. Les centres commerciaux des bords de Seine ont été contraints de stopper toute activité. Plus de 25 000 propriétés ont été sinistrées. Le bilan de l'époque est établi à 182 000 sinistres déclarés auprès des assureurs pour un coût total supérieur à 1,4 milliard d'euros selon la FFA (Fédération Française de l'Assurance). Dans son rapport post-catastrophe, elle précisait alors que les indemnisations des sinistres s'étaient élevées à 96% pour l'habitation, 97% pour l'automobile, 90% pour les entreprises et près de 95% pour l'agricole.

1.3.4 Aménagements modernes

Au fil des années et des crues dévastatrices du début du XX^{ème} siècle, l'Etat se décida à intervenir afin de tenter de maîtriser ces phénomènes. Des travaux sont alors entrepris afin de créer un grand lac-réservoir en amont de Paris pour réguler le cours du fleuve. Celui-ci fut inauguré en 1949 et nommé le lac-réservoir de Pannecière. Du haut

7. Évolution du niveau de la Seine à Paris lors de la crue de juin 2016 en regard d'une estimation des précipitations moyennes dans le bassin versant (Orly, Troyes, Reims)

de ses 5,2 km² de superficie et ses 82,5 millions de m³ de capacité, il devint à l'époque le plus grand lac artificiel de France. L'objectif de cet aménagement était d'absorber une partie des inondations qui menacent la capitale. Ainsi, en période de pluie (généralement de novembre à juin), le lac-réservoir se remplit progressivement au moyen de canaux d'aménée et de barrages. Des prélèvements dans le fleuve sont également effectués en période de crues pour limiter l'eau s'écoulant vers la capitale et limiter les inondations. Cette méthode est appelée "écrêtement des crues". A l'inverse, en période de sécheresse l'eau contenue dans le lac est restituée et participe au soutien d'étiage⁸.

Dix-sept ans plus tard et alors que Paris a connu entre temps une nouvelle crue exceptionnelle (7,24 mètres le 23 janvier 1955), un nouveau lac-réservoir est inauguré. Il s'agit du lac-réservoir d'Orient situé à une dizaine de kilomètres de Troyes, dans l'Aube. Il dépasse de près de cinq fois la superficie et trois fois la capacité de son prédécesseur. Deux autres suivront pour constituer l'ensemble des lacs-réservoir de la Seine. Au total, leurs capacités cumulées atteignent les 800m³ d'eau.



FIGURE 1.27 – Carte des quatre lacs réservoirs de la Seine. *Source : Nord Nord West.*

Le problème majeur de ces lacs est qu'ils sont très éloignés de la capitale. Ils protègent alors contre les pluies tombant en amont de leur localisation, mais s'il pleut en aval, ils ne sont d'aucun effet. En chiffres, ils ne permettent de contrôler que 17% du bassin versant de la Seine. De plus, ils permettent de limiter des petites crues mais lors d'un épisode de crue extrême type 1910, ils ne serviraient à réduire la hauteur d'eau que de 60cm au maximum. Cela reste cependant non négligeable, même si insuffisant. En 2005, Pascal POPELIN à l'époque directeur de l'IIBRBS (Institution Interdépartementale des Barrages-Réservoirs du Bassin de la Seine) devenue EPTB Seine grands lacs, estimait à 30% l'économie réalisée sur le coût total des dommages estimés du fait de ces 60cm d'écrêtement, ce qui représente alors plusieurs milliards d'euro d'économies.

8. Niveau le plus bas d'un cours d'eau.

1.4 Le bassin parisien

1.4.1 La Seine à Paris

Paris est traversé en son centre par la Seine, qui s'étend sur environ 13 kilomètres avec une largeur variant de 30 à 200 mètres. Mais si tout le monde connaît la Seine sous son tracé actuel, elle n'a pas toujours été ce qu'elle est aujourd'hui. En effet, à la fin du XXème siècle, la Seine possédait un méandre supplémentaire au nord des environs de la gare de Lyon et de la Bastille jusqu'au pont de l'Alma. Il s'est ensasé avec le temps mais faisait alors à l'époque de cette zone intermédiaire une zone marécageuse et inondable.



FIGURE 1.28 – Ancienne cartographie de la Seine à Paris.

Aujourd'hui, la Seine à Paris comporte deux véritables îles qui sont l'île de la Cité et l'île Saint- Louis. On en compte cependant quatre au total avec l'île Louviers et l'île des Cygnes. En 1856, le sculpteur français Georges Diebolt édifia la statue du Zouave du pont de l'Alma, bien connue aujourd'hui. Pendant des années et en l'absence d'échelle hydrologique, elle servit de référence pour mesurer l'importance des crues dans la capitale. Les petites crues (à partir de 5 mètres environ) voyaient alors l'eau monter jusqu'aux souliers du zouave, signifiant ainsi par la même occasion l'inondation d'une partie importante des actuelles voies sur berges de Paris.

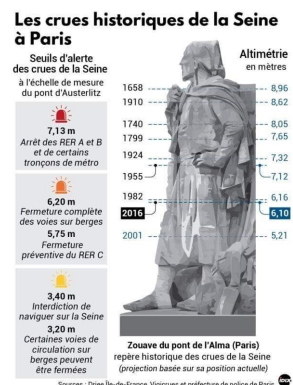


FIGURE 1.29 – Zouave du pont de l'alma et seuils d'alerte. *Source : DRIEE*

FIGURE 1.30 – Zouave du pont de l'alma se voyant lors de la crue de 1910.

1.4.2 Impacts sociaux-économiques d'une grande crue de la Seine

Comme une grande partie des grandes métropoles française, Paris a été construit autour d'un fleuve, la Seine. Le cours d'eau traverse d'ailleurs la capitale en son centre, ce qui explique la présence d'un grand nombre d'infrastructures historiques sur les bords de Seine. L'inondation partielle d'un certain nombre d'entre eux serait alors une catastrophe pour la ville.

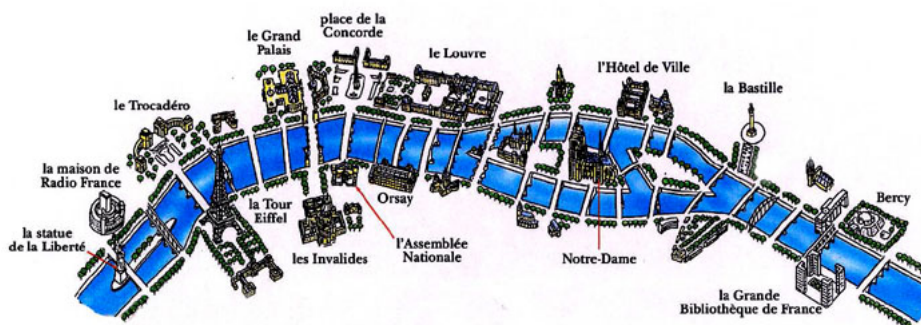


FIGURE 1.31 – Infrastructures historiques en bords de Seine.

En plus de ces bâtiments, un grand nombre de logements seraient touchés par une crue de type 1910. Le problème est que la vulnérabilité du territoire a augmenté depuis 100 ans et que les conséquences seraient aujourd'hui encore plus violentes. Selon l'Institut d'Aménagement et d'Urbanisme, au 31 octobre 2010, près de 435 000 logements sont situés en zone inondable. On recense pas moins de 100 000 logements construits en zones inondables entre 1980 et 2010 et si l'on regarde décennie par décennie, les chiffres sont en constante augmentation (26 100 dans la décennie 1980, 36 000 dans la décennie 1990, 38 000 dans la décennie 2000). Et depuis 2010, le chiffre grimpe à plus de 55 000.

Aujourd'hui, différentes estimations ont été réalisées par plusieurs experts, relayés notamment par l'OCDE (Organisation de Coopération et de Développement Économiques) ou par la ville de Paris. Voici en résumé les impacts d'une crue semblable à celle de 1910 qui surgirait en 2021 :

- **850 000** : Nombre de logements hypothétiquement touchés par une telle crue en Ile de France, dont 240 000 dans Paris ;
- **700 000** : Emplois potentiellement menacés par une telle crue ;
- **5 millions** : Population possiblement impactée en région parisienne par une telle crue ;
- **30 milliards** : Estimation défavorable en euros du coût généré par une telle crue.

Car une crue ne touche pas que les biens des locataires parisiens, caves et parkings compris. Dès l'infiltration de l'eau dans les sous-sols, l'immeuble entier est menacé. En effet, la dégradation de postes électriques ou de mécanismes d'ascenseur par exemple situés dans les niveaux inférieurs peut mettre en berne la totalité du fonctionnement de l'immeuble et couper l'accès au monde extérieur. Tout ceci constitue des facteurs de vulnérabilité supplémentaires.

A un niveau plus macroscopique, c'est toute la ville de Paris qui serait à l'arrêt. L'OCDE alerte dans son rapport de 2014 (effectué sur demande de l'EPTB) sur les conséquences d'une crue majeure de la Seine : plus d'1,5 million de clients pourraient voir leur électricité coupée après l'inondation ou la coupure préventive de près d'un quart des infrastructures de transformation électrique. Les transports publics seraient interrompus dans leur majeure partie avec la fermeture de 140 km du réseau de métro sur environ 250. Les ponts traversant la Seine se retrouveraient alors fragilisés et interdits à la circulation, bloquant l'accès d'une rive à l'autre. L'alimentation en eau potable pourrait également être interrompue dans Paris avec près de cinq millions d'habitants qui verraient apparaître des coupures d'eau prolongées et une dégradation de sa qualité. Les dommages directs d'une telle catastrophe ont été estimés de 3 à 30 milliards d'euros selon les scénarios.

Ces conséquences paraissent énormes mais sont pourtant réelles. On peut observer ci-dessous les impacts déjà visibles de la crue de 2016, pourtant près de 2,5 mètres plus basse que celle de 1910 :

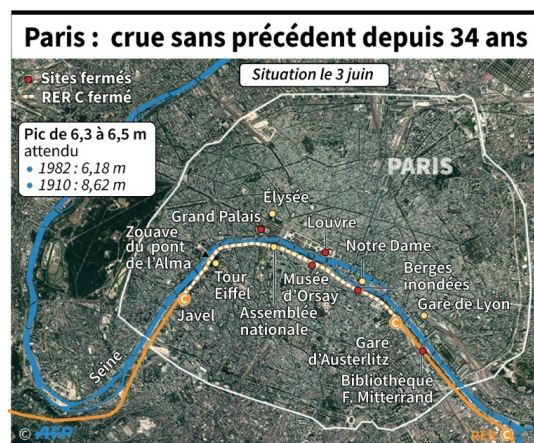


FIGURE 1.32 – Impacts de la crue de 2016 à Paris. *Source : AFP.*

La question n'est pas de savoir si une crue centennale frappera à nouveau Paris et sa région, mais quand. C'est pourquoi le seul levier d'action à ce jour est d'appréhender ce risque de la meilleure façon possible pour être prêt le jour où cela se produira.

1.5 Modélisation d'une crue et de ses impacts

1.5.1 Démarche générale d'un modèle CAT NAT

Un modèle CAT NAT diffère en plusieurs points d'un modèle classique d'évaluation du risque. Sa portée géographique constituée de multiples zones d'impact est un aspect primordial là où il demeure souvent négligeable pour la plupart des autres risques. En effet, la répartition géographique du portefeuille assuré et l'étendue de la catastrophe jouent un rôle fondamental. L'ampleur des dégâts est ainsi très volatile en fonction de la localisation et de l'importance de la catastrophe qui survient. L'évaluation d'un tel

risque suit toujours une démarche bien précise, au travers de l'étude de différents modules indispensables :

- **Module exposition** : C'est l'étude du portefeuille permettant d'apporter les informations nécessaires sur la localisation, les caractéristiques et les valeurs des biens assurés ;
- **Module aléa** : Il consiste à modéliser le risque physique, c'est-à-dire à retranscrire statistiquement le risque d'apparition d'une telle catastrophe naturelle. Afin de la décrire, il s'agit de modéliser sa localisation, sa fréquence de survenance et son intensité ;
- **Module vulnérabilité** : Ce module va allier les résultats des deux premiers. Il sert à déterminer, en fonction de l'intensité de l'évènement modélisé (module aléa), les pertes économiques. Pour cela, nous devons évaluer comment les biens sont touchés en fonction de leurs caractéristiques (module exposition) ;
- **Module financier** : C'est le module final qui va permettre de dégager les pertes réelles de l'assureur à partir des pertes brutes évaluées dans le module vulnérabilité, corrigées des différents éléments contractuels (franchises, limites, coassurance, réassurance, etc...).

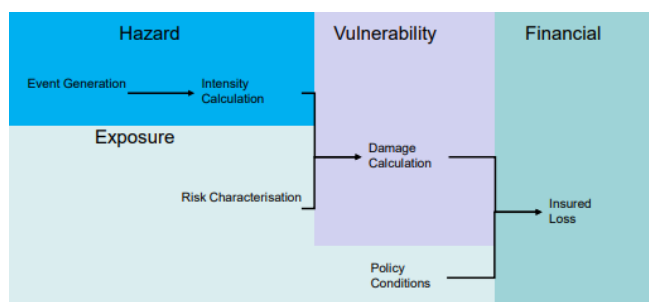


FIGURE 1.33 – Démarche d'un modèle CAT NAT.

Nous nous intéresserons dans ce mémoire uniquement aux trois premiers modules ci-dessus. Le module aléa étant la base de la modélisation et également le plus complexe à mettre en place, c'est celui qui a fait l'objet de la majeure partie du travail et qui sera donc le plus détaillé dans ce rapport. On notera également que le module financier peut avoir plusieurs applications que nous ne souhaitons pas forcément traiter dans ce mémoire.

1.5.2 Module exposition : Cartographie des risques

Le module exposition est un module moins technique mais n'en reste pas moins important. Il consiste en l'analyse la plus fine possible du portefeuille assuré. On trouve ainsi diverses variables ayant une utilité différente en fonction de leur typologie. Dans ce mémoire, nous étudions une base de données MRH avec les informations contrat. On y recense une centaine de variables par police. Nous n'avons retenu qu'un sous-ensemble de ces variables, utile à notre modélisation. Le tableau ci-dessous dresse l'inventaire

catégorisé des variables retenues :

Type de variable	Variable	Détail	Type
Contractuelle	Numéro de contrat	Clé de notre base	integer
	Exposition	Exposition du contrat	numeric
	Capital assuré	Valeur des biens assurés	numeric
Géographique	Latitude	Latitude du bien assuré	numeric
	Longitude	Longitude du bien assuré	numeric
Caractéristique	Habitation	Appartement ou Maison	factor
	Étage	Etage de l'habitation	factor
	Dependence	Nombre de dépendance(s)	factor
	Superficie	Superficie du bien assuré	factor
	Année de construction	Ancienneté du bien	factor
	Nombre de pièces	Nombre de pièces du bien	factor

TABLE 1.3 – Répertoire des variables de la BDD MRH.

On retrouve ici des variables indispensables aux trois prismes définis plus tôt. La valeur des biens assurés définit la perte potentielle de l'assureur face à l'apparition d'une catastrophe. Les caractéristiques des biens assurés tels que l'année de construction par exemple, définissent à quel point le bien est vulnérable à une destruction totale. La localisation du bien par ses coordonnées GPS ou l'étage de l'appartement sont eux indispensables pour déterminer si le bien assuré peut être touché ou non. Nous présenterons plus en détail cette base dans le chapitre 2.

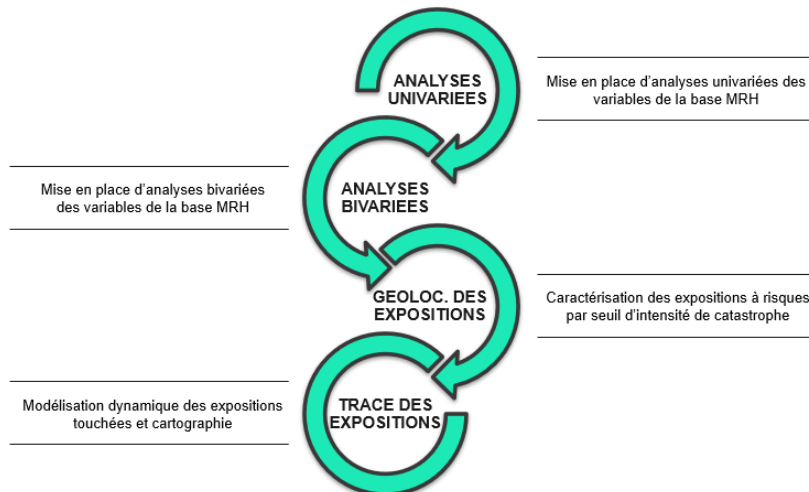


FIGURE 1.34 – Démarche du module exposition.

1.5.3 Module aléa : Calibrage du péril

Comme nous l'avons vu, le module aléa consiste à modéliser le phénomène en lien avec la catastrophe naturelle étudiée, ici l'inondation. L'objectif ici est de simuler les hauteurs d'eau de la Seine sur le bassin parisien, témoin et révélateur de l'apparition d'une inondation.

Pour commencer, il est nécessaire de récupérer les hauteurs d'eau historiques de la Seine en plusieurs points puis de les modéliser. Cette historique permet d'en déduire la distribution empirique de la hauteur d'eau en un point donné et d'en déduire les périodes de retour de crues typiques. Le principal problème de cette méthode est que dans le cadre de Solvabilité II et même plus généralement afin d'obtenir une modélisation consistante, nous manquons de données. C'est d'ailleurs la principale source de difficulté de la modélisation des CAT NAT à ce jour. En effet, il paraît difficile de penser pouvoir capter de façon raisonnable des périodes de retour bicentennaires lorsqu'on ne possède que quelques décennies de relevés météorologiques ou hydrologiques fiables. Car au vu de la plage de données disponible -on rappelle par exemple que l'échelle hydrométrique de référence située au pont d'Austerlitz date du milieu du XIXème siècle- on peut raisonnablement penser que nous n'avons pas encore observé certains événements extrêmes, n'apparaissant que très rarement, et c'est pourtant précisément ce genre d'évènements que l'on souhaite appréhender. Il est évidemment impossible de créer de l'historique, mais un moyen de contourner ce problème est d'étudier en profondeur les phénomènes explicatifs de la montée des eaux d'un fleuve.

Comme nous l'avons vu dans les parties précédentes, de nombreux facteurs peuvent influencer sur le risque de crue, mais le premier et principal paramètre reste la pluviométrie. Le fait de modéliser la pluie et ses impacts sur la Seine peut nous permettre de créer de nouveaux événements et de modéliser de façon plus fine le risque de crue de la Seine sur le bassin parisien. Afin de témoigner l'enrichissement de la modélisation par la considération des précipitations sur le bassin de la Seine, prenons un exemple simple. Supposons un cours d'eau présentant deux affluents de tailles relativement équivalentes. Ce cours d'eau est, comme tous les autres, régit selon une dynamique spatio-temporelle. Sa hauteur et son débit évoluent au cours du temps en fonction notamment des précipitations, toujours sous ce spectre dual. Par exemple, si l'on observe un jour donné des précipitations abondantes en un point A, le fleuve va monter et on risque d'observer en aval une crue au point à risque C. De même si l'on observe de fortes précipitations prolongées au point B.

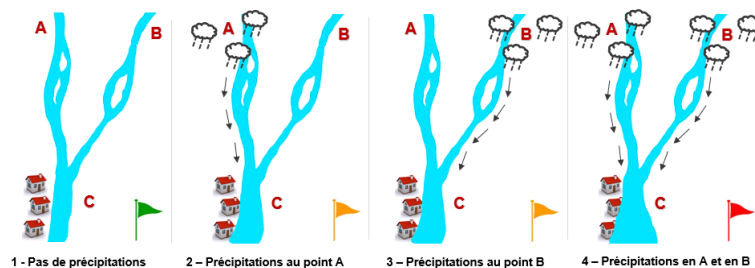


FIGURE 1.35 – Pluies simultanées et crues concomitantes.

Supposons que l'on ait observé ce phénomène mais jamais les deux en même temps. Décrire la dynamique du fleuve selon les précipitations aux points A et B permettra, si l'on réussit à modéliser correctement les pluies en A et en B, ainsi que leur impact sur la hauteur d'eau en C, de créer une nouvelle observation correspondant au schéma 4, et qui risque de témoigner d'une crue encore plus sévère que dans notre historique au point C. Ce type de crue résultante de pluies abondantes sur plusieurs de ses affluents en même temps est appelée une crue concomitante. Ce phénomène a été très bien décrit par Huguette Vivian dans sa thèse hydrométéorologique réalisée à l'université de Lille [Vivian, 1977]. Il faut alors bien garder en tête qu'une crue à un point donnée n'est pas forcément la résultante de pluie à cet endroit, mais parfois de pluies bien en amont de la zone touchée.

À plus fine échelle, il reste maintenant à connaître le temps que mettent les gouttes de pluies tombant dans un secteur à arriver au point C. Considérons par exemple que l'eau de la branche droite s'écoule plus doucement que sur l'autre branche du fait d'une distance ou d'une pente plus élevée. Si l'eau met trois jours à s'écouler du point A au point C et quatre jours du point B au point C, l'évènement qui sera le plus dangereux est d'observer de fortes pluies au point B à un instant donné et au point A le lendemain. Ce phénomène des crues concomitantes est complexe à modéliser et nécessite de connaître la cartographie de l'écoulement du bassin versant. Mais ce n'est pas tout, car il s'agit de savoir modéliser des précipitations sur plusieurs stations à la fois, et on se doute que le fait de savoir qu'il pleut à un instant donné au point B laisse à penser qu'il est possible d'observer des précipitations au point A au même moment, voir après ou avant. La modélisation des précipitations et plus généralement la science de la météorologie est également très complexe du fait de la dynamique spatio-temporelle des épisodes météorologiques.

Ainsi, la première étape de notre démarche sera de modéliser les précipitations indépendamment sur plusieurs stations du bassin versant de la Seine. Afin de prendre en compte les différentes tendances saisonnières (les précipitations ne suivent pas la même dynamique selon les saisons) nous considérerons un modèle différent par mois de l'année. Ensuite, nous traduirons le phénomène de dépendance spatio-temporelle des précipitations selon les stations via l'utilisation de copules paramétriques calibrées avec notre historique de données. Enfin, nous effectuerons un modèle final permettant d'expliquer la hauteur d'eau observée à Austerlitz avec les précipitations relevées sur l'intégralité du bassin versant.

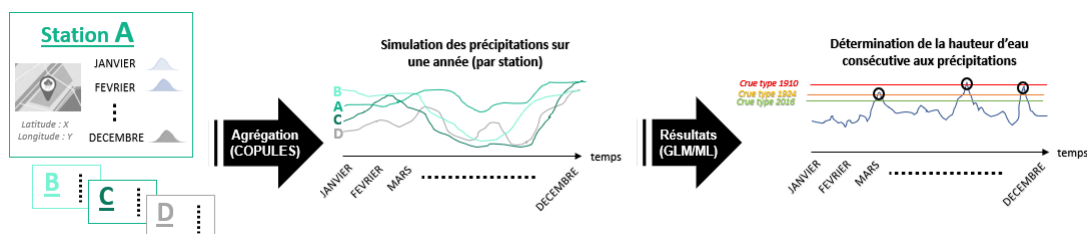


FIGURE 1.36 – Évaluation des périodes de retour par simulations.

La détermination des périodes de retour s'effectuera par la simulation d'un grand nombre d'années entières. Nous modélisons la pluie en continue sur l'année entière par pas de temps réguliers (12h) sur l'ensemble du bassin versant et avec le modèle explicatif on en déduit la hauteur d'eau à Austerlitz au même moment. Ainsi, nous pouvons repérer les années simulées qui observent une crue d'une ampleur donnée, comme celle de 1910 par exemple. Cela nous donnera la réponse quant à la période de retour de cette crue si célèbre et nous permet d'appréhender de façon complète le risque de crue de la Seine et ainsi de réaliser différents types d'applications actuarielles.

En résumé, l'objectif du module aléa est d'élaborer un catalogue d'épisodes de crues fictives générées aléatoirement. On pourra ainsi évaluer les périodes de retour correspondant à différents niveaux de crue et les pertes engendrées sur le portefeuille.

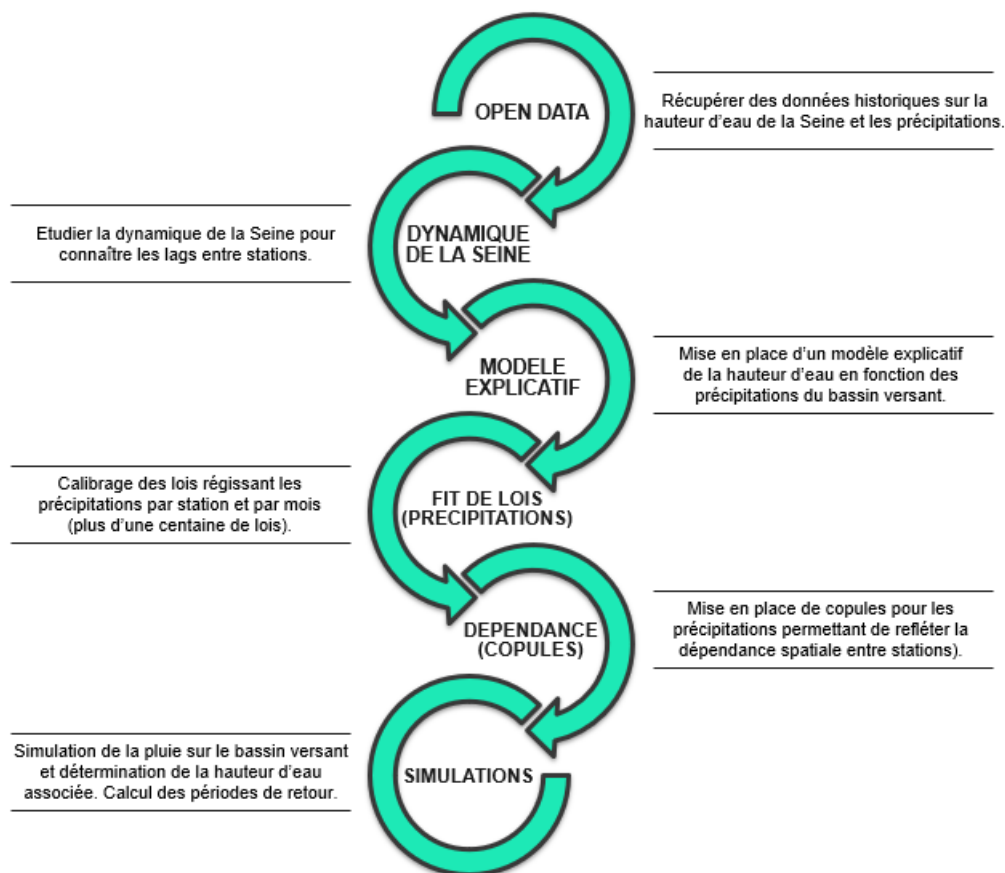


FIGURE 1.37 – Démarche du module aléa.

Toute cette démarche sera détaillée dans le chapitre 3.

1.5.4 Module vulnérabilité : Évaluation de l'impact

Comme signifié plus tôt, le module vulnérabilité relie les deux premiers modules. Il permet de décrire selon la gravité de la catastrophe, quelles seront les conséquences sur les expositions.

L'indicateur communément utilisé est le taux de destruction, défini comme le rapport entre la charge du sinistre et la somme assurée. Il met en lumière à quel point une exposition sera touchée en cas de survenance d'une catastrophe. Cependant, une exposition n'est pas touchée de la même manière selon l'intensité de la catastrophe. En effet, les dégâts ne seront pas identiques selon que l'on observe une petite crue ou une crue de type 1910. Il faut alors conditionner notre taux de destruction à l'intensité de la catastrophe. On parle alors de courbe de vulnérabilité où l'on retrouvera en abscisse cette intensité, correspondant dans notre cas à la hauteur d'eau, et en ordonnées le taux de destruction associé. Mais ce taux peut également être affiné selon les caractéristiques des biens assurés. En effet, une maison de plein pied ne subira pas les mêmes dégâts qu'un appartement au 1er étage.

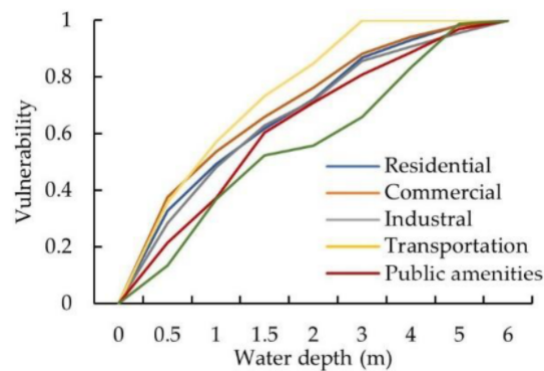


FIGURE 1.38 – Exemple de courbe de vulnérabilité.⁹

Ces taux de destruction sont complexes à définir de façon robuste, toujours du fait du manque d'historiques d'évènements catastrophes. Ils peuvent ainsi être déterminés à l'aide d'avis d'experts, étudiant par exemple la réaction de différents matériaux face à l'inondation, ou à l'aide de sinistres historiques, aussi peu nombreux soient-ils. Enfin, le calibrage des taux de destruction énoncé plus tôt, prend en considération une valeur moyenne que l'on applique ensuite à nos expositions. En réalité, une structure ne réagira pas toujours de la même manière en fonction de l'aléa qui surgit, quand bien même les taux de destruction seraient calibrés de façon optimale. On parle alors de distribution des taux de destruction, où à chaque hauteur d'eau donnée correspond une distribution du taux de destruction ayant pour moyenne la valeur explicitée sur la courbe. Pour se faire, nous évaluerons un taux de destruction moyen par zone et par étage.

Toute cette démarche sera détaillée dans le chapitre 2.

9. Courbes de vulnérabilité provenant de l'étude de Quntao Yang : *Improved Framework for Assessing Vulnerability to Different Types of Urban Floods*, 2020.

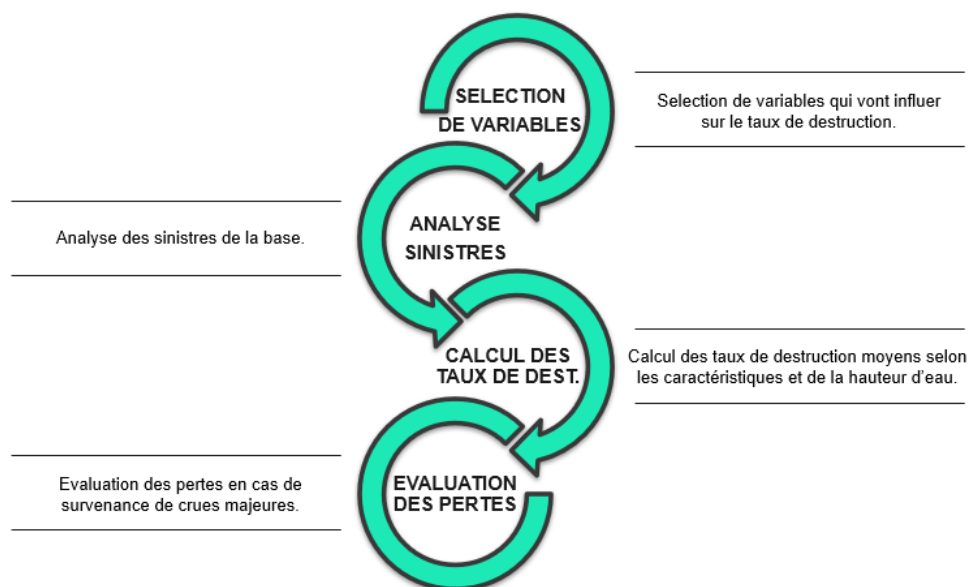


FIGURE 1.39 – Démarche du module vulnérabilité.

1.5.5 Module financier : Calcul des pertes

Le module financier permet de calculer à partir des pertes brutes évaluées dans le module vulnérabilité, les pertes nettes des conditions d'assurance et de réassurance. Il faut alors prendre en compte les différents éléments contractuels tels que les franchises ou les limites définis pour chaque contrat d'assurance, ainsi que la structure de réassurance propre à l'assureur. C'est notamment sur ce module que les assureurs possèdent un fort levier d'action. Dans ce mémoire, nous nous limitons aux dommages aux biens mais différents services en actuariat peuvent s'intéresser à un tel sujet. Modification des conditions contractuelles et tarifaires, optimisation de réassurance, risques humains, voici plusieurs domaines d'applications de la modélisation réalisée dans ce mémoire.

Le chapitre suivant s'intéresse aux modules d'exposition et de vulnérabilité, au travers d'une cartographie du risque réalisée à l'aide de données externes et à la réalisation d'une cartographie entière du bassin parisien.

Chapitre 2

Cartographie du risque inondation en région parisienne

Dans ce chapitre, nous allons cartographier le risque associé à un portefeuille MRH. Premièrement, nous allons déterminer pour chaque niveau de débordement de la Seine quels sont les biens qui seraient touchés et avec quel niveau d'immersion. Ensuite, nous allons calibrer des taux de destruction pour évaluer les dégâts moyens associés à chaque exposition sous risque. Ces deux paramètres permettent alors de calculer les pertes totales brutes pour différentes intensités de crue.

La première étape de cartographie du risque est primordiale. En effet, il convient de savoir comment notre portefeuille d'assurés sera touché en cas de crue, qu'elle soit mineure ou majeure. Cette problématique est typique d'une modélisation CAT NAT, car ce sont des événements qui touchent un grand nombre de polices en même temps. La géolocalisation de ces polices est primordiale pour la modélisation d'une inondation. C'est d'ailleurs un des périls les plus difficiles à modéliser car la localisation joue un rôle fondamental, plus encore que pour une tempête ou un épisode sécheresse par exemple, qui font intervenir une zone étendue mais unique. L'inondation touche quant à elle de multiples zones parfois limitées et distantes des autres. En effet, il peut arriver que deux foyers relativement proches ne soient pas touchés de la même façon, l'un peut être sinistré tandis que l'autre, surélevé par rapport au premier ne l'est pas.

Ainsi, nous devons être capable d'identifier les polices à risque conditionnellement à l'intensité du phénomène. Quelle part de notre portefeuille sera touchée par une crue de type 1910 : 1%, 5%, 10%, 30% ? Une méthode dite "historique" serait de prendre l'historique d'un portefeuille et de regarder quelle part a été touchée lors d'un événement similaire historique. Mais plusieurs problèmes s'apparentent :

- **Répartition des risques** : Si nos risques ne sont pas géographiquement équirépartis (par exemple les risques les plus élevés concentrés en zones inondables), ce taux global reflètera mal le risque réel ;
- **Evolution de la souscription** : Le portefeuille peut vraisemblablement ne plus refléter le profil de risque actuel de l'assureur ;
- **Absence d'historique** : Les crues majeures se faisant rares, il est impossible de collecter de réelles données viables sur une crue type 1910.

Ainsi, une méthode plus complexe à mettre en place mais plus robuste est de cartographier le risque réel de l'assureur. Pour cela, il faut déterminer quelle zone sera touchée en cas de survenance d'une crue d'une intensité donnée, et effectuer un parallèle avec l'exposition totale sur cette zone. Dans le cas d'une crue, l'intensité se mesure avec la hauteur d'eau, qui déterminera les zones touchées (zones d'altitudes inférieures au point culminant de la crue).

Dans ce mémoire nous nous intéressons au bassin parisien uniquement¹, car c'est d'une part la zone où l'on possède le plus d'exposition et d'autre part le territoire français le plus à risques de par la concentration des risques en son sein. Néanmoins, la méthodologie peut se répliquer sur n'importe quel autre bassin, le point crucial étant la présence et la qualité des données.

Nos épisodes de crues se forment à Paris, avec une hauteur d'eau mesurée à Austerlitz, et notre exposition se limitera au bassin parisien, englobant tout ou une partie des départements de Paris (75), des Yvelines (78), des Hauts-de-Seine (92), de Seine-Saint-Denis (93), du Val-de-Marne (94) et du Val d'Oise (95). On notera que cette démarche est cependant répliquable à n'importe quel endroit du globe.

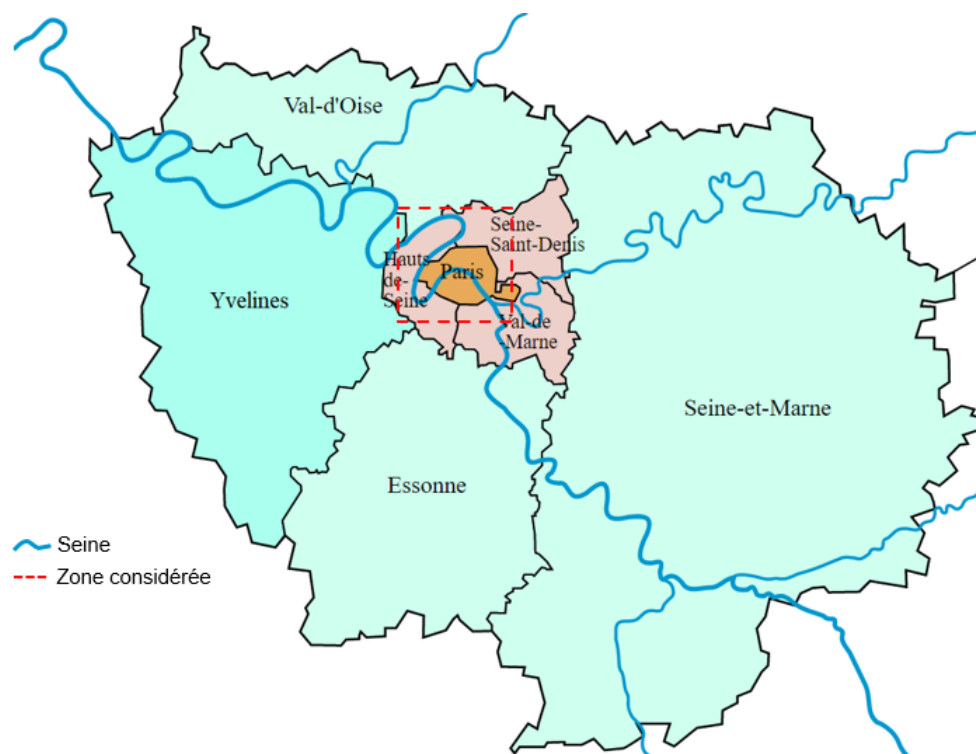


FIGURE 2.1 – Zone d'exposition considérée pour notre étude.

1. La zone d'exposition considérée est déterminée par un rectangle englobant le bassin parisien. Ses coordonnées sur le globe (latitude;longitude) sont : (48.7885;2.19551) - (48.93233;2.19551) - (48.93233;2.486168) - (48.7885;2.486168).

2.1 Topographie générale

2.1.1 Problématique

Topographie de la Seine

Des cartes réalisées par des organismes spécialisés permettent d’obtenir un aperçu des zones impactées en cas de crue. Ils s’appuient sur des crues historiques en déterminant les zones précédemment touchées pour ensuite les retranscrire sur une carte moderne. Cependant, on ne possède qu’une version figée de ces cartes, il est donc impossible selon une police donnée (longitude et latitude) de déterminer aisément si elle sera touchée ou non. En effet, ces cartes ne sont pas dynamiques, on ne peut donc pas déterminer avec précision la part d’expositions touchées par une crue d’intensité donnée. Enfin, Paris est en constante évolution et le Paris de 1910 n’est plus le Paris d’aujourd’hui, c’est pourquoi la crue n’avait à l’époque pas touchée exactement les mêmes zones que celles qui seraient touchées aujourd’hui.

Parmi elles, on peut citer les cartes des PHEC (Plus Hautes Eaux Connues). Pour la Seine à Paris, elle correspond à la crue de 1910.



FIGURE 2.2 – Carte des plus hautes eaux connues de la Seine sur le bassin parisien.
Source : DRIEE.

Topographie de Paris

Un autre élément majeur de cette problématique de cartographie des risques est l’altitude. Comme précisé précédemment c’est l’altitude de nos expositions qui fera figure de baromètre pour déterminer si une exposition est touchée ou non, mais cette information n’étant pas disponible dans notre base assurée, il nous faut trouver un moyen de l’obtenir.

C'est ce que nous avons fait au moyen d'une cartographie de la topographie du bassin permettant d'obtenir en fonction des coordonnées GPS (latitude et longitude) d'une exposition, son altitude et ainsi son risque d'être inondée. Pour ce faire, nous avons besoin de la maille la plus fine possible afin de déterminer avec précision l'altitude d'une exposition, et ainsi réussir à capter les disparités d'altitude entre zones voisines. Encore une fois, il existe sur internet des cartes de ce type mais la maille macroscopique et leur caractère figé ne nous permettent pas de les utiliser. Voici une de ces cartes :

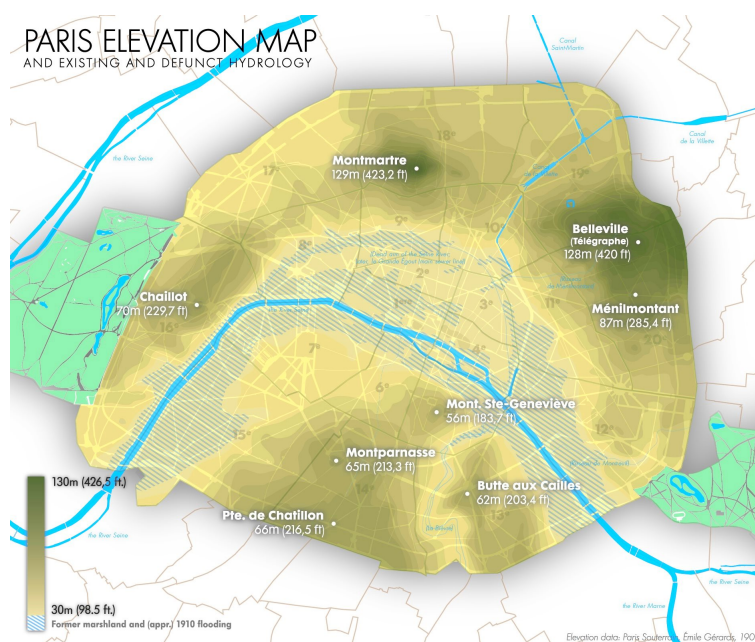


FIGURE 2.3 – Carte topographique de Paris. *Source : ThePromenader.*

2.1.2 Base de données COPERNIC

Afin de résoudre cette problématique de cartographie du risque, nous avons eu recours à l'open data. La base de données que nous avons utilisée est la base EU-DEM fournie par le programme européen *Copernicus*. C'est un programme coordonné par la commission européenne en partenariat avec l'Agence spatiale européenne, les membres de l'Union Européenne et diverses agences de l'Union Européenne. Les données collectées sont nombreuses et qualitatives. Elles contribuent notamment à la gestion de différentes problématiques : sécuritaires, environnementales et également à assurer le suivi des catastrophes. Ses satellites effectuent des missions d'observation du sol, des océans et de l'atmosphère, qui à l'aide de plusieurs instruments (caméras, radars, spectromètres, altimètres, radiomètres, ...) recueillent des données sur l'ensemble du continent européen et du globe tout entier.

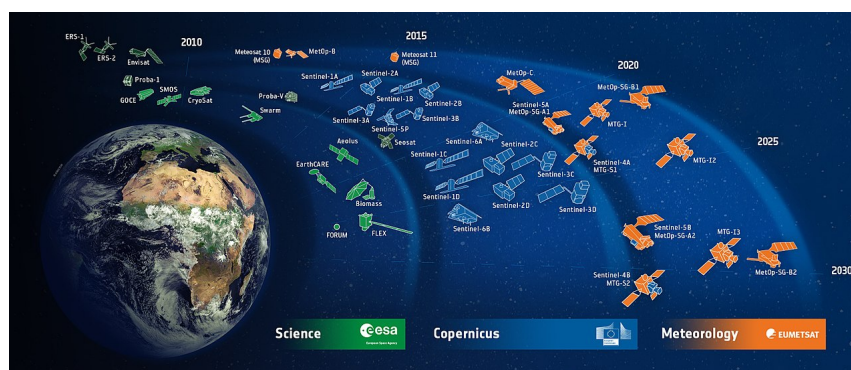


FIGURE 2.4 – Satellites d’observation de la Terre³. Source : ESA (European Space Agency).

Les données qui nous intéressent sont les données d’altitude. Le programme *Copernicus* a ainsi cartographié la quasi-totalité de l’Europe au travers d’un quadrillage fourni de tuiles de 100 x 100 km avec une précision d’environ 25 mètres en largeur et en hauteur. On devine alors le nombre gigantesque de données collectées et donc la lourdeur des fichiers associés. A une tuile correspond quelque 16 millions de pixels (40 000 x 40 000). Le quadrillage suit les grands méridiens du globe et est numéroté en fonction de ces derniers.

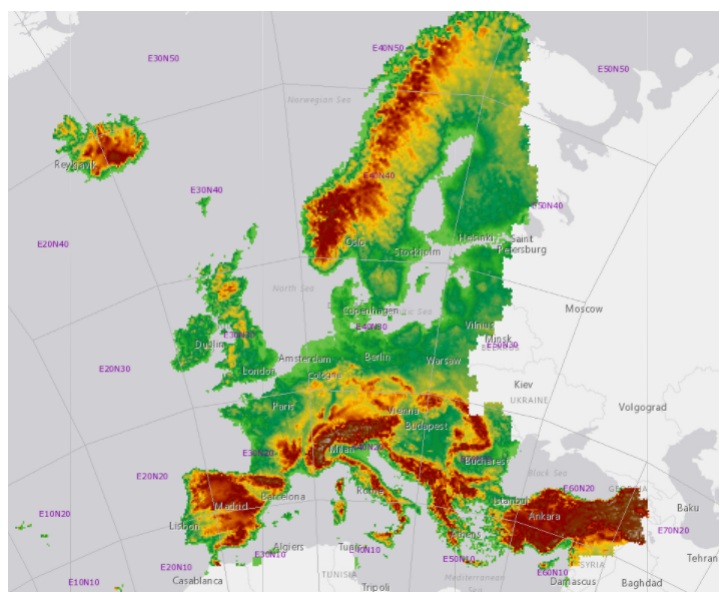


FIGURE 2.5 – Vue d’ensemble des données Copernicus⁴. Source : Copernicus.

3. Satellites opérés par l’ESA (Agence Spatiale Européenne) et l’EUMETSAT (Organisation européenne pour l’exploitation des satellites météorologiques) fournissant des données au programme Copernicus.

2.2 Cartographie de la topographie parisienne

2.2.1 Méthodologie

Les données décrites dans la partie précédente nous permettent ainsi de créer la topographie du bassin parisien. Dans la base de données EM-DAT, la France est séparée en quatre mailles, dont la principale recouvrant notamment Paris est la maille E30N20. Ces données ont été récoltées en 2011 et publiées en 2016. Elles sont au format Geotiff, nous les avons traitées à l'aide du logiciel **R**, avec la bibliothèque **raster** et le package **leaflet**, nous permettant de réaliser nos propres cartes topographiques de la zone qui nous intéresse, à savoir le bassin parisien. On notera que la lourdeur des fichiers ne permet pas avec **leaflet** l'affichage d'un pays entier, ou même de l'Europe comme on peut le voir dans la figure 2.8. Il permet cependant un affichage dynamique de la carte ce qui permet de zoomer sur les zones intéressantes. Ici, on présentera simplement différents zooms afin de donner au lecteur un aperçu de la cartographie réalisée. Le premier sur le bassin entier, puis sur Paris, et enfin un zoom au niveau du pont d'Austerlitz, là où les crues de la Seine sont relevées.

2.2.2 Résultats

On retrouve ci-dessous un aperçu de la cartographie de la topographie du bassin parisien obtenu grâce aux données EM-DAT et au package **leaflet** :

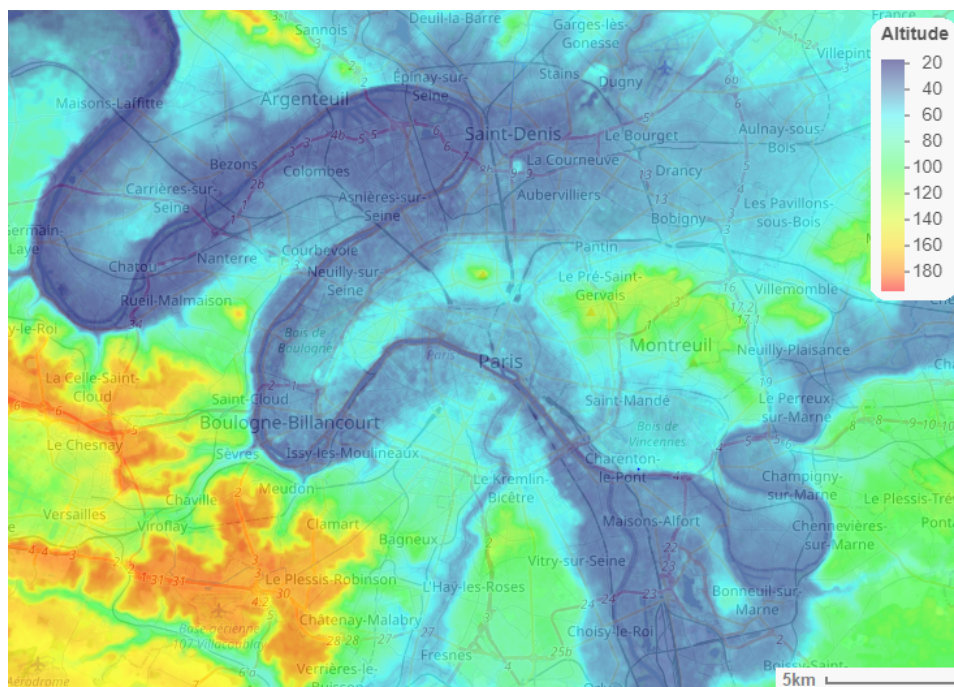


FIGURE 2.6 – Topographie : Zoom sur le bassin parisien. *Données : Copernicus.*

On présente maintenant un zoom sur la ville de Paris qui possède une altimétrie plutôt faible et assez uniforme :

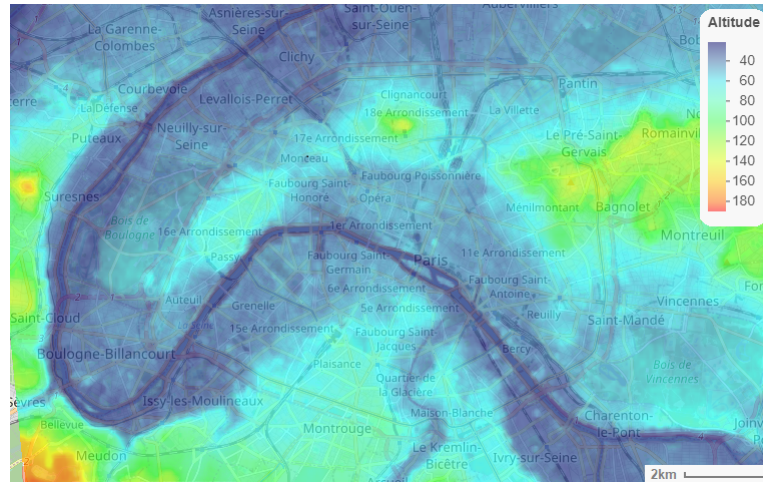


FIGURE 2.7 – Topographie : Zoom sur Paris. *Données : Copernicus.*

Enfin, voici un zoom au niveau du pont d'Austerlitz, permettant réellement d'apprécier la finesse de ces données :

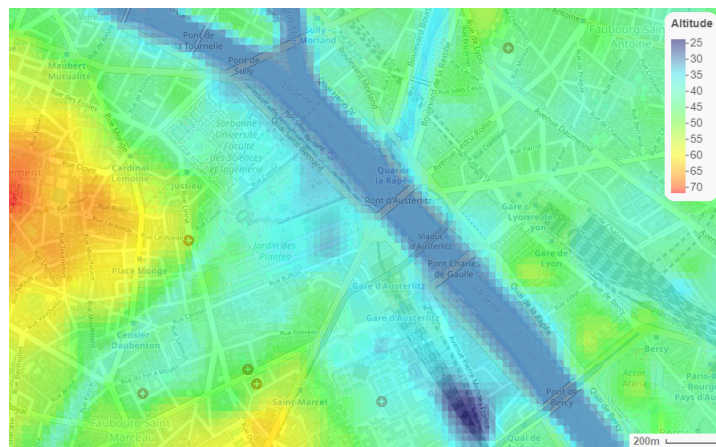


FIGURE 2.8 – Topographie : Zoom sur le pont d'Austerlitz. *Données : Copernicus.*

On notera que la largeur de la Seine traversant Paris est comprise entre 30 et 200 mètres, ce qui signifie que nous pouvons cartographier le fleuve grâce à ces données, car de 1 à 9 pixels cartographient le fleuve sur la totalité du bassin. De plus, cette cartographie du bassin parisien nous permettra plus tard de pouvoir associer à chaque exposition, une altitude donnée. Il reste alors à créer la cartographie d'empreintes de crues pour connaître les zones impactées par une montée des eaux et ainsi savoir quelles expositions seront touchées ou non.

2.3 Cartographie de la Seine

2.3.1 Méthodologie

Il est également possible de cartographier le fleuve grâce aux données *Copernicus*. La méthode utilisée a été la suivante : Cartographier la Seine en conditionnant l'altitude par un seuil. Ainsi, tous les pixels inférieurs à ce seuil seront considérés comme une partie du fleuve, et ceux au-dessus de ce seuil comme du terrain. La finesse de la maille étant une condition sinéquanone à la pertinence de la réalisation de cette cartographie par la méthode décrite ci-dessus. Les quais de Seine étant tous surélevés, cela garantit la bonne délimitation du fleuve sur ses rives, quant aux zones de basses altitudes de Paris étrangères au fleuve, il suffit de les retraiter en les enlevant de la cartographie.

Cependant, une difficulté apparaît lors de l'utilisation de cette méthode : la hauteur de la Seine n'est pas la même sur tout le bassin. En effet, la Seine possède un dénivelé d'environ 50 cm par kilomètre en moyenne, ce qui signifie qu'à un instant donné, si le fleuve est à une certaine altitude par rapport au niveau de la mer à l'entrée de Paris, il sera plusieurs dizaines de décimètres plus bas en sortie de Paris. Le seuil à partir duquel on considère un pixel comme fleuve est donc variable selon l'endroit. On notera que la profondeur du fleuve, très variable sur tout le réseau fluvial, ne nous est pas utile pour notre étude.

2.3.2 Résultats

Afin de déterminer ces différents seuils, nous avons représenté la cartographie de la Seine obtenue en fonction du seuil considéré. Cela nous permet d'observer l'altitude du fleuve dans chaque région au fur et à mesure que l'on augmente le seuil. On observe par exemple que pour obtenir la cartographie de la Seine autour de Maisons-Laffitte et ce jusqu'à Rueil-Malmaison, c'est un seuil de 24 mètres que nous allons utiliser. Entre Rueil-Malmaison et Bezons, on utilisera le deuxième seuil correspondant à 25 mètres d'altitude, et ainsi de suite.

Les résultats sont présentés ci-dessous avec la représentation des tracés pour chacun des huit seuils définis de 24 mètres à 31 mètres. On voit bien la nécessité d'effectuer un découpage spatial par seuil car si l'on souhaite capter la Seine aux abords d'Alfortville (Seuil 8), le fleuve sort complètement de son lit au niveau de Saint-Denis. Cela témoigne bien de la présence d'une dénivellation de la Seine dans Paris.

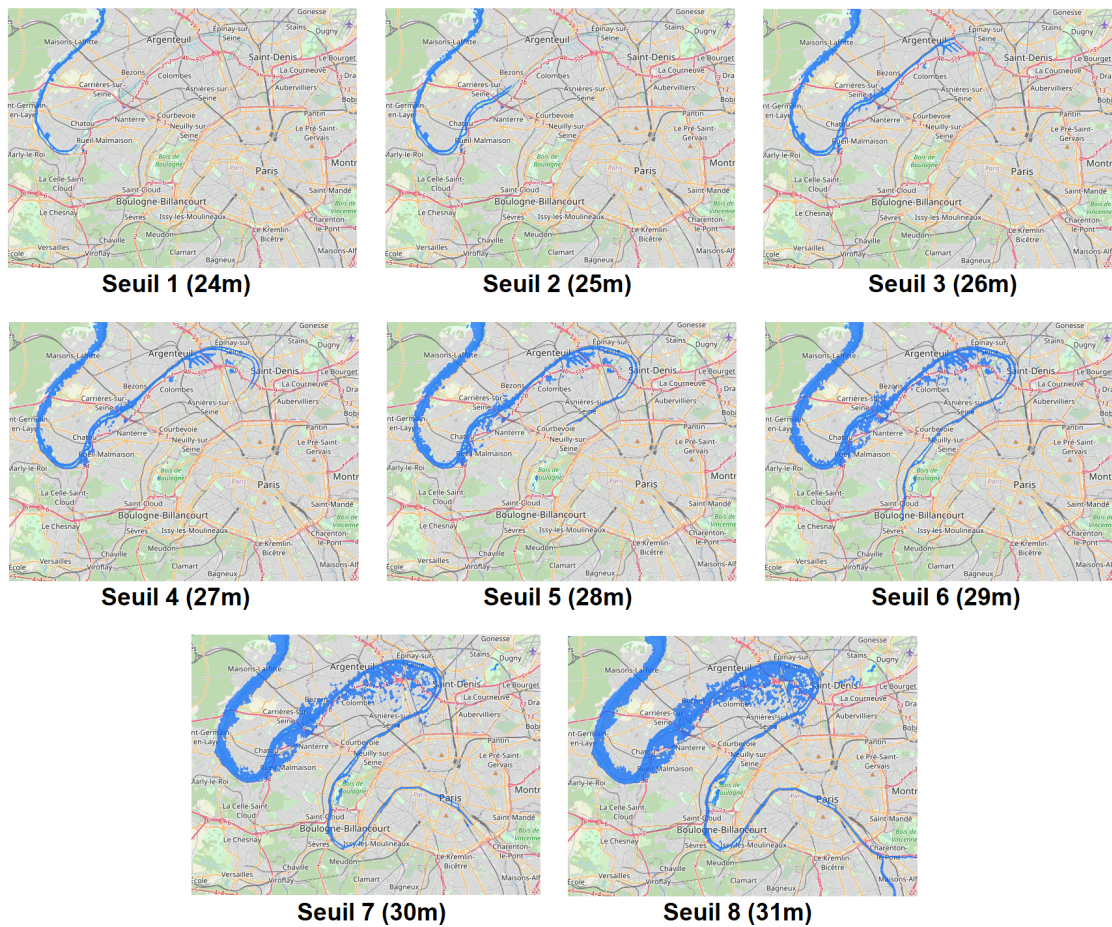


FIGURE 2.9 – Cartographie de la Seine selon les seuils. *Données : Copernicus.*

Le travail est maintenant de définir notre découpage spatial à l'aide de ces cartes afin d'obtenir une cartographie seuillée par zones et de réunifier ces zones pour obtenir la cartographie complète de la Seine sur le bassin parisien. La méthode pour réaliser le découpage a été de partir du seuil le plus bas, et de découper le fleuve à l'endroit où il n'apparaît plus sur la carte. Par exemple comme on l'a dit plus tôt, le seuil 1 permet de capter le fleuve entre Maisons-Laffitte et Rueil-Malmaison. À partir de Rueil-Malmaison on utilise le seuil 2, et ainsi de suite.

La première carte ci-dessous représente la concaténation des cartographies présentées ci-dessus. À chaque zone colorée correspond le seuil témoignant de la hauteur approximative du fleuve dans cette région. On présente ensuite une vue en éclaté de cette cartographie reflétant le découpage effectué, avec les différentes zones de différenciation.

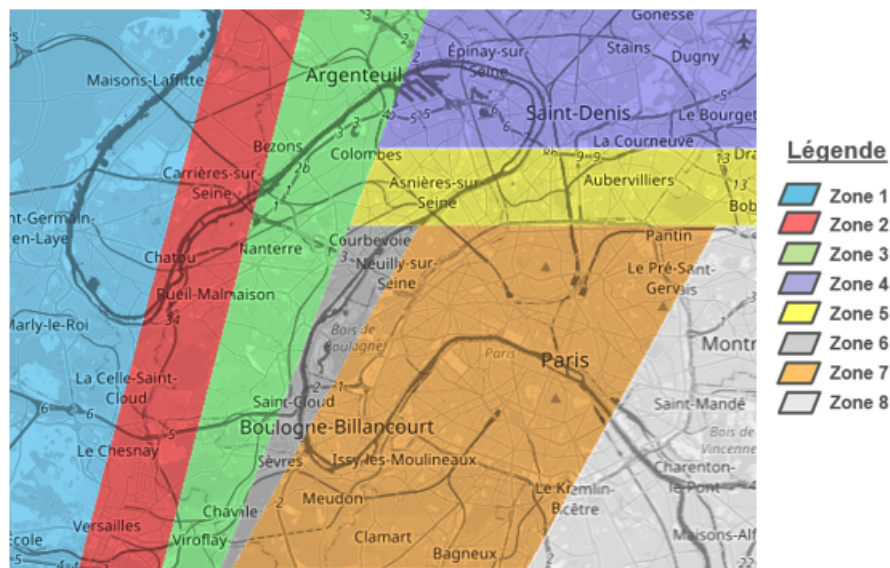


FIGURE 2.10 – Cartographie seuillée par zones.

Voici les huit parties que nous allons regrouper pour former la cartographie du fleuve :

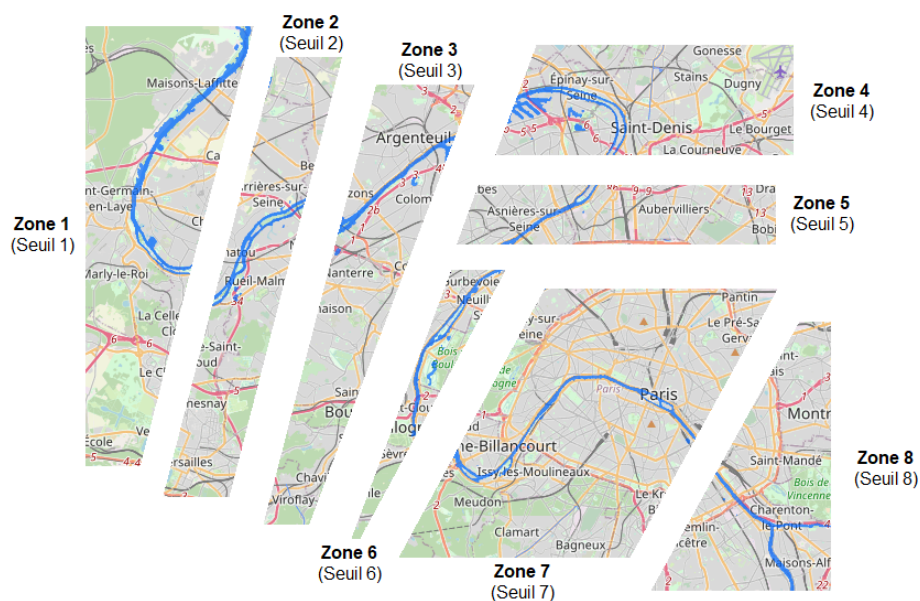


FIGURE 2.11 – Vue en éclaté des zones de différenciation.

Enfin, une fois les zones regroupées et les impuretés retirées, on obtient notre cartographie finale, que l'on peut comparer à celle de Google Maps par exemple :

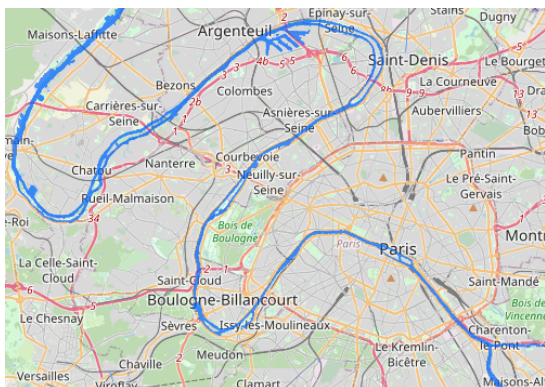


FIGURE 2.12 – Cartographie de la Seine (Modélisation).

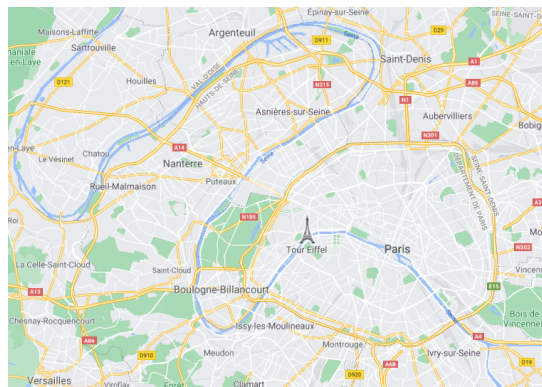


FIGURE 2.13 – Cartographie de la Seine (Google Maps).

2.3.3 Empreintes de crues

La cartographie présentée plus tôt nous sert alors de base, et nous allons pouvoir maintenant réaliser des cartes d’empreinte de crues. Pour se faire, nous partons de cette cartographie et nous augmentons le niveau d’eau en lien avec l’échelle d’Austerlitz. Selon la hauteur d’eau considérée, nous pouvons ainsi tracer les zones parisiennes inondées, et ainsi en déduire nos expositions touchées. On notera que dès le premier centimètre d’eau, la zone est considérée comme touchée.

On présente ci-dessous la carte d’empreinte de la crue de 1910 réalisée à partir de notre cartographie, que l’on compare encore une fois, à une carte provenant d’études. Celle-ci est la carte d’empreinte de crue référence réalisée par le ministère de la transition écologique et solidaire. On la retrouve sur le site geoportail. Encore une fois, les cartes sont relativement proches ce qui nous conforte sur la pertinence de notre module d’exposition et de cartographie.

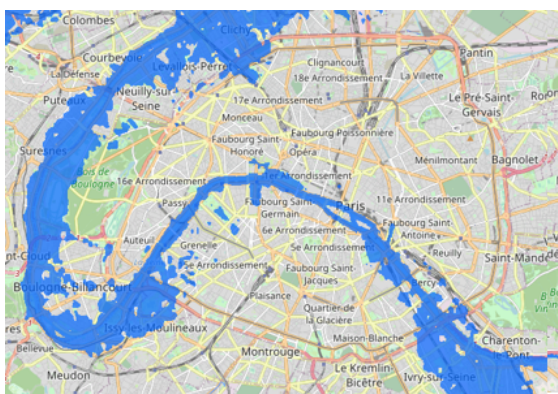


FIGURE 2.14 – Empreinte de la crue de 1910 (Modélisation).



FIGURE 2.15 – Empreinte de la crue de 1910 (Geoportail).

2.4 Cartographie des expositions

Maintenant que l'on possède toutes nos cartes topographiques, il nous reste à déterminer l'altitude de chaque exposition. Pour cela, il nous faut tout d'abord récupérer les coordonnées GPS de ces expositions à partir de leur adresse. Nous pourrions ensuite alors associer à chaque contrat de notre base, l'altitude associée afin de déterminer de façon binaire si une exposition est dans l'empreinte de crue d'une hauteur donnée.



FIGURE 2.16 – Obtention des altitudes de nos expositions.

2.4.1 Scrapping de données

La première étape de cette cartographie des expositions est la localisation précise de nos biens assurés. Ne disposant que de données macroscopiques (pays, département, commune) ou de données non exploitables en l'état (adresses), un travail de scrapping a été nécessaire. L'objectif est de géocoder ces adresses pour en tirer des variables géographiquement exploitables en vue d'une cartographie à maille fine (latitude et longitude). Pour effectuer ce travail, nous avons utilisé une API libre de droits mise en place par le gouvernement qui combinée avec la BAN (Base Adresse Nationale) permettant la récupération de coordonnées géocodées à partir de l'adresse complète. Nous avons donc récupéré ces données pour l'ensemble de nos contrats.

2.4.2 Analyse de la base de données contrats MRH

On rappelle que le portefeuille que nous considérons est un portefeuille MRH provenant d'une compagnie d'assurance française. Nous présentons ci-dessous quelques statistiques générales quant aux nombres de contrats, à l'exposition totale et aux sommes assurées totales de la base initiale ainsi que la base réduite à notre zone d'étude tel que défini plus tôt.

- **Nombre de contrats** : 1 287 097
- **Exposition totale** : 602 596
- **Sommes assurées totales** : 76 510 175 000
- **Nombre de contrats - zone d'étude (proportion)** : 109 156 (8,5%)
- **Exposition totale - zone d'étude (proportion)** : 50 591 (8,4%)
- **Sommes assurées totales - zone d'étude (proportion)** : 5 459 758 000 (7,1%)

Les habitations assurées possèdent des caractéristiques différentes, ce qui modifie leur vulnérabilité à une inondation. Nous avons décidé de nous intéresser particulièrement à certaines de ces variables qui sont pertinentes pour notre étude :

- **Type de logement** : Deux possibilités, "appartement" ou "maison". Il y a une majorité d'appartements dans la base, ce à quoi on pouvait s'attendre du fait de la concentration d'habitations de ce type dans la capitale.

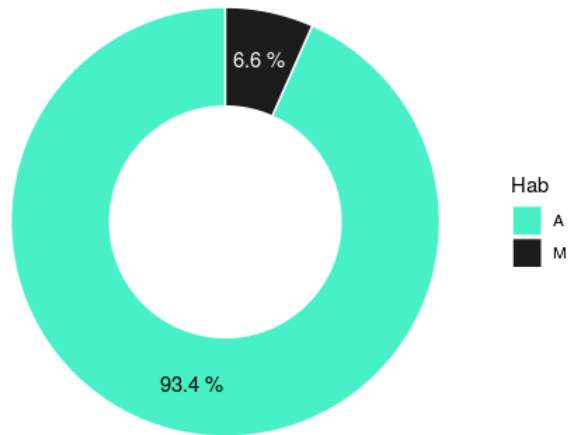


FIGURE 2.17 – Parts d'appartements et de maisons dans la base.

- **Étage** : L'étage est une variable très importante car un appartement au dernier étage ne sera pas touché de la même façon qu'un appartement au rez-de-chaussée. Pour la base de données que l'on possède, il est indiqué "maison" lorsque le bien assuré est une maison (on retrouve alors nos 6.6%) et lorsque l'on est en présence d'un appartement, l'information fournie est "Rdc" (rez-de-chaussée), "Inter" (étage intermédiaire) ou "Dernier" (dernier étage). Pour ce qui est de nos expositions, on observe que près de trois habitations sur quatre sont des appartements aux étages intermédiaires.

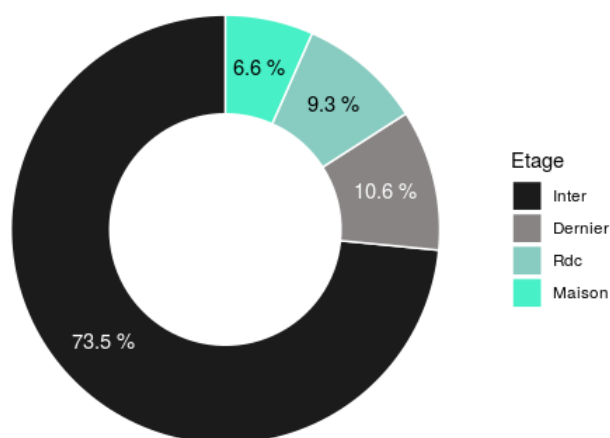


FIGURE 2.18 – Parts d'exposition selon l'étage de l'appartement.

- **Présence d'une dépendance** : La quasi-totalité des habitations ne possèdent pas de dépendance. Cela provient du fait que seules les maisons ont la possibilité d'en posséder et que peu en possèdent réellement une, alors qu'elles sont elles-mêmes déjà sous-représentées dans le portefeuille. La variable "présence d'une dépendance" paraît alors peu intéressante. Nous ne possédons pas d'informations sur les caves ou les garages indépendants, qui pourraient s'apparenter à des dépendances et qui cette fois s'adresseraient aux appartements. On notera toutefois que lors d'une inondation d'un sous-sol d'un immeuble par exemple, ce sera l'assurance auto qui jouera en cas de dommage de la voiture. Finalement, seulement 1% de nos expositions possèdent une ou plusieurs dépendances déclarées.

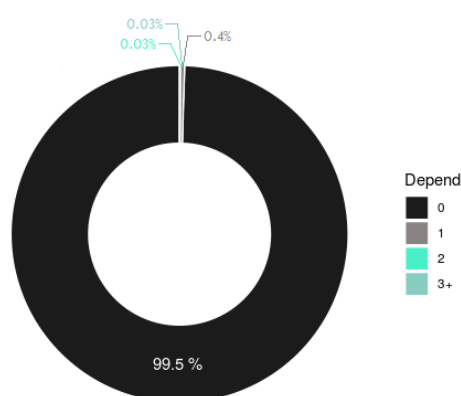


FIGURE 2.19 – Parts d'exposition selon la présence de dépendances dans l'habitation.

- **Année de construction** : Dans notre portefeuille, près de la moitié des habitations datent d'avant 1945. Deux tiers datent d'avant 1969. Moins de 4% datent du XXIème siècle. Avec environ un bien sur huit, la part de valeurs manquantes dans notre base est non négligeable.

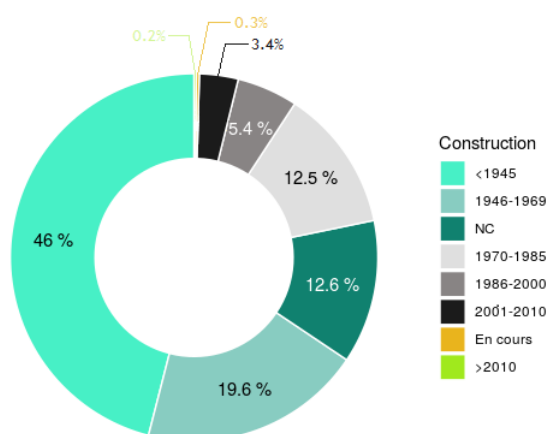


FIGURE 2.20 – Parts d'exposition selon l'année de construction de l'habitation.

- **Superficie** : En dehors du fait que l'on ne connaît pas les dimensions d'une habitation sur quatre, on remarque un équilibre de classes pour les propriétés d'une superficie inférieure à 75m². Ensuite, plus la surface grandit, moins l'habitation est représentée dans la base.

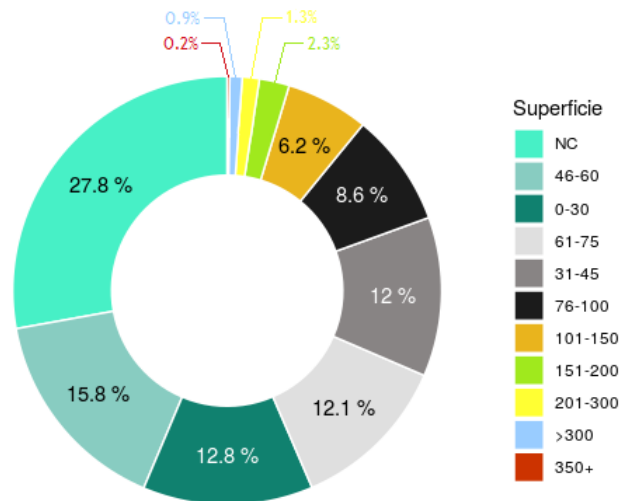


FIGURE 2.21 – Parts d'exposition selon la superficie de l'habitation.

- **Nombre de pièces** : La tendance pour le nombre de pièces est similaire à celle de la superficie. Hormis les appartements d'une seule pièce qui se classent troisième par ordre décroissant de part d'exposition, on observe une corrélation négative totale entre la part d'exposition et le nombre de pièces.

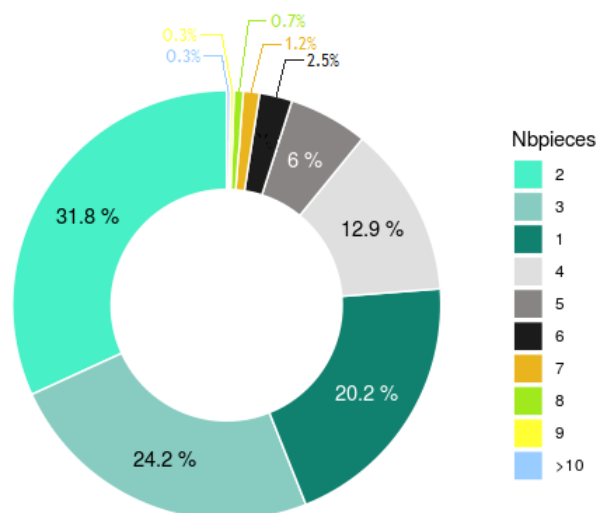


FIGURE 2.22 – Parts d'exposition selon le nombre de pièces de l'habitation.

- **Latitude/Longitude** : Les variables géographiques qui nous intéressent sont la longitude et la latitude car nous travaillons à la maille de l'adresse mais on ne peut pas représenter une maille aussi fine. On se propose alors de regarder la part d'exposition par département. Sans surprise, le 75 est majoritaire et les départements 95 et 78 sont très peu présents. Cela provient du découpage du bassin réalisé. Seule une très petite partie de ces départements est captée dans cette zone, d'où la faible proportion d'expositions provenant de ces derniers.

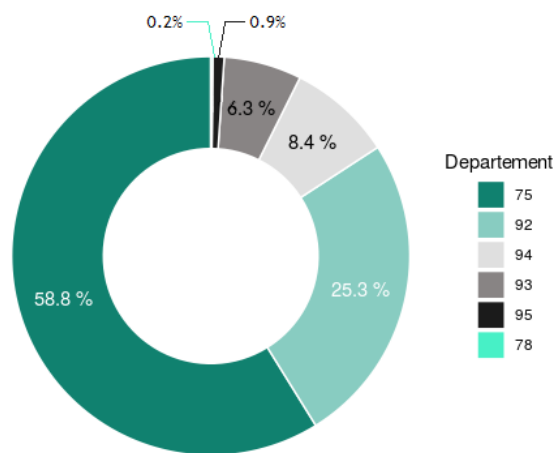


FIGURE 2.23 – Parts d'exposition par département.

- **Capital assuré** : Nous avons représenté un graphique permettant de mettre en lumière la proportion d'exposition par capital assuré croissant. On peut voir que la répartition des sommes assurées est concentrée à 95% en dessous de 100 000€.

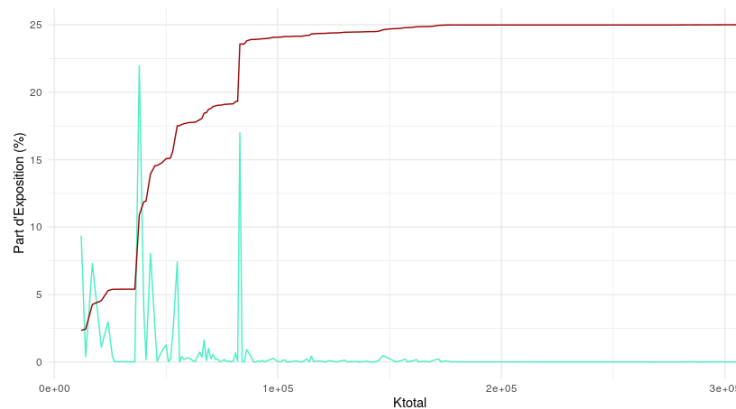


FIGURE 2.24 – Parts d'exposition selon le capital assuré.

2.4.3 Mailles d'exposition

Exposition brute

Afin de réaliser une cartographie des expositions et d'obtenir une idée des zones à risques, on peut opter pour différentes mailles d'observation. Pour commencer, on peut tracer les expositions brutes individuelles :

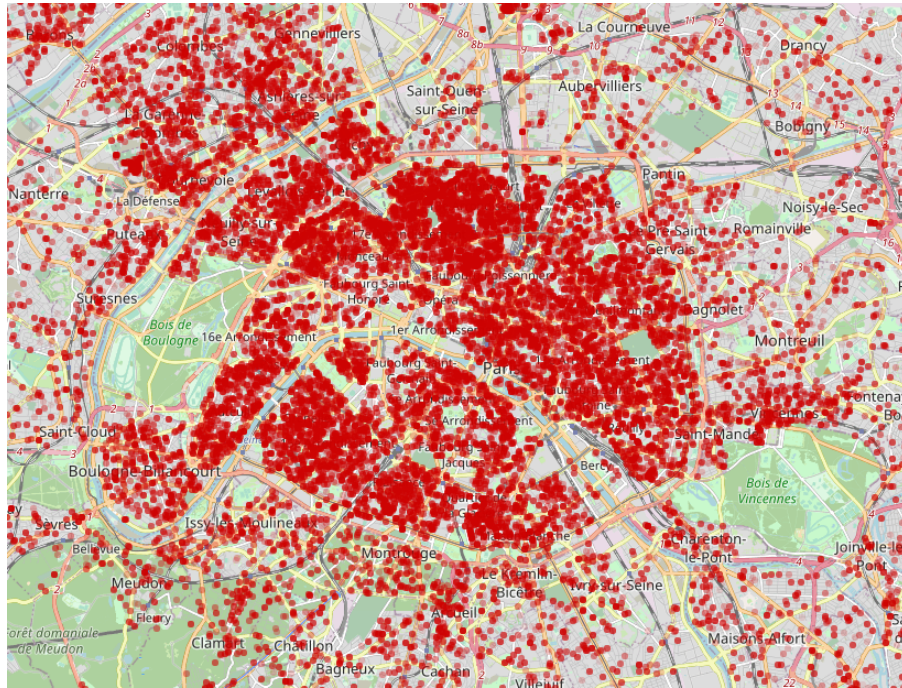


FIGURE 2.25 – Expositions sur le bassin parisien.

On observe que certaines zones possèdent une exposition très dense. Sans surprise, on ne repère aucune exposition dans les zones florales et fluviales. Cet aperçu est intéressant mais n'offre pas de vision permettant de classer les zones plus à risques que d'autres car trop peu visibles. Nous avons alors effectué plusieurs mailles d'exposition permettant de mieux mettre en lumière les concentrations locales d'exposition.

Maille Copernic

La première maille réalisée est la maille "Copernic". L'idée est de reprendre la résolution de la base de données *Copernic* et d'agréger les expositions selon cette maille. Cette agrégation permet de plus de transparaître directement la part totale d'exposition dans et hors de l'empreinte de crue. Pour ce faire, nous avons utilisé la méthode du plus proche voisin. Dans une grille rectangle, on sait que n'importe quel point du plan sera affilié à la cellule de laquelle il sera le plus proche de son centre. Les centres représentent ici les coordonnées latitude/longitude des cellules de la maille de la base de données *Copernic*.

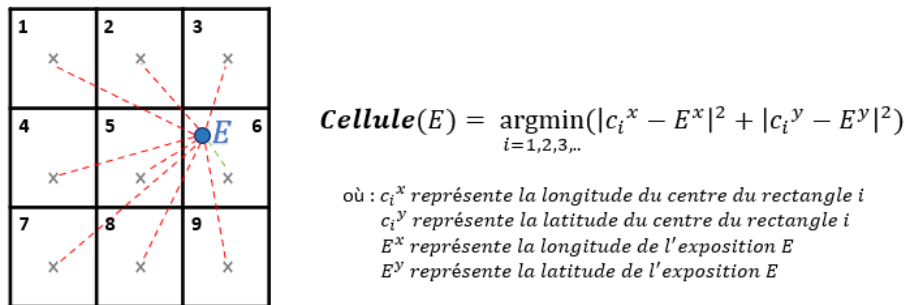


FIGURE 2.26 – Agrégation des expositions : Méthode du plus proche voisin (Rectangles).

Une fois l'agrégation des expositions effectuée, on peut représenter nos expositions par pixel de la maille "Copernic" :

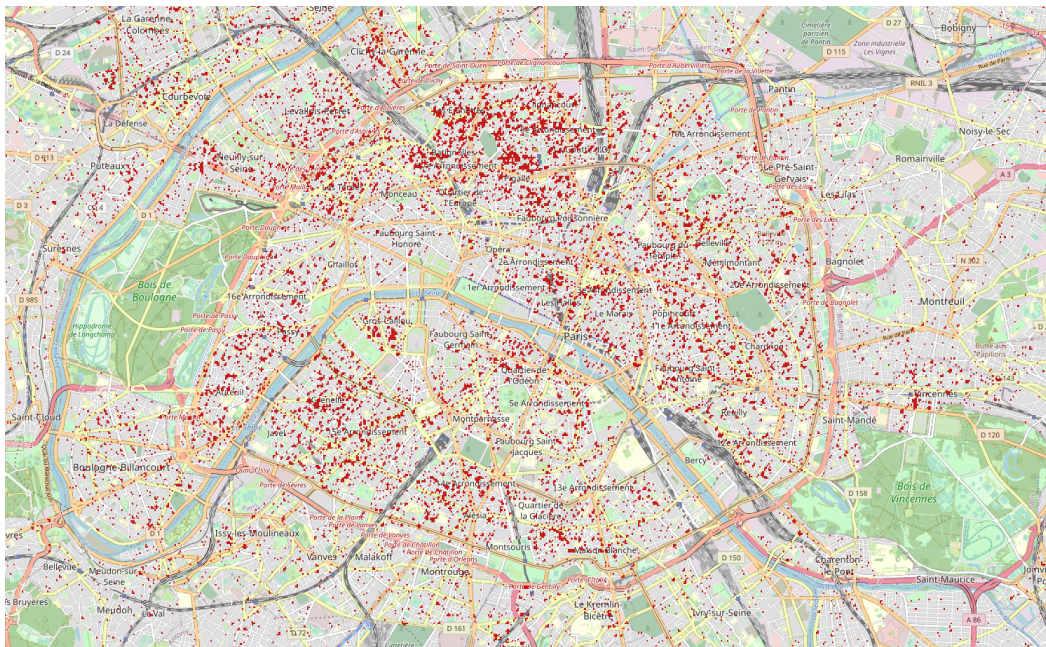


FIGURE 2.27 – Expositions sur le bassin parisien (Maille "Copernic").

Cette vision sera utile pour le calcul d'expositions touchées mais demeure encore une fois trop fine pour pouvoir observer de façon claire les zones à risques. Afin de réellement appréhender ces zones, il est nécessaire de zoomer sur différentes parties du bassin. Voici un zoom au niveau du pont d'Austerlitz mettant en lumière plusieurs secteurs d'expositions :

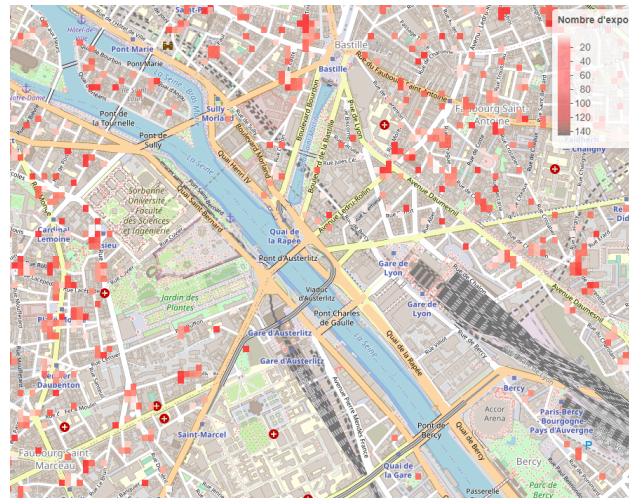
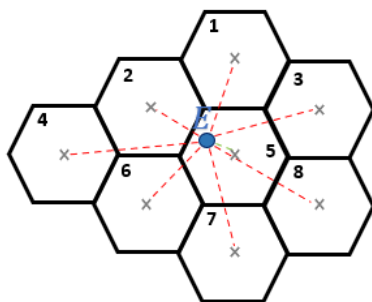


FIGURE 2.28 – Expositions sur le bassin parisien (Maille "Copernic") - Zoom.

Nous allons alors nous remettre à une agrégation plus macroscopique.

Maille Hexagonale

La dernière agrégation que nous avons réalisée est une maille hexagonale. Le bassin est alors scindé en une grille constituée d'une centaine d'hexagones, tous de la même taille (200 mètres de largeur). Pour cela, nous avons créé cette grille à l'aide d'un générateur de GeoJson. Ensuite, nous avons agrégé chacune des expositions avec la même méthode que pour la maille "Copernic".



$$Cellule(E) = \underset{i=1,2,3,..}{\operatorname{argmin}}(|c_i^x - E^x|^2 + |c_i^y - E^y|^2)$$

où : c_i^x représente la longitude du centre de l'hexagone i
 c_i^y représente la latitude du centre de l'hexagone i
 E^x représente la longitude de l'exposition E
 E^y représente la latitude de l'exposition E

FIGURE 2.29 – Agrégation des expositions : Méthode du plus proche voisin (Hexagones).

Cette agrégation permet d'observer les zones à risques dans Paris et de s'intéresser aux zones proches du fleuve. On peut imaginer que ces zones hexagonales seront touchées relativement identiquement en cas de crue car leur maille reste assez fine, cependant on verra que ce n'est pas forcément le cas. Cette maille reste donc utile surtout pour une cartographie plus lisible. Voici les résultats obtenus :

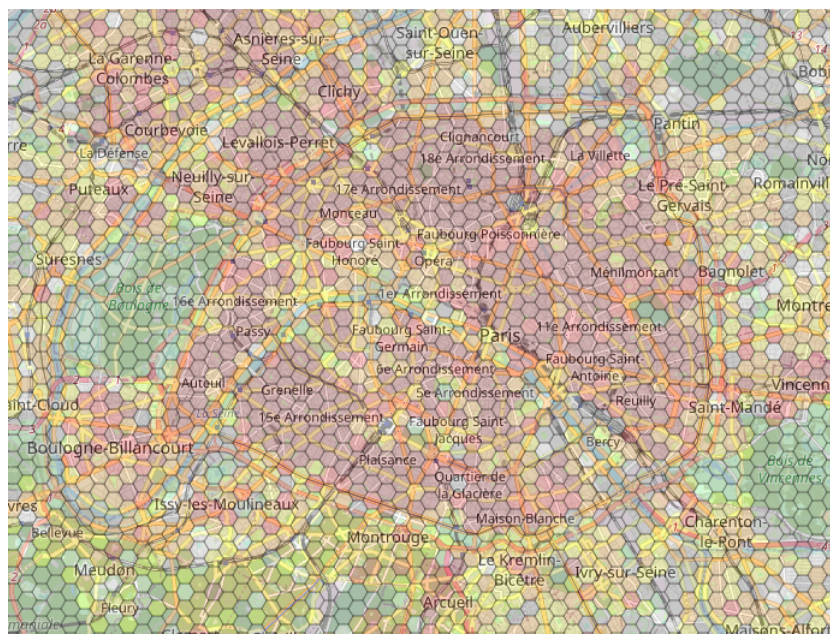


FIGURE 2.30 – Expositions sur le bassin parisien (Maille hexagonale).

2.4.4 Cartographie du risque et taux de sinistralité

Maintenant que nous avons réalisé notre cartographie d’empreintes de crues, de topographie du bassin ainsi que d’exposition, nous possédons tous les éléments pour cartographier le risque. Voici les trois étapes à suivre :

- **Évènement** : Pour une intensité donnée, on trace la carte d’empreinte de crue ;
- **Expositions** : On mappe les expositions avec la maille "Copernic" de sorte à affecter pour chaque zone touchée les expositions présentes à l’intérieur ;
- **Risque** : On calcule la part d’exposition dans l’empreinte de crue et donc potentiellement impactée.

Nous pourrons ensuite cartographier ces risques et reproduire des crues historiques, telles que celle de 1910 ou celle de 2016 par exemple. A partir de ce mapping des expositions et de l’empreinte de crue, nous avons pu tracer le graphique présentant la part d’exposition dans l’empreinte de crue en fonction de la hauteur d’eau relevée à la station Austerlitz. Pour cela, nous reprenons nos expositions agrégées avec la maille "Copernic".

On remarquera que jusqu’alors, l’exposition considérée était le nombre de risques (contrats). Dans le but d’analyser la dépendance de la modélisation à la définition de l’exposition, nous avons tracé ce graphique pour trois mesures différentes de l’exposition :

- **Exposition contractuelle** : Elle est comprise entre 0 et 1 par contrat et représente la part de l’année civile qui est couverte par le contrat. Ainsi, un contrat courant du 1er avril au 31 décembre aura une exposition de 0.75, soit trois trimestres sur les quatre de l’année ;
- **Nombre de risques** : Cette définition revient à simplement compter le nombre

de contrats, il correspond au nombre de lignes dans notre base contrats ;

- **Somme assurée** : C'est le capital assuré pour chacun des contrats de notre base. Il représente alors une borne supérieure des pertes potentielles brutes. C'est cette définition qui permettra ensuite de construire notre approche par taux de destruction.

Nous présentons ci-dessous les résultats des parts d'exposition touchées selon la hauteur d'eau relevée à la station d'Austerlitz.

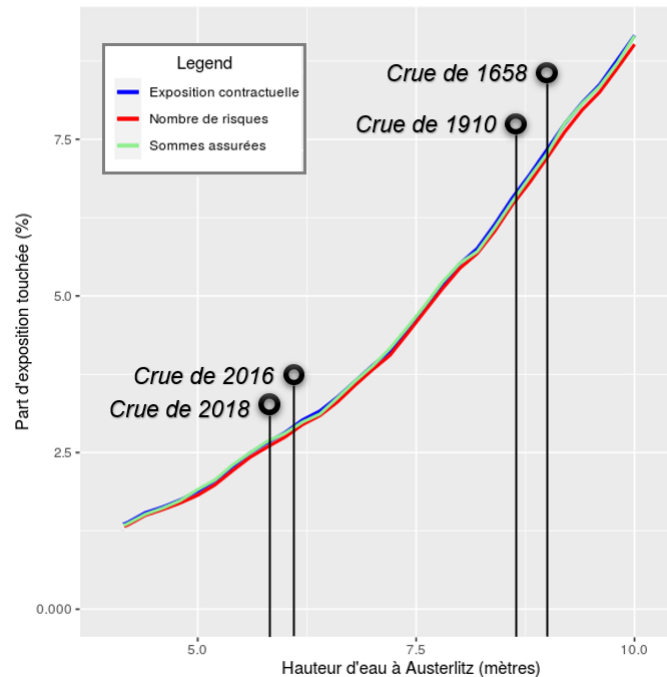


FIGURE 2.31 – Pourcentage d'exposition dans l'empreinte de crue selon la hauteur d'eau.

A partir de cette part d'expositions touchées, on peut estimer la population totale qui pourrait être touchée par une crue de type 1910. On peut ensuite comparer ces chiffres avec ceux de l'OCDE que l'on a présenté plus tôt. Notre part d'exposition pour une hauteur d'eau de 8,62 mètres est d'environ 7% pour le bassin parisien tel que nous l'avons défini. Ainsi, si l'on reporte ce chiffre aux quelques 1 386 846 logements de la capitale (INSEE, 2018), on trouve environ 97 000 logements impactés. L'OCDE prévoyait quant à elle 240 000 biens hypothétiquement touchés. Cependant, cette différence s'explique par différents facteurs. Premièrement, rien ne dit que notre portefeuille est représentatif de l'exposition parisienne réelle. Comme nous l'avons précisé plus tôt, nous possédons quelque 100 000 biens assurés dans notre zone d'étude, contre 1 386 846 logements dans Paris et 5 846 882 en Ile de France, notre base n'est donc logiquement pas exhaustive. En outre, nous avons considéré un bassin parisien à partir d'une portion de l'Ile-de-France centrée sur Paris et sur le fleuve et ayant une densité géographique différente à celle de Paris ou de l'Ile-de-France que nous utilisons pour notre calcul proportionnel. De

plus, si nous avons pris une zone différente, ces chiffres auraient changé. Par exemple, en étendant cette zone vers des zones éloignées nous aurions diminué artificiellement la part d'expositions touchées, et inversement en la concentrant sur les bords de Seine. Ces chiffres donnent donc une tendance pour un portefeuille donné, ce qui nous intéresse dans un cadre actuariel.

Notre module d'exposition nous permet maintenant d'associer à chaque intensité de crue une part d'expositions touchées. En cas de montée des eaux, nous connaissons en temps réel les biens impactés et pouvant alors calculer ensuite des pertes potentielles. Nous allons dans un premier temps présenter la cartographie de crues de 7, 8 puis 9 mètres en indiquant la part d'expositions touchées (nombre de contrats).



FIGURE 2.32 – Cartographie d'une crue de 7 mètres et part d'expositions touchées.



FIGURE 2.33 – Cartographie d'une crue de 9 mètres et expositions touchées .

Comme nous venons de le voir, nous pouvons à partir d'une crue de n'importe quelle

intensité déterminer la part d'expositions touchées. En plus de connaître ce chiffre, nous savons précisément quelles expositions sont sous l'eau et quelles sommes assurées sont en jeu. Il peut alors être intéressant de représenter les expositions touchées, ce qui permet d'observer les zones à risques et de pointer les expositions assurées en zone inondable. Afin de faire le parallèle avec des événements passés, nous présentons ci-dessous les cartes des biens assurés qui seraient impactées par les crues historiques de 2016 dans un premier temps puis ensuite celle de 1910, ainsi que les statistiques de sinistralité.

Crue de type 2016 - Maille à l'adresse

- Nombre d'expositions touchées : **3 190**
- Part d'expositions touchées : **2,91%**
- Sommes assurées touchées : **156 782 000**
- Part des sommes assurées touchées : **2,86%**

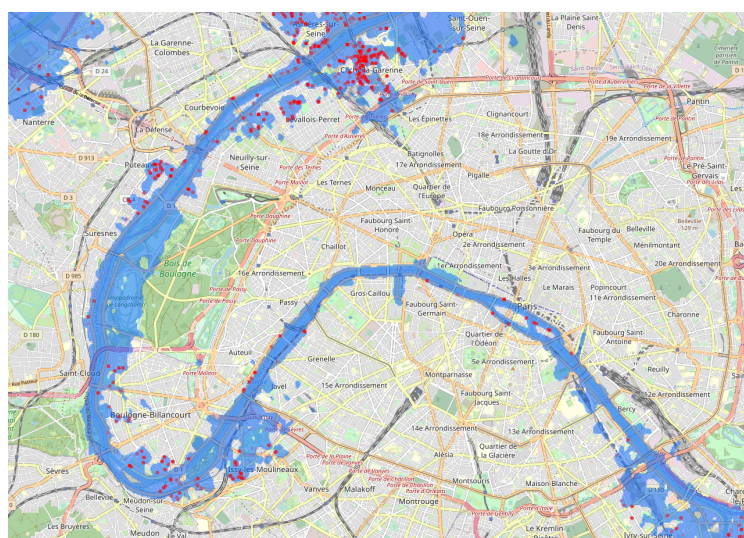


FIGURE 2.34 – Cartographie des expositions dans l'empreinte de crue de 2016 (Maille à l'adresse).

Crue de type 2016 - Maille hexagonale

- Nombre d'expositions touchées : **4 210**
- Part d'expositions touchées : **3,81%**
- Sommes assurées touchées : **211 655 700**
- Part des sommes assurées touchées : **3,85%**



FIGURE 2.35 – Cartographie des expositions dans l’empreinte de crue de 2016 (Maille hexagonale).

Crue de type 1910 - Maille à l’adresse

- Nombre d’expositions touchées : **7 327**
- Part d’expositions touchées : **6,71%**
- Sommes assurées touchées : **375 293 000**
- Part des sommes assurées touchées : **6,87%**

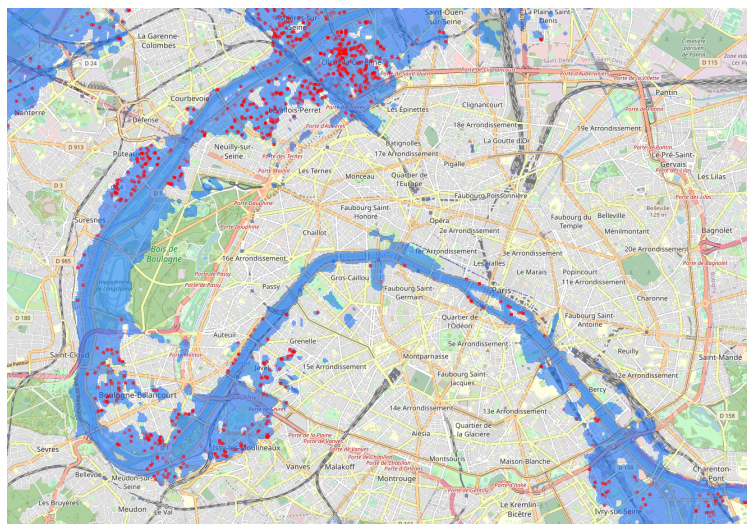


FIGURE 2.36 – Cartographie des expositions dans l’empreinte de crue de 1910 (Maille à l’adresse).

Crue de type 1910 - Maille hexagonale

- Nombre d'expositions touchées : **9 525**
- Part d'expositions touchées : **8,72%**
- Sommes assurées touchées : **484 128 000**
- Part des sommes assurées touchées : **8,87%**

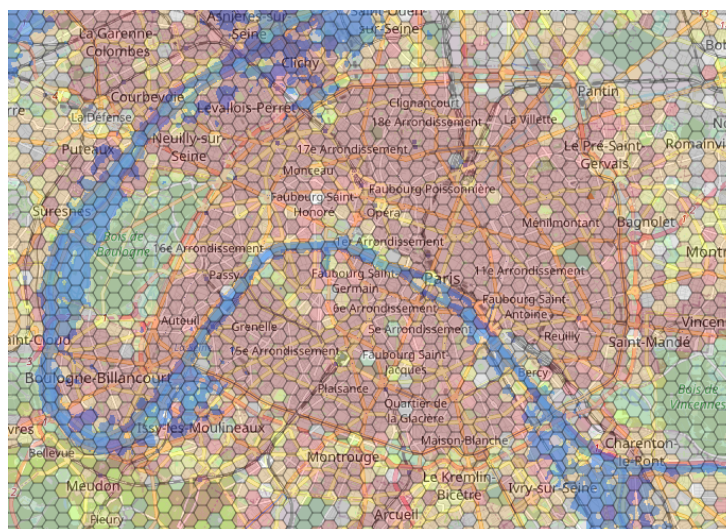


FIGURE 2.37 – Cartographie des expositions dans l'empreinte de crue de 1910 (Maille hexagonale).

Afin de pousser plus loin nos études, nous pouvons étudier la part d'expositions touchées (sommes assurées) en fonction de la hauteur d'eau et par caractéristiques du bien assuré. Par exemple, nous pouvons regarder la différence entre les appartements et les maisons :

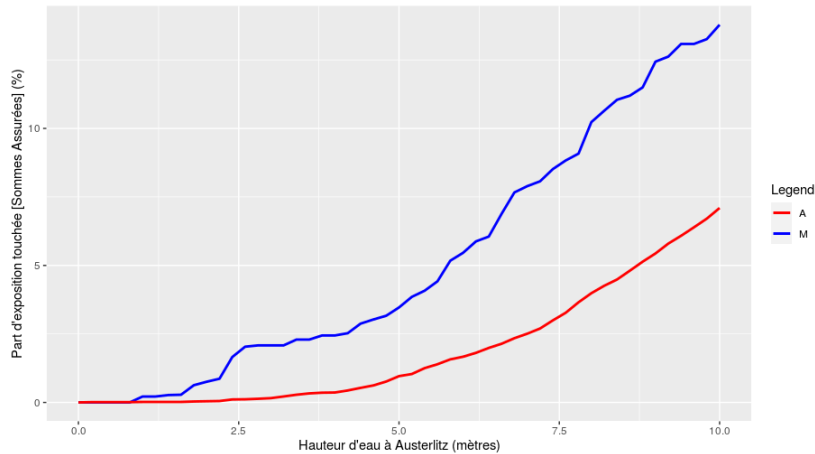


FIGURE 2.38 – Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (Appartement ou Maison).

On observe que la part d'expositions touchées des maisons croît plus rapidement que celle des appartements. Cela signifie que la part des maisons de notre portefeuille en zone à risque par rapport aux maisons en zones sécurisées est plus élevée que pour les appartements. Nous pouvons encore préciser cette analyse avec la décomposition des appartements par étage :

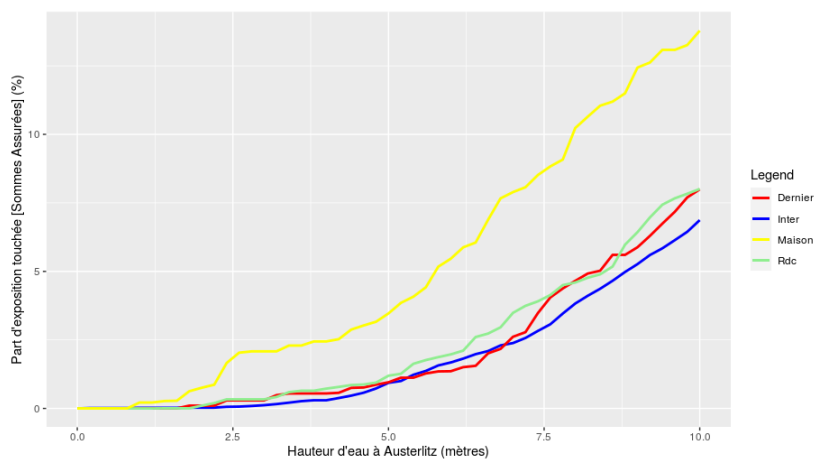


FIGURE 2.39 – Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (selon étage).

Les comportements des expositions touchées sont relativement similaires pour tous les types d'appartements. On notera tout de même que dans notre base, ce sont les appartements au rez-de-chaussée qui sont légèrement plus impactés proportionnellement pour presque toutes les hauteurs d'eau. Ce résultat est intéressant car ce sont les expositions les plus à risques parmi les appartements.

On s'intéresse maintenant à l'année de construction du bâtiment. Voici les résultats que nous avons obtenus :

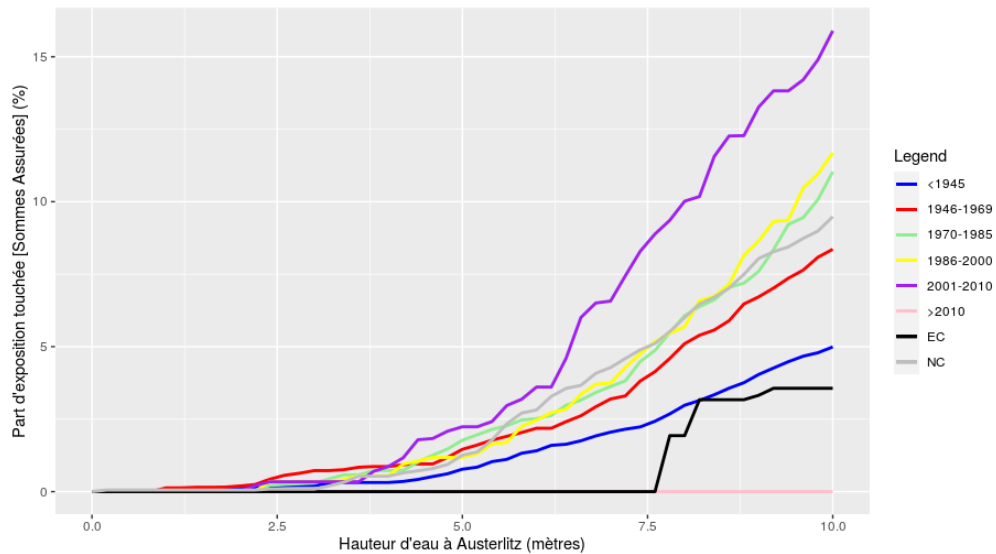


FIGURE 2.40 – Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (selon année de construction).

On remarque alors une surreprésentation des expositions touchées pour les biens les plus récents. Cela confirme la tendance exposée par le gouvernement signifiant que l'on continue encore depuis des années à construire en zone inondable malgré notre connaissance accrue sur le risque de débordement de la Seine. On peut même dire avec ces chiffres que l'on construit proportionnellement de plus en plus en zone inondable.

Enfin, il reste l'étude des parts d'exposition touchées par nombre de pièces. Nous avons placé le graphique correspondant en Annexe A car il nous paraissait peu intéressant. On notera tout de même que les biens les plus vulnérables sont plutôt ceux de taille intermédiaire (trois, quatre ou cinq pièces) qui sont proportionnellement plus impactés que les très petits biens (souvent plus loin du centre de Paris) et les très grands biens (souvent carrément hors de Paris).

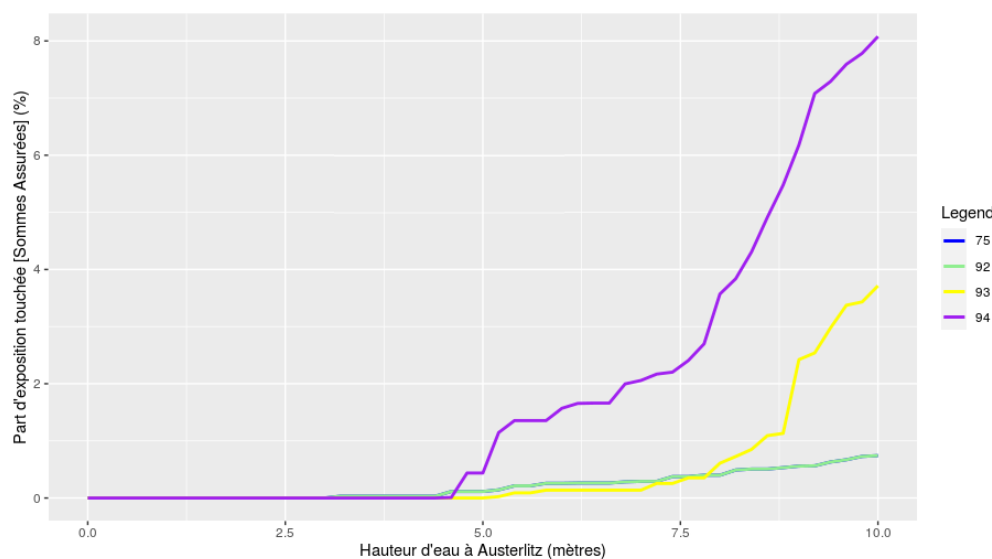


FIGURE 2.41 – Analyse bivariable : Part d'expositions touchées selon hauteur d'eau (selon département).

2.5 Évaluation de la Vulnérabilité

2.5.1 Méthodologie

Dans un modèle classique d'estimation de la fréquence de sinistralité, on se base sur l'historique et on essaie de déterminer la répartition de cette fréquence. Cependant, pour un modèle CAT NAT d'inondation la méthode est différente. En effet, la localisation est un paramètre indispensable et témoin de l'apparition ou non d'un sinistre. Ainsi, la sinistralité et la localisation sont deux variables totalement dépendantes et lorsque l'on connaît la localisation du bien assuré et les zones touchées par l'inondation, on connaît nos polices théoriquement sinistrées. Cette méthode est subordonnée à la cartographie de la topographie. Si cette dernière possède un biais à certains endroits, alors on peut capter du risque superflu ou inversement. Une seconde limite est de ne pas capter certaines fraudes à l'assurance. En effet, dans le cadre de catastrophes naturelles, le nombre conséquent de sinistres empêche l'envoi systématique d'experts pour l'évaluation des sinistres, et l'état de CAT NAT étant décliné à la maille communale, certaines zones géographiques peuvent se voir attribuer l'état de catastrophe naturelle à tort. Cependant, la considération qu'une exposition est systématiquement touchée (à des degrés divers) lorsqu'elle est dans l'empreinte reste raisonnable. En effet, cette méthodologie reste bien plus précise que l'application de simples proportions moyennes d'expositions touchées sur le portefeuille entier.

L'historique des sinistres servira alors à construire nos taux de destruction pour la partie "coût" mais la partie "sinistralité" est déterminée à l'aide de : la dynamique de la crue ainsi que la localisation et l'altitude de l'exposition. Ce sont les trois facteurs qui vont

déterminer si l'exposition sera touchée. Il faut ensuite définir à quel point l'exposition sera touchée. C'est le module vulnérabilité qui répondra à cette problématique. C'est alors qu'interviendront les autres variables telles que l'étage (un appartement au rez-de-chaussée ou une maison seront directement impactés, un appartement au 1er étage le sera beaucoup moins, seuls les parties communes, la cave ou le garage peuvent être touchés par exemple). Nous verrons tout cela par la suite.

2.5.2 Analyse de la base de données sinistres MRH

Nous possédons une base de données de sinistres inondation d'un assureur français. On rappelle que nous nous intéressons au risque de destruction du bâtiment et de son contenu. Nous ne nous intéressons pas à la sinistralité auto par exemple car nous ne possédons pas de données correspondantes. Nous allons alors calibrer nos taux de destruction moyens en fonction des caractéristiques de l'habitation assurée et de la dynamique de crue. Cependant, par souci de confidentialité nous avons retraité ces données afin de déformer la répartition et les montants de ces sinistres.

La première variable influant sur le taux de destruction est le type d'habitation (appartement/maison). Nous représentons ci-dessous la somme assurée moyenne ainsi que le coût moyen des sinistres (volontairement sans échelle) :

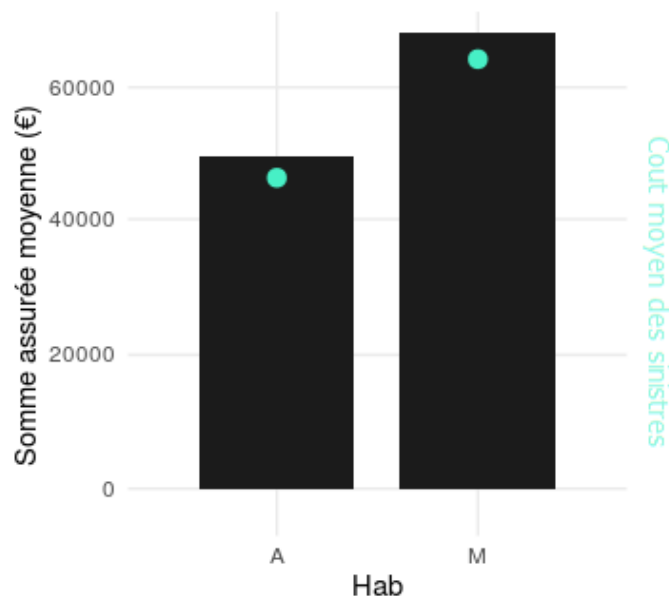


FIGURE 2.42 – Sommes assurées et coût total des sinistres par type d'habitation.

Comme nous pouvions nous y attendre, un sinistre sur un appartement coûte moins cher qu'un sinistre sur une maison. Le fait que des appartements puissent ne pas être de plein pied en est l'explication principale.

On peut aller plus loin en regardant par étage directement. La tendance décrite plus tôt se confirme bien. Plus le bien assuré est à un étage élevé, moins il est touché.

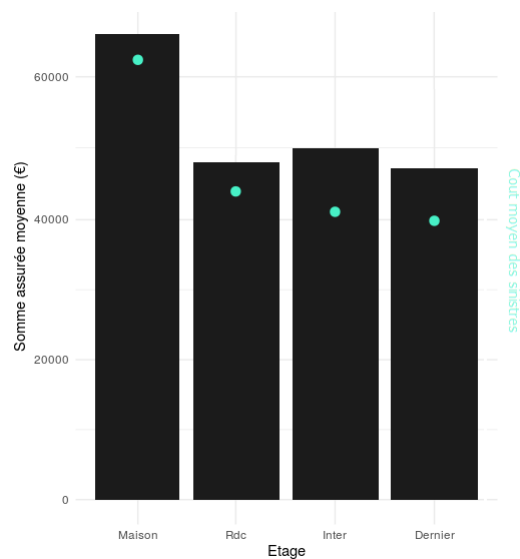


FIGURE 2.43 – Sommes assurées et coût total des sinistres par étage.

Intéressons-nous maintenant à l'année de construction des logements. On s'attend à avoir une corrélation négative entre la vulnérabilité du bien et son année de construction.

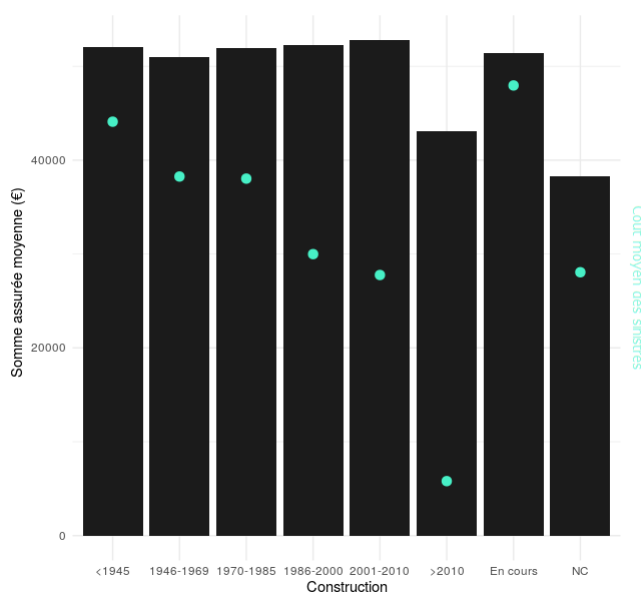


FIGURE 2.44 – Sommes assurées et coût total des sinistres par année de construction.

En effet, on peut observer qu'au plus le bien est récent, au plus le taux de destruction au global (*moyenne des sommes assurées / moyenne des coûts de sinistres*) est faible. On remarque que les logements les plus anciens (avant 1945) sont les plus vulnérables.

Pour terminer, nous allons regarder ces mêmes graphiques pour la superficie et le nombre de pièces. On devrait observer un comportement similaire entre ces deux variables, comme c'était le cas dans la base contrat. En effet, la superficie d'un logement et son nombre de pièces sont évidemment corrélés.

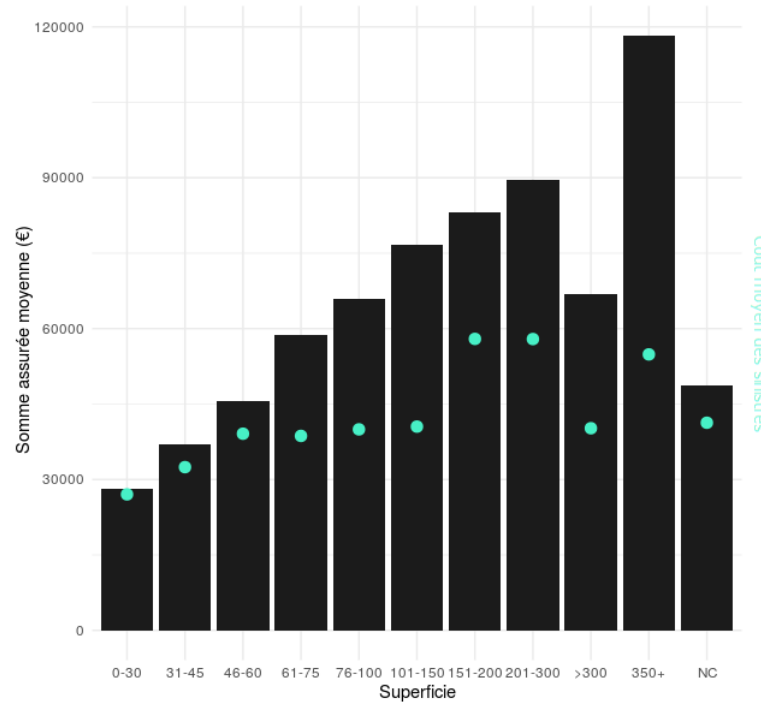


FIGURE 2.45 – Sommes assurées et coût total des sinistres par superficie.

Cette fois il est plus compliqué de dégager une tendance claire. Les biens plus grands ont tout de même tendance à posséder un taux de destruction global moyen plus élevé. Cependant, la part de valeurs manquantes est conséquente et leur taux de destruction est élevé également. Nous avons donc fait le choix de considérer la superficie comme indépendante de la destruction du bien, et cela paraît être une hypothèse vraisemblable dans la réalité. Nous observons la même tendance pour le nombre de pièces, dont le graphique est en Annexe A.

2.5.3 Calibrage des taux de destruction

Afin de pouvoir déterminer selon une crue donnée les pertes pour l'assureur, il nous faut définir des taux de destruction moyens. Un taux de destruction est défini comme tel :

$$\text{Taux de destruction} = \frac{\text{Coût du sinistre}}{\text{Somme Assurée}} \quad (2.1)$$

Cependant les taux de destruction sont différents pour chaque sinistre et possèdent une volatilité intrinsèque impossible à déterminer. Généralement, on se base alors sur des

taux de destruction moyens que l'on applique à nos expositions touchées. Ils sont évidemment variables selon l'intensité de la catastrophe, c'est pour ça que nous construisons des courbes de vulnérabilité comme nous l'avons présenté dans le chapitre précédent. Un point d'attention particulier est que l'axe des abscisses de ces courbes (intensité de la catastrophe i.e. hauteur d'eau notamment) n'est plus identique aux précédents graphiques présentés. En effet, la hauteur d'eau considérée jusqu'alors était celle sur l'échelle d'Austerlitz. Ici, nous nous intéressons à la hauteur d'eau supérieure au sol du bâtiment de notre exposition. Ainsi, une hauteur d'eau d'un mètre signifiera que le bâtiment est inondé avec 1 mètre d'eau. Nous calculons alors cette valeur à partir de la hauteur d'eau à l'échelle d'Austerlitz et de l'altitude de l'exposition :

$$HE^5 \text{ dans un bâtiment} = HE \text{ à l'échelle d'Austerlitz} - \text{Altitude de l'exposition} \quad (2.2)$$

Comme nous l'avons vu plus tôt, l'année de construction du bien assuré influe grandement sur les pertes potentielles. Plus un bâtiment est récent, moins il est vulnérable. Nous avons donc décidé de conditionner nos taux de destruction en fonction de l'année de construction du bien assuré, pour avoir une meilleure représentation des pertes réelles. Nous avons effectué de même avec la variable habitation car une maison et un appartement ne seront pas touchés de la même façon. Pour un appartement, l'étage est un paramètre primordial. Dans le cas d'un appartement au rez-de-chaussée, le taux de destruction varie selon une hauteur d'eau comprise entre 0 (minimum) et la hauteur standard d'un plafond -à savoir 2,50m- (maximum). Au-dessus de 2,50m, le taux de destruction est plafonné à 1. Pour un appartement à étage intermédiaire, nous considérerons que dès lors que le bâtiment est touché, le bien est impacté avec un taux de destruction constant non nul (correspondant aux dégâts potentiels dans les caves/parkings) tant que la hauteur d'immersion n'a pas atteint l'étage. Ensuite, dès lors que l'eau atteint le sol de l'appartement, le taux de destruction varie de façon similaire à un appartement au rez-de-chaussée entre 0 et 2,50m.

Comme nous ne possédons pas beaucoup d'historiques d'inondations, nous sommes obligés d'effectuer une méthode d'interpolation sur des taux de destruction moyens regroupés par intervalles de 0,5m. Nous présentons ci-dessous les résultats par catégories.

Taux de destruction : Maisons

Pour commencer, nous avons évalué nos taux de destruction pour les maisons. Celui-ci varie entre 0% à 0 mètre d'eau et 100% pour 5 mètres d'eau (maison totalement immergée du rez-de-chaussée à l'étage) sous l'hypothèse qu'une maison possède généralement 1 étage. Cette approximation n'est pas un problème étant donné qu'un tel niveau d'immersion n'est presque jamais atteint. Pour calibrer nos taux de destruction, nous avons étudié nos sinistres inondation pour les maisons assurées. Il en ressort un taux de destruction d'environ 10% pour les premiers décimètres d'eau, qui sont les plus impactant (sol, plaintes et les meubles touchés de plein fouet). Ensuite, la progression des dégâts est de

5. HE=Hauteur d'eau

plus en plus lente, jusqu'à atteindre plus de 50% aux alentours d'1 mètre d'eau. Nous ne possédons pas de données sur les hauteurs d'eau supérieures, nous avons donc extrapolé la courbe afin de la prolonger jusqu'au point (5m;100%). Afin d'obtenir une courbe lisse, nous avons finalement décidé d'utiliser une courbe de tendance logarithmique du fait de la forme de la courbe, très pentue pour les premiers centimètres d'eau et plus linéaire ensuite.

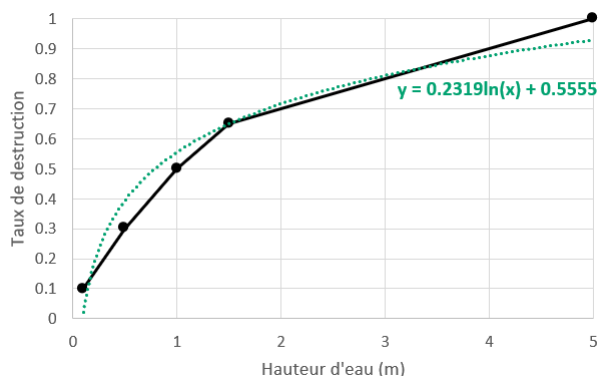


FIGURE 2.46 – Courbe de vulnérabilité pour une maison quelconque.

Nous possédons trop peu de données pour créer des courbes de vulnérabilité différentes selon l'année de construction, nous avons utilisé des approximations. Ainsi, nous avons fait le choix d'inflater de 5% supplémentaire sur base fixe de (5m,100%) le taux de destruction lorsque l'on descend à la classe d'ancienneté inférieure à partir de la classe 2000-2010.

Une autre variable influant sur la vulnérabilité d'un habitat est la durée d'immersion. Devant l'impossibilité d'effectuer des courbes de vulnérabilité avec nos historiques de données, nous avons utilisé des avis d'experts. Ainsi, selon un rapport du CEPRI (Centre de Prévention du Risque Inondation) [CEPRI, 2010], un bâtiment a plus de deux fois plus de chance d'être endommagé si sa durée d'immersion est comprise entre deux et trois jours par rapport à une immersion courte (moins d'une journée). Nous avons donc effectué une seconde inflation conditionnelle à la durée d'immersion du bien assuré. Ainsi, notre taux de destruction dépend non seulement de la hauteur d'eau atteinte par la crue, mais également de la durée de la crue. Pour résumer, nous avons défini les règles suivantes :

- Inflation de 5% supplémentaire par saut de classe d'ancienneté de construction à partir de la classe 2000-2010 sur la base fixe de (5m;100%)⁶ ;
- Inflation de 50% par saut de classe de jour supplémentaire d'immersion (inférieure ou égale à deux jours, entre deux et quatre jours, supérieure ou égale à quatre jours).

Hauteur d'eau maximale atteinte et altitude de l'exposition, durée d'immersion, type d'habitation, année de construction du bien assuré, voici les paramètres qui définissent un

6. Pour les biens à année de construction non connue, nous avons utilisé les taux de destruction les plus élevés par principe de prudence.

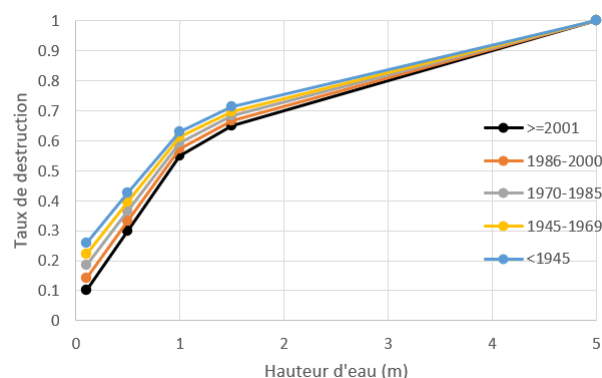


FIGURE 2.47 – Courbe univariée de vulnérabilité pour les maisons selon l'année de construction (durée d'immersion fixée à " $\leq 2j$ ").

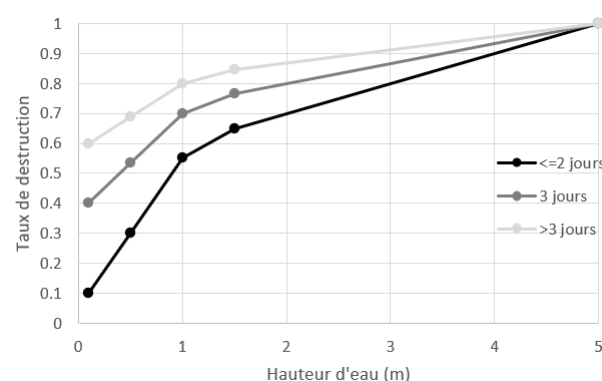


FIGURE 2.48 – Courbe univariée de vulnérabilité pour les maisons selon la durée d'immersion (durée d'immersion fixée à " $\leq 2j$ ").

taux de destruction spécifique pour chaque exposition, permettant de refléter au mieux les dégâts réels lors de la survenance d'une catastrophe. On présente ci-dessous un tableau récapitulatif pour trois hauteurs d'immersion données

Il est impossible de présenter toutes les courbes de destruction sur un même graphique du fait du trop grand nombre de paramètres influant sur ce dernier. Nous pouvons cependant tracer un graphique en trois dimensions en fixant deux paramètres. Voici un exemple de représentation 3D des taux de destruction en fonction de la hauteur d'immersion et de l'année de construction du bien assuré pour une maison avec une durée d'immersion de trois jours.

Nous pouvons également représenter quelques taux de destruction dans un tableau de vulnérabilité résumant les taux de destruction classiques (pour 0,1m, 0,5m, 1m et 1,5m)

8. L'année de construction est une variable qualitative avec cinq classes, la représentation sous forme d'un plan est purement visuelle. Il s'agit d'une simple régression linéaire point par point entre les cinq courbes de destruction, mais ces valeurs n'ont pas de sens théoriquement parlant.

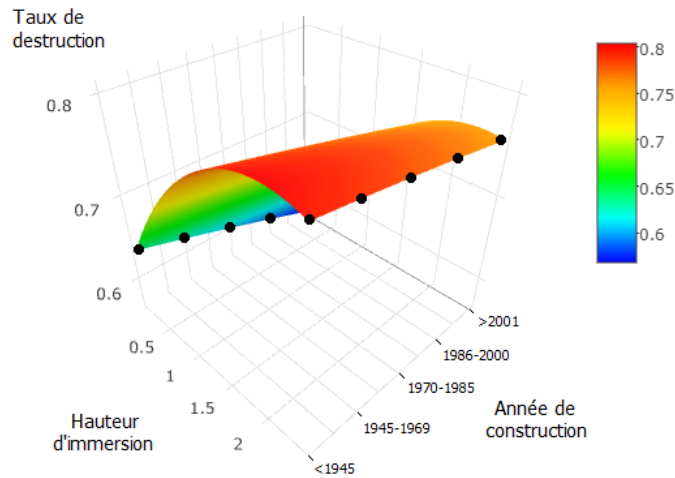


FIGURE 2.49 – Taux de destruction selon la hauteur d’immersion et l’année de construction du bien assuré (pour une maison et une durée d’immersion de trois jours) ⁸.

que nous considérerons pour les maisons :

Durée d'immersion ↘ Année de construction ↓	Hauteur d'immersion											
	0.1 mètre			0.5 mètre			1 mètre			1,5 mètres		
	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j	≤ 2 j	2 à 4 j	≥ 4 j
>=2001	0.32	0.55	0.70	0.49	0.66	0.77	0.56	0.70	0.80	0.60	0.73	0.82
1986-2000	0.36	0.57	0.71	0.51	0.67	0.78	0.58	0.72	0.81	0.62	0.74	0.83
1970-1985	0.39	0.59	0.73	0.53	0.69	0.79	0.60	0.73	0.82	0.63	0.76	0.84
1945-1969	0.42	0.61	0.74	0.56	0.70	0.80	0.62	0.74	0.83	0.65	0.77	0.85
<1945	0.44	0.63	0.75	0.58	0.72	0.81	0.63	0.76	0.84	0.67	0.78	0.85

FIGURE 2.50 – Taux de destruction retenus adaptés aux caractéristiques de la maison.

Selon le CEPRI : "Les retours d’expérience sur les effets des crues passées montrent que dès que la hauteur de l’eau dépasse celle de l’allège⁹, 70 à 80 % du potentiel d’endommagement du rez-de-chaussée est atteint.". Nous retrouvons alors des valeurs globalement cohérentes avec les dires du rapport, ce qui nous conforte sur le calibrage de ces taux.

Taux de destruction : Appartements

La même méthode a été appliquée pour les appartements. La différence réside dans l’étage du bien assuré. Comme nous avons pu le préciser plusieurs fois déjà, un appartement au 5ème étage ne sera que très modestement touché par le biais des biens communs et des sous-sols. Nous avons estimé cet impact à 5% en moyenne. Ainsi, nos taux de destruction pour un appartement à un étage sera fixé à 5% jusqu’à ce que l’eau monte au sol de l’étage en question. Pour un appartement au rez-de-chaussée, la courbe de destruction est proche de la courbe de destruction utilisée pour les maisons. Ci-dessous la

9. L’allège désigne sur un bâtiment la partie du mur située entre le plancher et l’appui de la fenêtre. En moyenne, elle est à environ 1 mètre du sol.

représentation de la courbe de destruction de base :

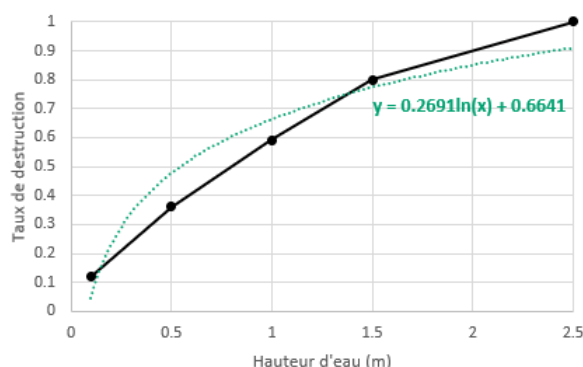


FIGURE 2.51 – Courbe de vulnérabilité pour un appartement au rez-de-chaussée.

Pour un appartement au 1er étage, nous avons donc fixé à 5% le taux de destruction pour une hauteur d'immersion sous 2,5 mètres, puis laggé la courbe de 2,5 mètres vers les valeurs positives. Comme nous n'avons pas le détail précis des étages, nous avons considéré les appartements de classe "Intermédiaire" comme au 1er étage et ceux de la classe "Dernier étage" comme au 2ème étage. Cela influera peu sur nos résultats car très peu d'immeubles seront immergés de deux mètres et plus.

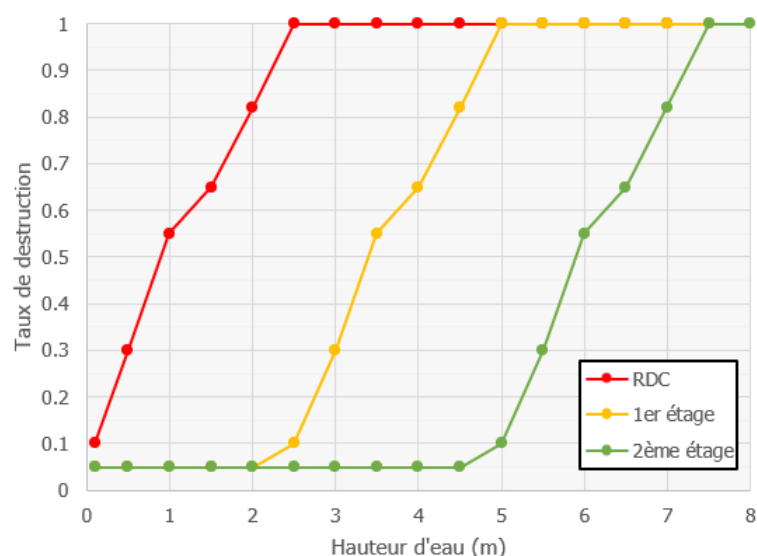


FIGURE 2.52 – Courbe de vulnérabilité pour un appartement selon l'étage.

Cela revient simplement à considérer les mêmes courbes mais avec l'abscisse laggée de 2,5 mètres ou de 5 mètres. Ainsi, on ne représente plus la hauteur d'eau à partir du niveau 0 de l'immeuble mais à partir du niveau 0 de l'appartement. Enfin, nous avons effectué les mêmes traitements que pour les maisons sur l'inflation causée par la

durée de l'immersion ou l'année de construction. Nous présentons ci-dessous des tableaux récapitulatifs des taux de destruction retenus selon les caractéristiques du bien, la durée d'immersion et la hauteur d'eau maximale de la crue pour les appartements au rez-de-chaussée dans un premier temps, au premier étage dans un second temps, et enfin au deuxième étage.

Appartement au RDC : Hauteur d'immersion (à partir du niveau 0 de l'immeuble)									
Durée d'immersion ↘ Année de construction ↓	0.5 mètre			1 mètre			1.5 mètres		
	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours
≥ 2001	0.58	0.72	0.81	0.66	0.78	0.85	0.71	0.81	0.87
1986-2000	0.60	0.74	0.82	0.68	0.79	0.86	0.73	0.82	0.88
1970-1985	0.62	0.75	0.83	0.70	0.80	0.86	0.74	0.83	0.88
1945-1969	0.64	0.76	0.84	0.71	0.81	0.87	0.75	0.83	0.89
<1945	0.66	0.77	0.85	0.72	0.82	0.88	0.76	0.84	0.89

FIGURE 2.53 – Taux de destruction retenus adaptés aux caractéristiques de l'appartement et à la crue (Appartement RDC).

Appartement au 1er étage : Hauteur d'immersion (à partir du niveau 0 de l'immeuble)									
Durée d'immersion ↘ Année de construction ↓	≤ 2.5 mètres			3 mètres			3.5 mètres		
	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours
≥ 2001	0.05	0.05	0.05	0.58	0.72	0.81	0.66	0.78	0.85
1986-2000	0.05	0.05	0.05	0.60	0.74	0.82	0.68	0.79	0.86
1970-1985	0.05	0.05	0.05	0.62	0.75	0.83	0.70	0.80	0.86
1945-1969	0.05	0.05	0.05	0.64	0.76	0.84	0.71	0.81	0.87
<1945	0.05	0.05	0.05	0.66	0.77	0.85	0.72	0.82	0.88

FIGURE 2.54 – Taux de destruction retenus adaptés aux caractéristiques de l'appartement et à la crue (Appartement 1er étage).

Appartement au 2ème étage : Hauteur d'immersion (à partir du niveau 0 de l'immeuble)									
Durée d'immersion ↘ Année de construction ↓	≤ 4.5 mètres			5 mètres			5.5 mètres		
	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours	≤ 2 jours	2 à 4 j	≥ 4 jours
≥ 2001	0.05	0.05	0.05	0.05	0.05	0.05	0.58	0.72	0.81
1986-2000	0.05	0.05	0.05	0.05	0.05	0.05	0.60	0.74	0.82
1970-1985	0.05	0.05	0.05	0.05	0.05	0.05	0.62	0.75	0.83
1945-1969	0.05	0.05	0.05	0.05	0.05	0.05	0.64	0.76	0.84
<1945	0.05	0.05	0.05	0.05	0.05	0.05	0.66	0.77	0.85

FIGURE 2.55 – Taux de destruction retenus adaptés aux caractéristiques de l'appartement et à la crue (Appartement 2ème étage).

2.6 Évaluation des pertes

2.6.1 Méthodologie

Dans cette partie, nous cherchons à déterminer la fonction de perte reliant notre portefeuille aux taux de destruction définis précédemment. La caractérisation multivariée de nos courbes de destruction et la considération d'un module aléa continu dans le temps nous permet de définir une fonction de perte dépendante à la fois de la hauteur d'eau mais également de la durée de crue pour chaque exposition donnée. C'est donc toute la dynamique de la crue qui va influencer sur les résultats. Les paramètres en entrée de cette fonction de perte sont alors la dynamique complète de crue sur l'ensemble de sa durée (période durant laquelle la crue est supérieure à 4 mètres) ainsi que les caractéristiques de l'exposition (altitude, type d'habitation, etc...).

$$\forall C_{h(T)} \text{ crue de dynamique } h(T), f_p(C_T) = \sum_{e=1}^E t_D(e; C_T) * K_e \text{ où :} \quad (2.3)$$

- e est un bien assuré du portefeuille
- E est le nombre total de contrats dans le portefeuille
- K_e est le capital assuré pour l'exposition e
- $t_D(e; C_T)$ est le taux de destruction du bien e pour une crue à dynamique C_T

Afin de mieux comprendre le fonctionnement de l'évaluation des pertes, voici une simulation de la dynamique C_T de la hauteur d'eau à Austerlitz sur une plage de temps quelconque :

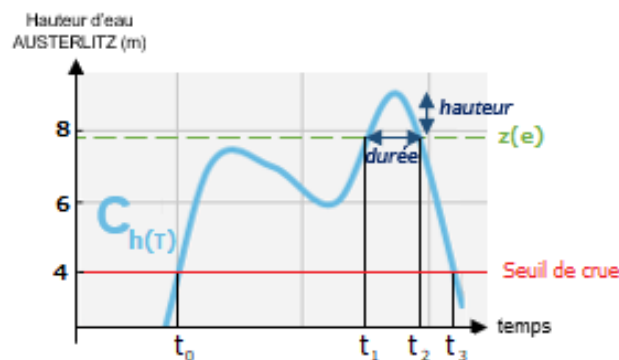


FIGURE 2.56 – Étude de la perte brute pour une exposition du portefeuille et une crue donnée.

On a alors la formule de pertes pour l'exposition e :

$$\text{pertes}(e) = t_D(e; C_T) * K_e \quad (2.4)$$

À la date t_0 , la Seine entre en crue. À la date t_1 l'exposition e est touchée car l'eau atteint son niveau 0 (altitude), notée $z(e)$. Ce n'est qu'à la date t_2 que l'eau quitte

l'habitation. La durée d'immersion de e est donc de $t_2 - t_1$, et la hauteur maximale atteinte est de $\max(C_h(T)) - z(e)$ où $z(e)$ correspond à l'altitude de l'exposition. On en déduit alors les pertes correspondantes en prenant le taux de destruction correspondant tel que spécifié dans la partie précédente. Une fois ce calcul de pertes par exposition réalisé, la somme donne alors la perte totale brute associée au portefeuille pour la crue $C_h(T)$.

2.6.2 Représentation graphique de la fonction de perte

Nous avons vu plus tôt que les taux de destruction des biens assurés et donc les pertes associées dépendent de la hauteur d'eau et de la durée d'immersion. Nous ne pouvons donc pas tracer toutes les combinaisons possibles de dynamiques de crue $C_h(T)$. Afin d'obtenir une idée des pertes potentielles, nous allons tracer les courbes des pertes brutes à durée d'immersion figée. Ainsi, nous supposons que chacune de nos expositions possède la même durée d'immersion. Cela revient à avoir une dynamique de crue $C_h(T)$ égale à une fonction porte de support égal à la durée d'immersion considérée (\leq à deux jours, entre deux et quatre jours ou \geq à quatre jours). Voici le graphique des pertes brutes associées une dynamique de crue :

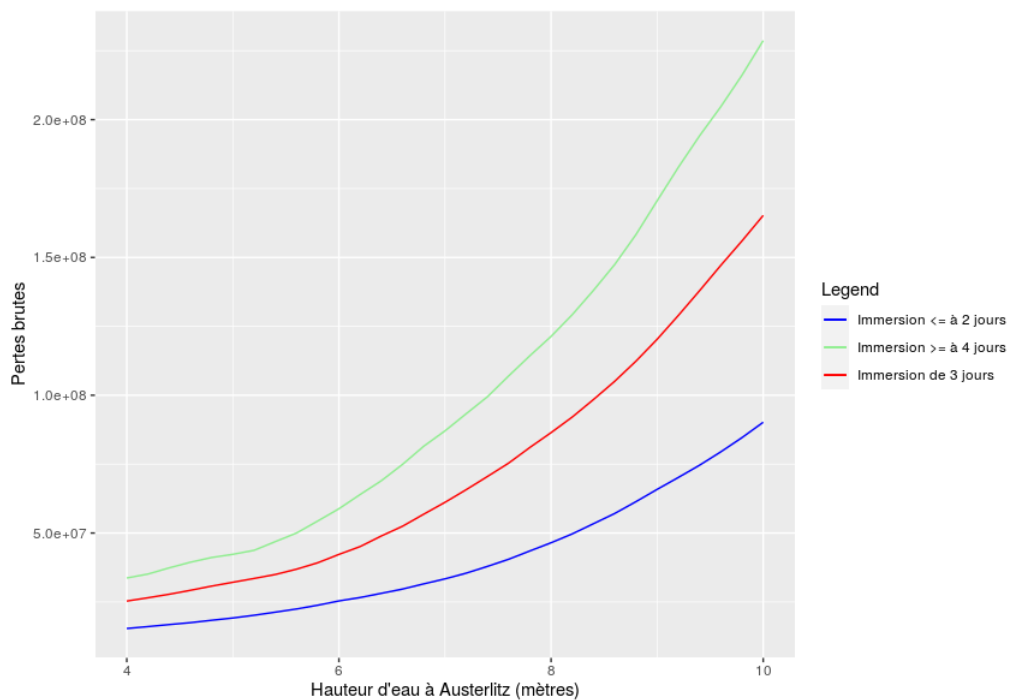


FIGURE 2.57 – Étude de la perte brute pour une exposition du portefeuille et une crue donnée.

On peut également représenter en trois dimensions ce graphique. Si nous avons considéré une séparation encore plus fine des jours d'immersion, nous aurions obtenu un plan

et au moyen d'une interpolation entre les points, une valeur donnée pour des paramètres continus de hauteur d'eau et de durée d'immersion.

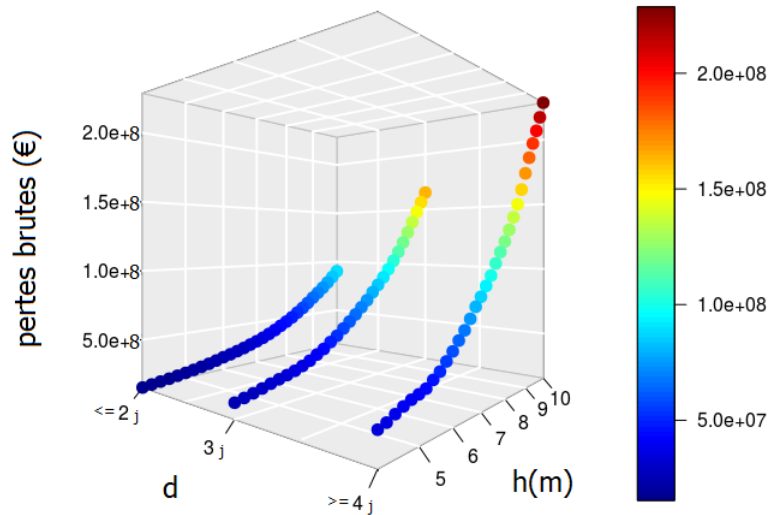


FIGURE 2.58 – Étude de la perte brute selon la dynamique de crue (hauteur et durée d'immersion).

Maintenant que nous savons déterminer les pertes brutes associées à n'importe quelle crue, il nous reste à déterminer la probabilité associée à ces pertes en fonction de l'intensité des crues. Pour cela, nous allons modéliser la hauteur d'eau relevée à Austerlitz sur une année et calculer les pertes brutes associées en cas de crue apparue dans cette simulation. Nous connaissons alors la distribution des pertes annuelles brutes en effectuant un grand nombre de simulations. Nous pourrions également comptabiliser l'apparition de crues d'une hauteur donnée pour déterminer la période de retour associée.

Chapitre 3

Modélisation de l'aléa et probabilisation des pertes brutes annuelles

Nous allons dans ce chapitre modéliser le phénomène de crue. L'indicateur principalement utilisé pour déterminer la présence et l'intensité d'une crue parisienne de la Seine est la hauteur d'eau relevée à l'échelle d'Austerlitz. Afin de la modéliser, nous allons étudier l'hydrométrie et la pluviométrie sur l'ensemble du bassin versant. Le but est de déterminer le lien entre les précipitations tombant dans le bassin de la Seine à une date donnée, et l'apparition d'une crue dans Paris à une date ultérieure. Nous allons pour cela étudier la dynamique du fleuve afin de connaître la durée de propagation d'une onde de crue dans le bassin. Ensuite, nous allons modéliser les précipitations tombant en amont de Paris et les relier avec la dynamique spatio-temporelle du fleuve pour déterminer la date et l'intensité de la crue apparaissant dans Paris à la suite des pluies tombées précédemment plusieurs dizaines voire centaines de kilomètres en amont de la capitale. Nous ne nous limiterons pas à calculer les périodes de retour de certaines crues, mais nous allons modéliser les hauteurs d'eau journalières à Austerlitz sur une année complète. Nous déduirons alors les pertes annuelles associées à chacune des simulations. Cela nous permettra alors d'en déduire des périodes de retour de pertes annuelles, utiles par exemple pour un calcul de SCR en modèle interne. Voici les étapes du module aléa en résumé¹ :

Modélisation

- Étude de la dynamique du fleuve dans le but de connaître les vitesses de propagation des ondes de crue ;
- Étude des précipitations du bassin et calibrage de lois et de structures de dépendances sur ces dernières ;
- Mise en place d'un modèle explicatif de la hauteur d'eau à Austerlitz selon les précipitations apparentes dans le bassin à J-1, J-3, ...

1. Un schéma résumé est disponible en Annexe B.

Simulation

- Simulation des précipitations annuelles pour chaque station ;
- Détermination de la hauteur d'eau journalière à Austerlitz consécutive aux précipitations ;
- Calcul des pertes annuelles consécutives à ces précipitations.

3.1 Hydrométrie

3.1.1 Base de données HYDRO

La DREAL (Directions Régionales de l'Environnement, de l'Aménagement et du Logement) relève depuis plus de 100 ans des données hydrométriques sur le sol français avec plus de 3 000 stations de mesure. La plus ancienne, celle d'Austerlitz génère des données depuis 1885. Ces données sont cependant très disparates, avec des stations créées et d'autres supprimées régulièrement ce qui fait qu'il est difficile de constituer un panel de stations temporellement homogènes. Pour notre étude, nous avons étudié les relevés disponibles sur le bassin versant de la Seine en amont de Paris. Nous avons ensuite sélectionné une dizaine de stations les plus équidistantes possibles. Finalement, nous avons étudié 19 stations au travers du bassin en plus de celle d'Austerlitz, toutes réparties parmi la Seine et ses trois principaux affluents en amont de Paris, à savoir la Marne, l'Yonne et l'Aube.

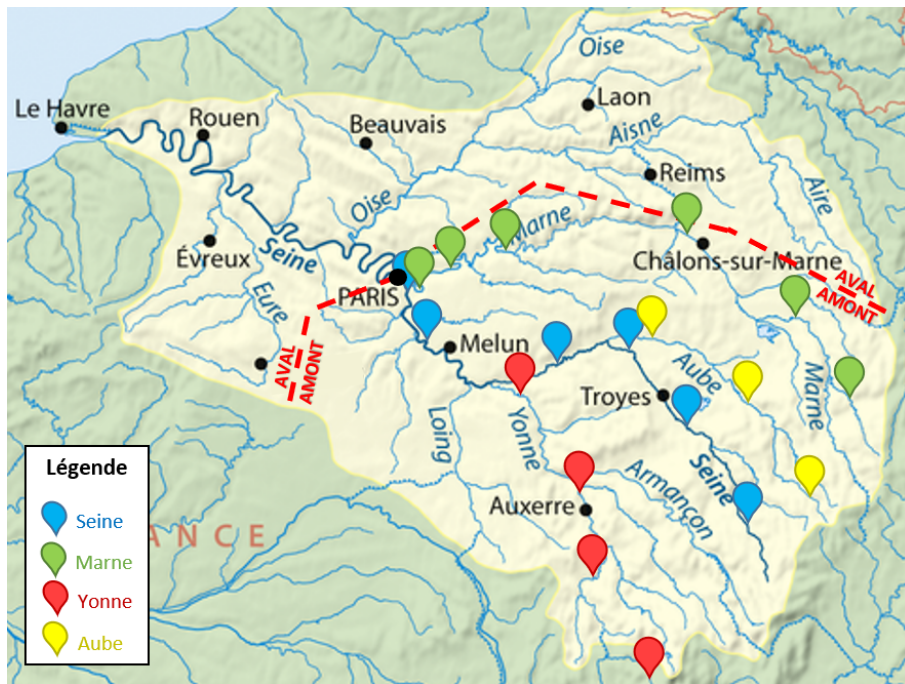


FIGURE 3.1 – Cartographie des stations étudiées.

On peut remarquer que cette cartographie des stations est dans l'ensemble spatialement homogène. Pour ce qui est des relevés par station, on constate certaines différences que ce soit par rapport à la plage de mesures où à l'intervalle de temps entre ces dernières. Voici une représentation en réseau des stations considérées classées selon la dynamique d'écoulement du fleuve :

Le but est alors de déterminer les lags temporels entre chaque station, témoins de la durée moyenne d'écoulement du fleuve sur chaque affluent. Ils représentent ainsi la

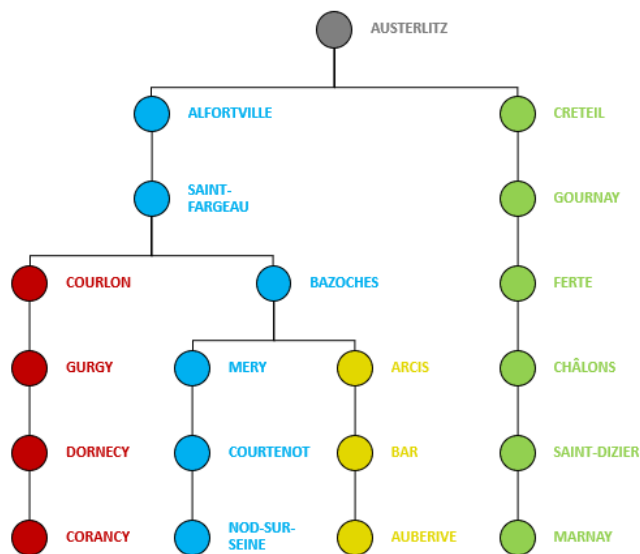


FIGURE 3.2 – Réseau constitué des stations *HYDRO* étudiées.

vitesse de propagation de l'onde de crue. En plus d'être utile dans le domaine de la prévision des crues -externe à ce mémoire-, cette démarche est primordiale afin d'obtenir une idée précise du temps pris par une goutte de pluie tombant dans le bassin versant pour atteindre Austerlitz et ainsi considérer de façon la plus réaliste possible le phénomène de crues concomitantes et ainsi obtenir une modélisation consistante.

Nous avons donc récupéré les données historiques de ces 20 stations. Le tableau 3.1 regroupe quelques indicateurs afin de prendre connaissance des caractéristiques des données.

La station qui nous intéresse pour modéliser les crues de la Seine sur le bassin parisien est celle d'Austerlitz. Voici quelques statistiques permettant de bien assimiler l'ordre de grandeur des données :

- **Minimum observé** : $-0,50\text{m}^2$ le 5 janvier 1891 à 8h00
- **Moyenne observée** : 1,26m
- **Maximum observé** : 8,62m le 29 janvier 1910 à 12h00
- **Seuil d'alerte de niveau 1** : 3,20m : certaines voies de circulation ferment.
- **Seuil intermédiaire** : 3,40m : la circulation sur la Seine est interdite.
- **Seuil d'alerte de niveau 2** : 5,75m : fermeture du RER C.
- **Seuil intermédiaire** : 6,20m : fermeture complète des voies sur berge.
- **Seuil d'alerte de niveau 3** : 7,13m : fermeture des RER A, RER B et d'une grande partie des tronçons de métro.

2. Le zéro de l'échelle n'est pas au fond du fleuve, ce qui peut expliquer des valeurs négatives.

Affluent	Station	Plage de temps	Relevés
Seine	Austerlitz	1885-2021	375 000
Seine	Alfortville	1992-2006/2012-2021	120 000
Seine	Saint-Fargeau	1999-2021	150 000
Seine	Bazoches	1999-2021	115 000
Seine	Mery	1982-2021	300 000
Seine	Courtenot	1983-2021	140 000
Seine	Nod	1900-2021	77 000
Marne	Créteil	2010-2021	425 000
Marne	Gournay	2008-2016	255 000
Marne	Ferté	1993-2021	110 000
Marne	Châlons	1989-2021	350 000
Marne	Saint-Dizier	1983-2021	120 000
Marne	Marnay	1983-2021	115 000
Yonne	Courlon	1982-2021	250 000
Yonne	Gurgy	1967-2021	125 000
Yonne	Dornecy	1989-2021	80 000
Yonne	Corancy	1990-2021	60 000
Aube	Arcis	1983-2021	425 000
Aube	Bar	1987-2021	215 000
Aube	Auberive	2000-2021	45 000

TABLE 3.1 – Caractéristiques des stations *HYDRO* considérées.

Après avoir nettoyé nos données des valeurs aberrantes, nous pouvons tracer l'historique des hauteurs d'eau relevées à Austerlitz depuis la fin du XIXème siècle :

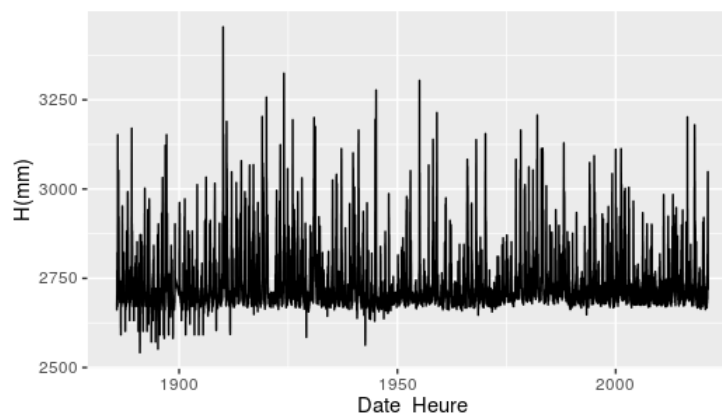


FIGURE 3.3 – Historique des hauteurs d'eau relevées à Austerlitz depuis 1885.

Nous avons effectué le même traitement pour toutes les autres stations. On rappelle que les données sont très disparates car calquées sur des plages historiques différentes et avec des intervalles de mesures hétérogènes. On présente ci-dessous les tracés historiques d'une station aléatoire par branche du réseau tel que définit en figure 3.2 :

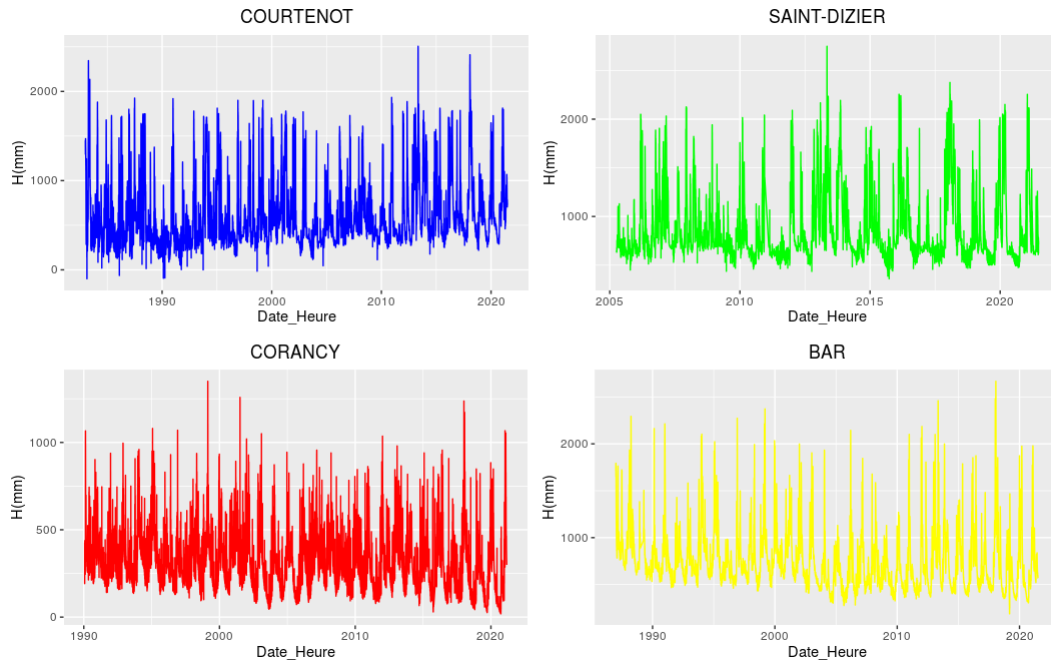


FIGURE 3.4 – Différents aperçus des données hydrométriques.

Ci-dessous les données par mois. On remarque que certains mois de l'année n'ont jamais observé de crue majeure. On constate également à quel point la crue de 2016 était exceptionnelle, car très loin du niveau habituel du fleuve en mai, et conséquence d'une pluviométrie inédite pour un tel mois.

3.1.2 Étude de la dynamique de la Seine

Comme nous avons pu le voir précédemment, la hauteur d'eau à Austerlitz n'est pas seulement la conséquence de pluies tombant à Paris. Le fleuve possède une dynamique entière pilotée par son bassin versant. L'écoulement des eaux d'un point A à un point B et le phénomène de crues concomitantes expliqué dans le chapitre 2 nous pousse à devoir modéliser la hauteur d'eau du bassin tout entier. On cherche alors à capter comment une goutte d'eau tombée dans un affluent de la Seine tel que la Marne ou l'Yonne par exemple va se comporter et combien de temps va t'elle mettre pour arriver à Paris et ainsi influencer sur la hauteur d'eau relevée dans la capitale. Cette dynamique spatio-temporelle est complexe à mettre en place mais néanmoins primordiale afin de chercher ensuite à expliquer la hauteur d'eau à Paris en fonction de la pluviométrie du bassin tout entier.

Pour ce faire, nous avons estimé les lags temporels entre chacune des stations du réseau

et la station d’Austerlitz, repère de l’apparition d’une crue parisienne. Cette cartographie définissant la dynamique du fleuve nous servira ensuite à expliquer la hauteur d’eau au pont d’Austerlitz en fonction des pluies tombées en amont de la capitale, lagées des heures/jours caractérisant la vitesse d’écoulement de l’eau.

Afin d’illustrer notre démarche, nous allons représenter un graphique de hauteurs d’eau relevées entre janvier et avril 2020 à Austerlitz ainsi qu’à Arcis, première station de l’Aube en amont de Paris de notre réseau. Elle se situe à environ 200 kilomètres de Paris.

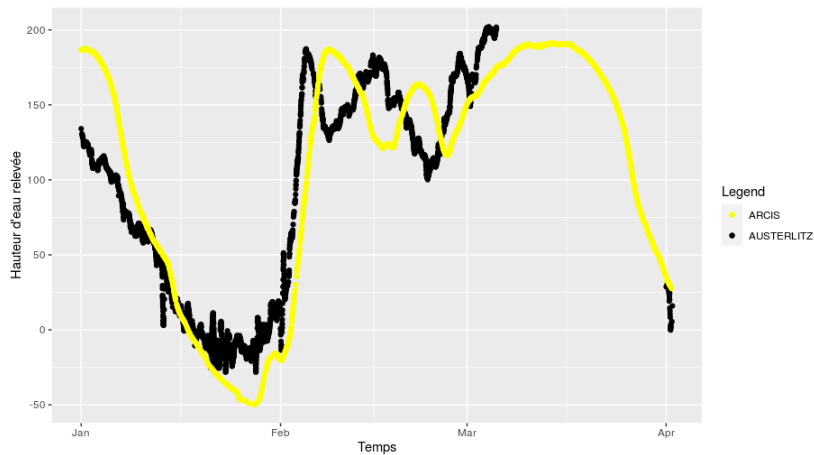


FIGURE 3.5 – Relevés hydrométriques à Austerlitz et à Arcis entre janvier et avril 2020.

Afin de déterminer ces lags, nous avons dû retraiter les données en amont. La première étape a été de repérer les valeurs aberrantes (sauts de hauteurs d’eau en un laps de temps minime, valeurs extrêmes, ...) et de les supprimer. Ensuite, afin de pouvoir comparer nos différentes stations entre elles, nous avons dû normaliser chacune de ces séries. Ce sont les séries retraitées de ces deux caractéristiques qui sont représentées ci-avant. Le traitement de normalisation est le suivant :

$$\forall t \text{ date de relevé, } f_n(t) = f(t) - \frac{1}{N} \sum_{i=1}^N f(i) \text{ où :} \quad (3.1)$$

- f est une fonction de relevés de hauteur d’eau à une station (Austerlitz ou Arcis par exemple)
- f_n est la fonction de relevés normalisée associée à f
- N est le nombre total de relevés dans l’historique de la station

Un autre problème apparaît sur le graphique 3.5, la différence de séquantage temporel. Comme nous l’avons relaté lors de la présentation de la banque de données *HYDRO*, les relevés sont effectués à des périodes de temps différentes et à lag temporel différent. Ainsi, nous pouvons avoir un relevé toutes les heures précises lorsque nous possédons un relevé toutes les vingt-cinq minutes à Arcis. Afin de remédier à ce problème, nous avons

procédé à une interpolation linéaire de chacune des séries par pas de temps de 10 minutes (pas de temps minimal observé). :

$$\forall s, f(s) = f(t_i) + \frac{f(t_{i+1}) - f(t_i)}{t_{i+1} - t_i} \text{ où :} \quad (3.2)$$

- s est une date-heure de relevé entre t_0 et t_N multiple de 10 minutes (exemple : "01/01/2020 09h20", "01/01/2020 09h30", ...)
- f est une fonction de relevés de hauteur d'eau à une station (Austerlitz ou Arcis par exemple)
- t_i est le relevé précédant la date-heure s . S'il existe un relevé à la date-heure s , alors $t_i = s$
- t_{i+1} est le relevé suivant strictement la date-heure s

Voici un aperçu de l'interpolation des relevés d'eau à Austerlitz du 1er janvier 2020.

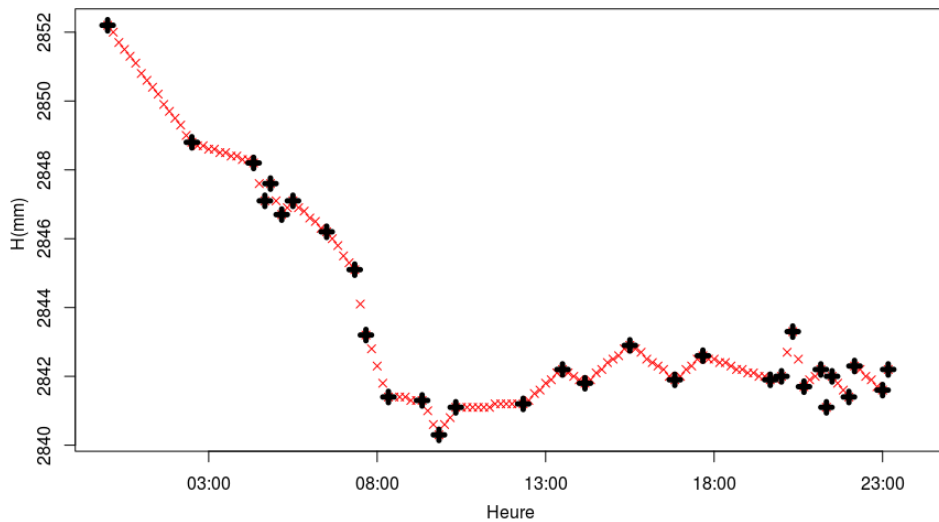


FIGURE 3.6 – Interpolation des relevés d'eau du 1er janvier 2020 à Austerlitz.

Cette interpolation nous permet maintenant de comparer deux séries de façon quasi-continue. Notre objectif est maintenant de déterminer l'onde de crue entre ces deux stations, c'est-à-dire lorsque le fleuve monte à Arcis (du fait de précipitations), combien de temps met-il pour monter à Austerlitz. Il est évident que du fait de la complexité du fonctionnement du bassin, nous savons que le fleuve peut monter à Austerlitz alors qu'il est totalement stable à Arcis en cas de fortes pluies entre ces deux zones par exemple. Cependant, lorsque l'on observe le graphique 3.9, on remarque une tendance globale similaire avec des pics et des creux identiques mais lagés d'un laps de temps. C'est ce temps que nous cherchons à déterminer. Il correspond pour nous à la durée moyenne de propagation de l'onde de crue d'Arcis à Austerlitz, que nous assimilons à la durée entreprise par la pluie tombant dans le bassin Arcisien pour arriver jusqu'à la capitale.

Pour déterminer ce lag, nous utilisons une méthode assimilable aux moindres carrés.

Nous allons calculer la somme des écarts au carré entre les deux hauteurs d'eau normalisée pour chaque pas de temps (possible grâce à l'interpolation), celle d'Austerlitz et celle de l'autre station. Cette valeur, que nous nommerons *ErrOnde*, définit pour chaque station du réseau (hors Austerlitz) le taux de concordance temporelle entre les deux courbes, c'est-à-dire à quel point elles "collent" l'une sur l'autre.

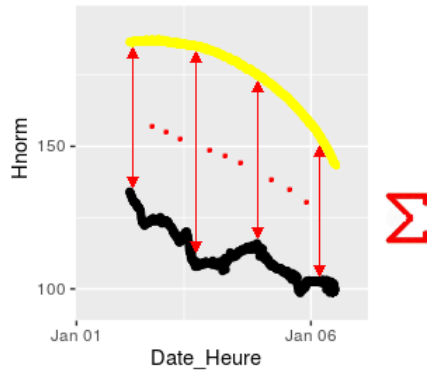


FIGURE 3.7 – Méthode des moindres carrés.

Voici la formule utilisée :

$$ErrOnde(g_n^{int}) = \sum_{i=1}^I |f_{n,AUS}^{int}(s_i) - g_n^{int}(s_i)|^2 \text{ où :} \quad (3.3)$$

- s est une date-heure de relevé entre t_0 et t_N multiple de 10 minutes (exemple : "01/01/2020 09h20", "01/01/2020 09h30", ...)
- $f_{n,AUS}^{int}$ est la fonction de relevés de hauteur d'eau d'Austerlitz interpolée et normalisée
- g_n^{int} est une fonction de relevés de hauteur d'eau à une station du réseau (hors Austerlitz)
- s_1, \dots, s_I sont les laps de temps de la période d'interpolation considérée

L'idée est ensuite de lagger la courbe de la 2ème station afin de faire en sorte que les pics et les creux se retrouvent au même instant t , ce qui aura pour conséquence de minimiser *ErrOnde*. Ainsi, un lag d'un pas correspondra à décaler la série de la 2ème station étudiée (ici Arcis) de 10 minutes en arrière. Un lag de m pas correspondra alors à un lag temporel de $10 \times m$ minutes. Commençons par définir une série temporelle laggée de m pas. On utilise pour cela l'opérateur de retard Δ :

$$Lag(g_n^{int}, m)(s_i) := \Delta^m g_n^{int}(s_i) := g_n^{int}(s_{i-m}) \text{ où :} \quad (3.4)$$

- s_i est la i -ème valeur de l'intervalle de temps d'interpolation
- m est le lag considéré

Nous cherchons alors le lag optimal m^* qui va déterminer notre lag temporel entre les deux stations. Pour cela, nous utiliserons la fonction **optimize** de **R** en plus de nos propres fonctions de normalisation, d'interpolation et de calcul d'*ErrOnde*.

$$m^*(g) = \operatorname{argmin}_m [\operatorname{ErrOnde}(\operatorname{Lag}(g_n^{\text{int}}, m))] \quad (3.5)$$

Afin d'avoir des données homogènes et assez conséquentes pour contrer le phénomène décrit plus tôt expliquant que certaines mini-crués peuvent apparaître sur une station et pas sur l'autre, nous avons considéré une plage de temps identique pour chacune des stations et assez grande. Ainsi, nos calculs de lags temporels sont effectués sur la plage de temps : "01/01/2014"- "31/12/2016" soit trois années complètes. Cette valeur nous paraît suffisante et permet de réduire les temps de compilation.

Nous avons effectué ce travail pour chacune des 19 stations du réseau pour connaître la dynamique complète du fleuve dans son bassin versant. Les résultats sont présentés ci-dessous. On notera que nous ne présentons pas de valeur pour les stations de Créteil et d'Alfortville, ces stations étant en amont direct de Paris, leur lag avec Austerlitz est très faible et donc peu intéressant.

Affluent	Station	Lag temporel ³ (jours)
Seine	Austerlitz	0
Seine	Saint-Fargeau	0,3
Seine	Bazoches	1,6
Seine	Mery	1,9
Seine	Courtenot	2,4
Seine	Nod	3,0
Marne	Gournay	0,8
Marne	Ferté	1,4
Marne	Châlons	3,1
Marne	Saint-Dizier	4,6
Marne	Marnay	6,3
Yonne	Courlon	1,7
Yonne	Gurgy	2,3
Yonne	Dornecy	3,1
Yonne	Corancy	3,9
Aube	Arcis	2,0
Aube	Bar	2,9
Aube	Auberive	4,0

TABLE 3.2 – Vitesse de propagation d'une onde de crue vers Austerlitz.

Des ordres de grandeurs ont été données dans le rapport de la mission par la Direction régionale et interdépartementale de l'environnement et de l'énergie (DRIEE) : "Les crués de la Seine en Île-de-France mettent entre trois à quatre jours depuis l'amont de l'Yonne ou du Loing et six à huit jours depuis l'amont de la Seine et de la Marne pour parvenir dans l'agglomération parisienne", [DRIEE, 2017]. Si l'on regarde de plus près

3. On rappelle que ces lags sont tous effectués par rapport à la station Austerlitz.

nos résultats, nous sommes globalement en phase avec ces propos. Nous pouvons alors maintenant cartographier l'ensemble du bassin. Ces valeurs nous serviront lors de la mise en place de notre modèle explicatif de la hauteur d'eau à Austerlitz en fonction de la pluviométrie du bassin. Nous y reviendrons par la suite.

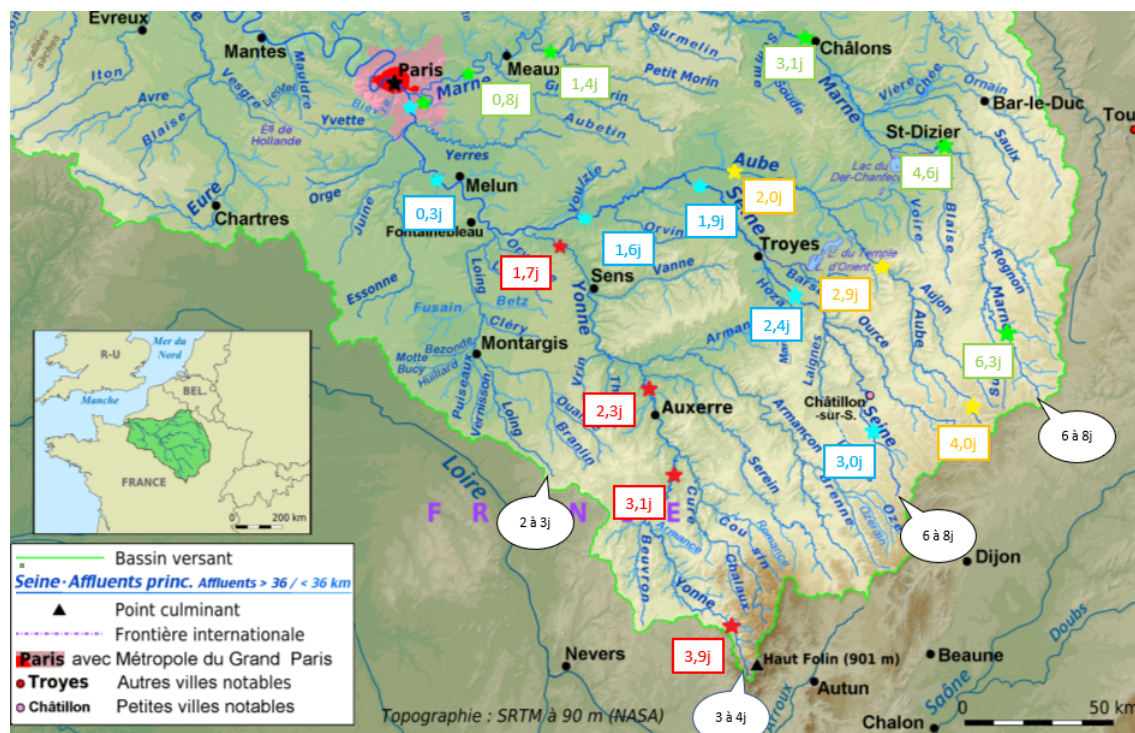


FIGURE 3.8 – Dynamique des ondes de crue du bassin versant de la Seine.

3.2 Pluviométrie

3.2.1 Base de données ISD

Les données météorologiques en France ne sont malheureusement pas facilement accessibles. Météo France, établissement public administratif, est le service officiel de la météorologie et de la climatologie du pays. Cependant, la mise à disposition de leurs historiques est payante et nous n'avons donc pas pu les utiliser dans le cadre de ce mémoire. Il est également possible de trouver en libre accès et gratuitement quelques historiques météorologiques pour certaines grandes métropoles, mais la profondeur des données est trop souvent faible.

La NOAA (National Oceanic and Atmospheric Administration), agence américaine responsable de l'étude de l'océan et de l'atmosphère, a lancé un programme nommé *ISD* (Integrated Surface Data) au milieu du XX^{ème} siècle. C'est une base de données composée d'observations climatiques mondiales via de multiples stations réparties sur le

globe. Différents indicateurs y sont relevés tels que la vitesse et la direction du vent, les rafales de vent, la température, la part de nuages dans le ciel, la pression au niveau de la mer, le calage altimétrique, la pression de la station, la visibilité, les quantités de précipitations, la profondeur de neige, etc... La NOAA laisse le libre accès à ces données, ce qui nous a permis de récupérer des données météorologiques sur le sol français avec une profondeur d'historique suffisante. Comme pour nos données hydrométriques, les données *ISD* sont très disparates avec des plages d'observations et des intervalles de mesure souvent différents. Pour la France, 334 stations ont été ou sont toujours utilisées pour relever des données météorologiques. Voici les stations présentes dans les régions traversées par la Seine ou un de ses affluents :

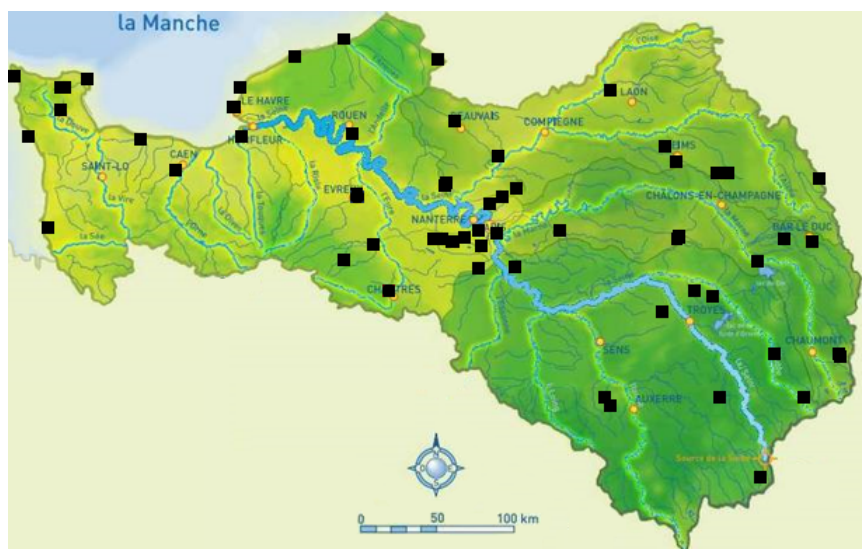


FIGURE 3.9 – Stations ISD sur le bassin de la Seine.

L'idée est de réaliser un maillage des relevés de pluie en amont de Paris, puis de faire concorder les relevés pluviométriques avec les relevés des stations hydrométriques voisines afin de connaître le temps parcouru par les gouttes de pluies de chaque bassin pour arriver jusqu'à Paris (station d'Austerlitz). Cependant, nous savons que les précipitations possèdent une dynamique spatio-temporelle. Ainsi, nous ne pouvons pas modéliser les pluies des différentes stations indépendamment, car s'il pleut à Sens par exemple, il y a de fortes chances pour qu'il pleuve également à Auxerre située à moins de 50 kilomètres à vol d'oiseau. Nous avons donc besoin de données avec une plage d'observation commune. Certaines stations sont également hors service depuis des décennies. Nous avons donc effectué un tri pour ne garder que les stations exploitables et nous avons finalement retenu neuf stations au cœur de notre bassin versant. Notre historique de données est de l'ordre d'une quarantaine d'années de profondeur. Au vu de l'évolution actuelle du climat sur le globe et des problématiques de réchauffement climatique, cet historique reste suffisant car remonter à trop loin pourrait également biaiser le modèle si on considère que le climat est une variable mouvante au fil des décennies.

Pour représenter la cartographie de ces stations et déterminer leur zone d'action approximative (zones homogènes en loi et modélisées par la même pluviométrie) nous avons tracé les cellules de Voronoï déterminées par ces stations sur bassin de la Seine en amont de Paris. Cela revient à associer chaque point de la carte à la station météorologique la plus proche. Ainsi, nous avons une idée approximative du lag temporel qui existe lorsque qu'une goutte de pluie tombe dans une zone donnée proche d'une station *HYDRO*.

Cellules de Voronoï : soient $M \subset \mathbb{R}^2$ (plan affine) et S un ensemble de n points dans M . On appelle cellule de Voronoï de $i \in S$ l'ensemble :

$$Vor(i) := \{x \in M / \forall j \neq i \in S, \|x - i\| \leq \|x - j\|\}$$

Autrement dit, étant donné un point i de S , la cellule de Voronoï de i correspond à l'ensemble des points plus près de i que de tout autre point de S .

Voici un aperçu de la cartographie que l'on obtient :

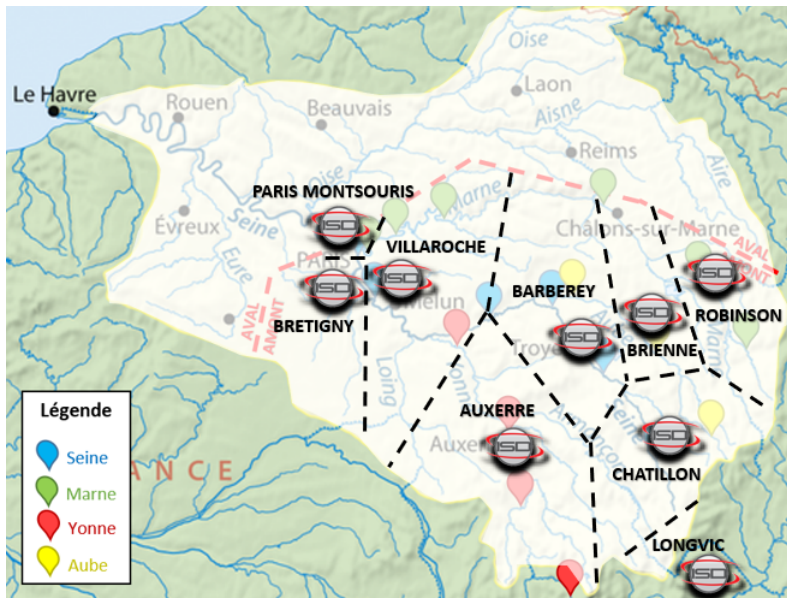


FIGURE 3.10 – Cartographie des stations ISD étudiées.

Pour souci de lisibilité, nous noterons par la suite les stations *ISD* entièrement en majuscule et les stations *HYDRO* avec une unique majuscule.

3.2.2 Analyse des données pluviométriques

Traitement des données

Les données pluviométriques constituent le cœur de notre étude, elles sont donc très importantes. Comme nous l'avons précisé plus tôt, nous nous sommes intéressés à neuf stations : PARIS MONTSOURIS, BRETIGNY, VILLAROCHE, BARBEREY,

AUXERRE, BRIENNE, CHATILLON, ROBINSON et LONGVIC. Chacune est plus ou moins loin de Paris et va donc plus ou moins influencer sur la hauteur d'eau relevée dans la capitale et avec un lag temporel différent des autres.

Présentons par exemple l'ensemble de nos données pour la station PARIS MONT-SOURIS. Cette station est située au sud de Paris, très proche donc d'Austerlitz. On imagine alors que la pluviométrie relevée sur cette station va influencer de façon quasi-instantanée sur la hauteur d'eau de la Seine dans la capitale. On remarquera que la profondeur d'historique pour PARIS MONT-SOURIS est d'environ 40 ans, même si l'on remarque une absence de relevés sur la période 1993-1999.

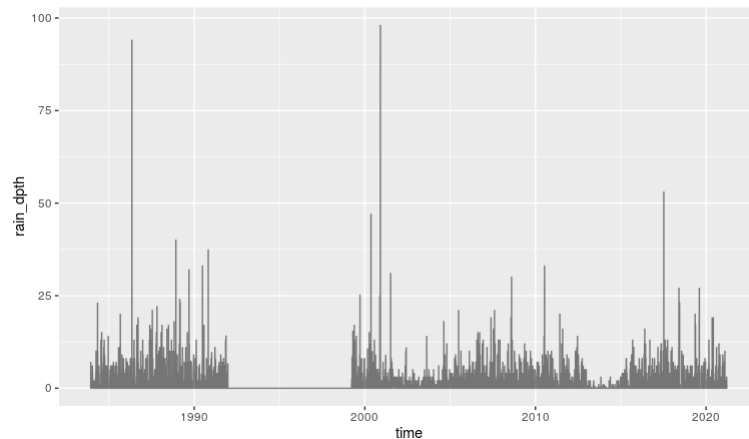


FIGURE 3.11 – Historique des pluies relevées à PARIS MONT-SOURIS.

Il peut être intéressant de zoomer sur certaines dates particulières afin de repérer des crues historiques. On peut par exemple regarder la pluviométrie de Paris lors de la période où est apparue la crue de 2016. On notera que la pluviométrie affichée est la pluviométrie agrégée par demi-journée contrairement au graphique précédent où certains relevés sont toutes les heures par exemple, ce qui explique des valeurs plus basses. Nous y reviendrons plus tard.

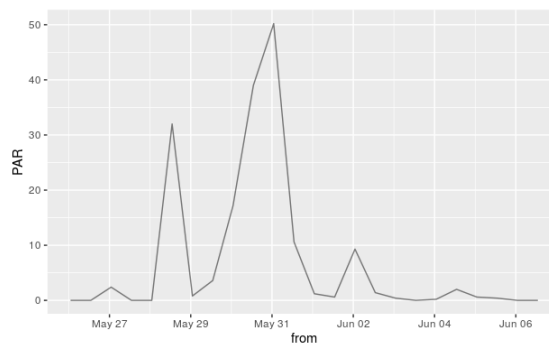


FIGURE 3.12 – Précipitations à PARIS MONT-SOURIS lors de la crue de 2016.

On observe bien une augmentation soudaine de la pluie dans la capitale à partir de la nuit du 28 au 29 mai. Cependant, cela ne suffit pas de regarder une unique pluviométrie, car comme nous l'avons fait remarquer plusieurs fois, l'apparition d'une crue peut provenir de pluies observées en amont de la capitale.

Nous pouvons également comparer nos données au graphique 1.32 présenté dans le chapitre 2 mêlant hydrogramme et crue et précipitations relevées à Paris lors de la crue de 2016. On observe bien une similitude entre la dynamique pluviométrique présentée et nos données.

Sur le graphique de la hauteur d'eau relevée à PARIS MONTSOURIS entre le 27 mai et le 4 juin 2016, on remarque que l'on possède un relevé toutes les 12 heures. Cependant, ce n'est pas le cas dans la base initiale. En effet, comme pour les données *HYDRO*, les dates de relevés sont variables. Il faut donc agréger les valeurs par tranches afin d'obtenir des relevés sur des intervalles de temps fixes et communs à chaque station. C'est ce qui a été réalisé sur le graphique présentant les précipitations à Paris lors de la crue de 2016. Nous avons choisi de considérer des plages d'agrégation de 12h car les lags de nos stations calculés plus tôt étaient de cet ordre de grandeur entre deux stations voisines du réseau. De plus, la seine monte doucement comme nous l'avons vu plus tôt, donc considérer un relevé toutes les 12h des pluies (et donc de la hauteur d'eau à Austerlitz) reste amplement suffisant.

On notera que les données pluviométriques représentent une profondeur de pluie, c'est-à-dire la totalité de la pluie tombée R_{t_i} entre la date t_i du relevé et la date t_{i-1} du relevé précédant $R_{t_{i-1}}$. Ainsi, il faut sommer la quantité totale de pluie tombée entre 00h et 11h59 puis entre 12h et 23h59 pour chaque journée de l'historique. Malheureusement, nous possédons parfois plusieurs données dans une tranche de 12h et parfois aucune. De la même façon que pour les données *HYDRO*, nous avons donc dû interpoler nos données, mais cette fois pour posséder un relevé toutes les 12h. Lorsque nous possédons plusieurs relevés à l'intérieur d'une même tranche nous les sommons, et lorsque l'on n'en possède pas ou qu'un seul relevé entre deux plages, nous distribuons la pluie tombée pour ce relevé dans les deux tranches voisines proportionnellement à la durée considérée dans chacune des tranches.

Soient T correspondant à une date donnée (le 23 décembre 2014 par exemple) et à une horaire (deux possibilités : soit $T = 00h00$, soit $T = 12h00$). On note $T' := T + 12h$ la date-heure de fin de la plage de temps $]T; T']$ pour lequel nous souhaitons avoir une valeur de profondeur de pluie. Notre plage d'observation est alors l'ensemble $\bigcup_{T_1}^{T_{2N}}]T, T + 12h]$ où T_1 est la première date de relevé et N le nombre de jours de relevés. On considère alors tous les relevés $R(t_0), \dots, R(t_n)$ pour lesquelles les dates de relevé t_0, \dots, t_n sont dans la plage de temps $]T, T']$. On note t_{-1} la date de relevé tout juste précédant t_0 et t_{n+1} la date de relevé tout juste suivant t_n . On obtient alors, en posant $P_{]T; T']}$ valeur des précipitations estimés entre T et T' :

$$P_{]T; T']} = \frac{R(t_0) - R(t_{-1})}{t_0 - T} + \sum_{i=1}^n R(t_i) + \frac{R(t_{n+1}) - R(t_n)}{T' - t_n} \quad \text{où :} \quad (3.6)$$

Un schéma représentant les plages de temps est placé en Annexe C afin de mieux comprendre la formule ci-dessus. On notera que l'on a fait le choix de supprimer les relevés où le temps entre deux relevés de stations est supérieur à 1 jour car nous les considérons peu fiables. Finalement, voici le graphique que l'on obtient après retraitement :

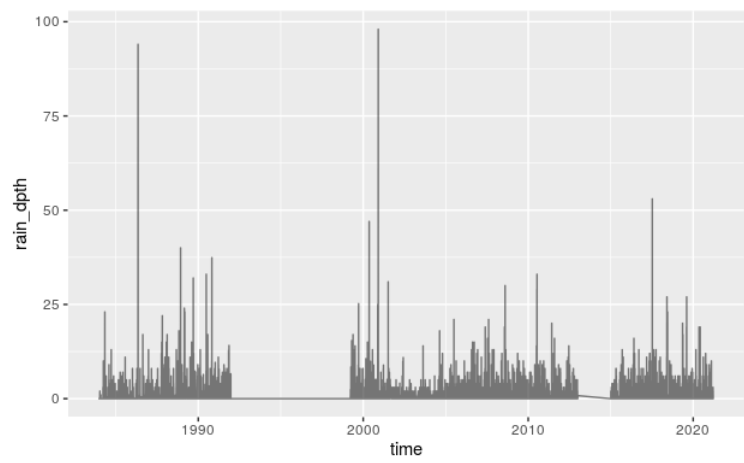


FIGURE 3.13 – Historique des pluies relevées à PARIS MONTsouris entre 1983 et 2021 après interpolation bi-hebdomadaire .

Séparation mensuelle

Comme nous l'avons expliqué plus tôt, nous savons que les précipitations fluctuent selon le mois de l'année. On peut alors observer des pluies plus fortes ou plus répétées selon que l'on soit en janvier ou en juillet par exemple. Ainsi, il devient pertinent d'effectuer des adéquations de loi différente selon le mois de l'année. Ceci se vérifie avec nos données. Prenons encore une fois la pluviométrie relevée à PARIS MONTsouris, et intéressons-nous simplement aux moyennes observées sur l'historique par mois.

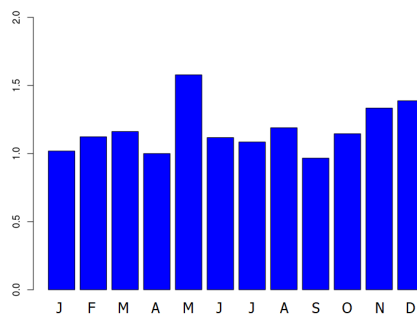


FIGURE 3.14 – Moyennes historiques des pluies relevées par demi-journée à PARIS MONTsouris par mois.

Cet indicateur nous montre déjà une dissemblance entre les lois mensuelles des précipitations. On peut également regarder la dispersion (écart-type) de ces mêmes données :



FIGURE 3.15 – Écart-types historiques des pluies relevées par demi-journée à PARIS MONTSOURIS par mois.

Là aussi on observe une différence nette entre les différents mois de l'année. L'écart-type est plus élevé l'été ce qui peut s'expliquer par un plus grand nombre de jours sans pluie mais avec des averses violentes ponctuelles, là où les mois d'hiver sont plus souvent soldés de pluies continues. Cela justifie le choix d'effectuer des calibrages des pluviométries conditionnellement au mois de l'année.

3.2.3 Calibrage paramétrique

Méthodologie

Afin de pouvoir simuler les pluies tombant dans le bassin à un instant donné, nous avons besoin de connaître les lois des précipitations. Comme nous l'avons précisé plus tôt, afin d'être le plus réaliste possible, nous modélisons séparément les pluies des différentes parties du bassin via les neuf stations de relevés décrites plus tôt, ainsi que les 12 mois de l'année, car nous avons pu remarquer que la loi régissant les précipitations est différente selon le mois en question. Ainsi, nous allons calibrer 108 lois différentes (9 stations \times 12 mois) permettant ainsi de modéliser l'ensemble du bassin.

Pour réaliser ces ajustements de loi, on considère les échantillons de données $x_S^M(t_1), \dots, x_S^M(t_{N_S^M})$ correspondant à l'historique des précipitations tombées sur les dernières 12 heures dans un mois M pour la station S . On effectue l'hypothèse que cet échantillon représente des mesures indépendantes de loi de probabilité F_S^M que nous cherchons à estimer. Nous allons utiliser des méthodes d'ajustement de l'échantillon observé à la loi théorique F afin de déterminer la loi optimale d'ajustement. Pour cela, nous utiliserons des tests statistiques ainsi que la comparaison des vraisemblances afin de mesurer la qualité statistique de l'ajustement.

Comme nous avons pu le voir dans la partie précédente, le nombre de demi-journées sans pluie est conséquent, ce qui provoque une inflation de zéro sur notre densité des pluies. Ainsi, calibrer directement nos précipitations à partir de lois usuelles n'est pas

pertinent car la distribution empirique des pluies n'est pas usuelle du fait de cette inflation de zéros. On peut en être sûr en regardant les distributions des pluies à PARIS MONTsouris en janvier par exemple. Représentons la densité empirique de la pluviosité brute et celle où l'on a retiré les demi-journées sans pluie. On remarque clairement que la première distribution possède une queue de distribution extrêmement longue car écrasée par le nombre de valeurs nulles. On voit très clairement qu'il est plus pertinent de calibrer la loi de droite, beaucoup plus régulière.

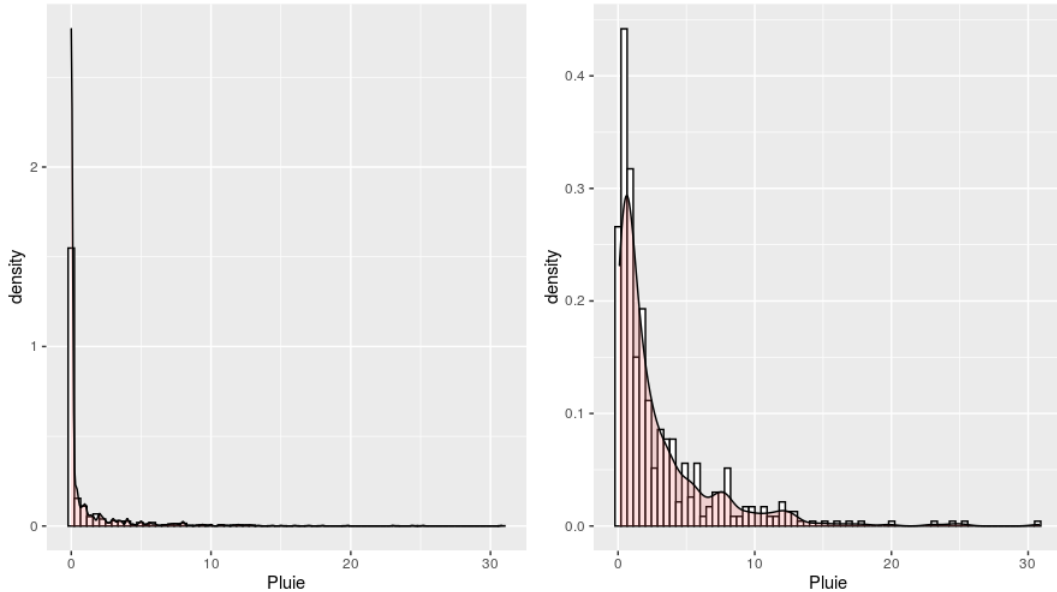


FIGURE 3.16 – Histogramme et densité empirique de la quantité de pluie tombée lors des demi-journées pluvieuses de janvier à PARIS MONTsouris.

Nous allons donc utiliser une loi combinée afin de paramétrer nos précipitations. Ainsi, pour chacun des mois, nous allons estimer d'une part la loi du nombre de demi-journées pluvieuses et d'autre part la distribution de la quantité de pluie lorsqu'il pleut. Ce modèle s'apparente à un modèle fréquence/sévérité. Nous allons donc estimer la loi de la variable aléatoire du nombre de demi-journées de pluies pour le mois M et la station S , que l'on notera N_S^M . Ensuite, nous estimerons la loi de la quantité de pluie tombée par demi-journée lorsqu'il pleut, notée Q_S^M .

Afin d'estimer N_S^M , nous allons utiliser les lois de comptage usuelles : Poisson (\mathcal{P}), binomiale (\mathcal{B}) et négative binomiale (\mathcal{NB}). L'utilité de ces deux dernières lois est de modéliser des situations où la variance n'est pas égale à l'espérance, ce qui est la propriété fondamentale de la loi de Poisson. La loi binomiale négative représentera une sur-dispersion, tandis que la loi binomiale représentera une sous-dispersion. Notons $n_S^M(1), \dots, n_S^M(A)$ les observations du nombre de demi-journées de pluies dans le mois où A représente le nombre d'années de l'historique. On notera $\hat{\lambda}_S^M$ la moyenne des demi-journées de pluie mensuelles de l'historique pour la station S et le mois M et $\hat{\sigma}_S^M$ l'écart-type empirique

du nombre de demi-journées de pluies mensuelles de l'historique. On note également m^M le nombre de demi-journées dans le mois M . On fera l'approximation $m^{Fevrier} = 28 \times 2$ en ignorant les années bissextiles.

On a alors :

$$\hat{\lambda}_S^M = \frac{1}{A_S^M} \sum_{i=1}^{A_S^M} n_{S^M}^M$$

$$\hat{\sigma}_S^M = \sqrt{\frac{1}{A_S^M} \sum_{i=1}^{A_S^M} (n_{S^M}^M - \hat{\lambda}_S^M)^2}$$

Ainsi, pour déterminer la loi de N_S^M nous avons décidé de comparer moyenne et variances empiriques et d'utiliser la règle suivante :

$$(\hat{\sigma}_S^M)^2 \leq 90\% \cdot \hat{\lambda}_S^M \Rightarrow N_S^M \sim \mathcal{B}(m^M, \frac{\hat{\lambda}_S^M}{m^M})$$

$$90\% \cdot \hat{\lambda}_S^M < (\hat{\sigma}_S^M)^2 < 110\% \cdot \hat{\lambda}_S^M \Rightarrow N_S^M \sim \mathcal{P}(\hat{\lambda}_S^M)$$

$$(\hat{\sigma}_S^M)^2 \geq 110\% \cdot \hat{\lambda}_S^M \Rightarrow N_S^M \sim \mathcal{NB}(m^M, \frac{\hat{\lambda}_S^M}{m^M})$$

Il nous reste maintenant à estimer la loi qui représente le mieux la variable aléatoire Q_S^M . Pour cela, nous avons utilisé le package **ftdistrplus** de **R** afin de calibrer 11 lois différentes : Weibull (\mathcal{W}), Gamma (\mathcal{GA}), Cauchy (\mathcal{C}), Log-normale (\mathcal{LN}), Pareto (\mathcal{PA}), Gumbel (\mathcal{GU}), Exponentielle (\mathcal{E}), Logistique (\mathcal{L}), Log-Logistique (\mathcal{LL}), Weibull inverse (\mathcal{WI}) et Burr (\mathcal{BU}). Nous présentons les fonctions de densité de ces 11 lois en Annexe D. Nous allons ensuite déterminer les paramètres des lois en maximisant la vraisemblance associée aux observations, et choisir la loi la plus adéquate (celle ayant la plus grande vraisemblance). Nous présenterons, là encore en Annexe E, la méthode de calcul du maximum de vraisemblance pour une loi arbitraire parmi les 11.

Nous présenterons la démarche d'adéquation de loi pour un seul exemple, les autres graphiques seront en Annexe F et nous synthétiserons les résultats dans des tableaux récapitulatifs là aussi placés en Annexe F. Nous prendrons de façon arbitraire la pluviométrie de janvier pour PARIS MONTsouris.

Démarche

La première étape est de déterminer la proportion de demi-journées pluvieuses pour le mois de janvier à PARIS MONTsouris. Commençons par regarder notre historique du nombre de demi-journées de pluie durant le mois de janvier par année :

4. Nous faisons ici l'abus de notation A au lieu de A_S^M par souci de lisibilité.

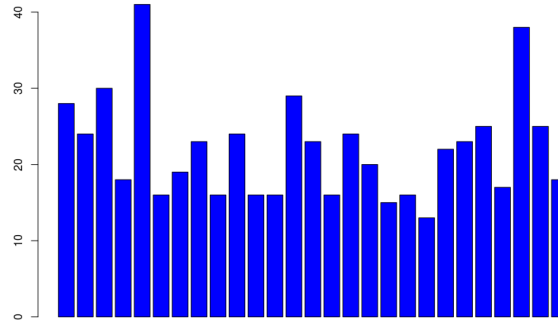


FIGURE 3.17 – Nombre de demi-journées de pluie durant le mois de janvier par année à Paris-Montsouris.

Sur cet échantillon, nous obtenons $\hat{\lambda}_S^M = 22,04$ et $\hat{\sigma}_S^M = 6,84$, soit $(\hat{\sigma}_S^M)^2 = 46,7 > 24,24 = 110\% \cdot \hat{\lambda}_S^M$. Nous considérerons donc une loi $\mathcal{NB}(62, \frac{22,04}{62} \approx 35,5)$.

Une fois que nous avons estimé le paramètre de la loi de N_S^M , il nous reste à trouver la meilleure adéquation de loi optimale pour Q_S^M . Reprenons la densité empirique présentée plus tôt. On observe que la loi est plutôt régulière avec une queue qui reste relativement fine. Afin de déterminer la meilleure adéquation de loi, nous utilisons le package **R fitdistrplus** permettant à la fois de calculer la vraisemblance du modèle ainsi que les paramètres optimaux. On peut ensuite tracer la densité théorique obtenue pour chacune de nos 11 distributions possibles. Nous présentons ci-dessous le résultat de l'adéquation de la loi logistique à la densité empirique de nos observations :

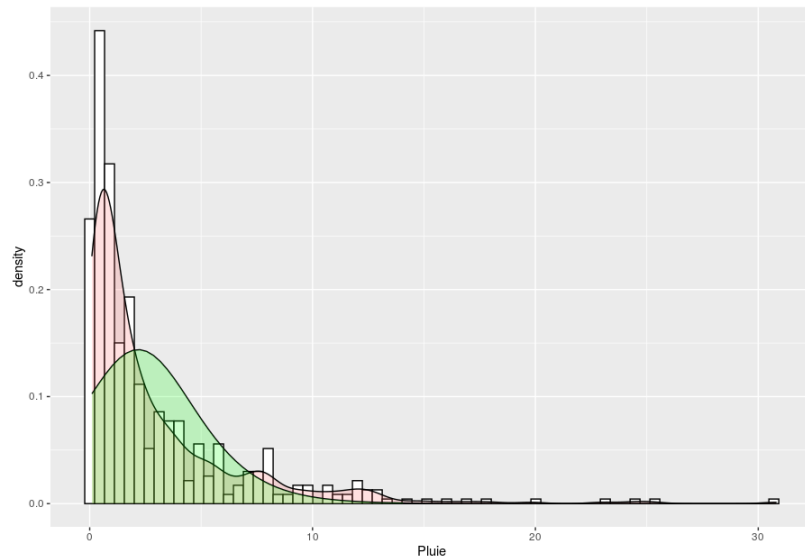


FIGURE 3.18 – Adéquation d'une loi logistique sur les précipitations de janvier à PARIS MONTsouris.

On observe en vert la distribution théorique de la loi logistique et en rouge la distribution empirique. On peut remarquer que la loi logistique capte mal les valeurs extrêmes et surreprésente les valeurs intermédiaires. Nous cherchons alors quelle sera la loi optimale pour nos modélisations. Pour cela, nous regardons la log-vraisemblance de chacun des modèles et nous choisissons la loi avec la log-vraisemblance la plus élevée. Pour le mois de janvier à PARIS MONTsouris, la loi optimale est une log-normale. En effet, on voit qu'elle colle mieux que la loi logistique en tout point par rapport à la distribution empirique.

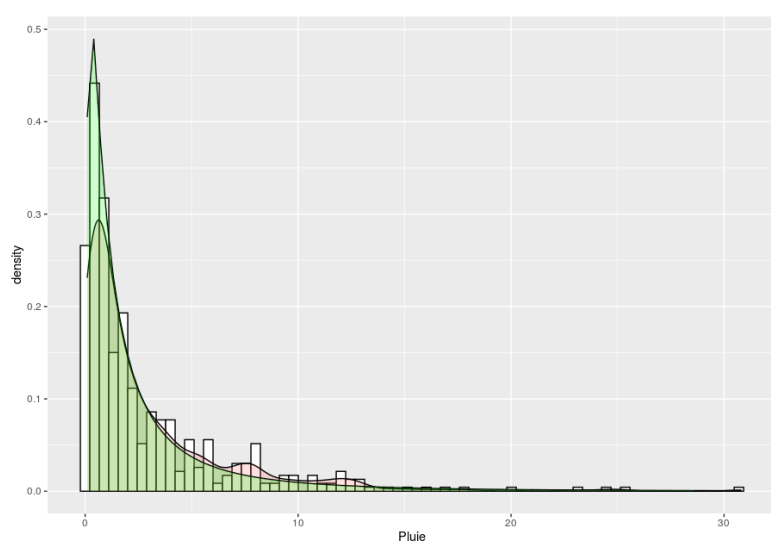


FIGURE 3.19 – Adéquation d’une loi log-normale sur les précipitations de janvier à PARIS MONTsouris.

Nous faisons de même pour chacun des 12 mois de l’année et nous récupérons les 12 meilleures adéquations de lois via le maximum des log-vraisemblances.

Weibull inverse	-1072	-998	-1137	-976	-1091	-868	-805	-878	-819	-1155	-1220	-1372
Weibull	-1065	-983	-1116	-964	-1105	-885	-835	-891	-816	-1161	-1229	-1349
Pareto	-1057	-976	-1109	-962	-1090	-868	-816	-881	-813	-1154	-1210	-1346
Log-normale	-1043*	-965*	-1100*	-952*	-1076*	-856*	-803*	-869*	-803*	-1137*	-1195*	-1334*
Log-Logistic	-1059	-975	-1113	-968	-1090	-867	-814	-881	-816	-1159	-1211	-1355
Logistic	-1371	-1228	-1382	-1213	-1486	-1194	-1166	-1202	-1078	-1520	-1622	-1696
Gumbel	-1257	-1131	-1277	-1122	-1376	-1098	-1074	-1112	-996	-1401	-1494	-1573
Gamma	-1072	-988	-1121	-968	-1122	-901	-854	-905	-825	-1172	-1247	-1357
Exponentielle	-1083	-993	-1127	-977	-1176	-934	-903	-948	-853	-1200	-1278	-1375
Cauchy	-1274	-1141	-1304	-1165	-1299	-1041	-983	-1059	-986	-1417	-1453	-1618
Burr	-1056	-974	-1108	-962	-1089	-867	-813	-881	-812	-1154	-1210	-1345
	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Decembre

FIGURE 3.20 – Tableau des log-vraisemblances des adéquations de lois pour PARIS MONTsouris.

Afin de s'assurer de la pertinence de l'adéquation, nous allons en plus de la méthode du maximum de vraisemblance effectuer des tests statistiques de conformité. Nous avons décidé de mettre en place le test de Kolmogorov-Smirnov (ks), utilisé pour tester des variables aléatoires continues. Nous l'avons défini en Annexe G. Nous allons également tracer le QQ-plot de la distribution théorique versus la distribution empirique afin de comparer les quantiles théoriques et observés. Regardons les résultats pour notre loi log-normale pour le mois de janvier à PARIS MONTsouris :

```

One-sample Kolmogorov-Smirnov test

data: Data[Mois == 1 & Pluie > 0, Pluie]
D = 0.056379, p-value = 0.05909
alternative hypothesis: two-sided
    
```

FIGURE 3.21 – Test de Kolmogorov-Smirnov pour les précipitations en janvier à PARIS MONTsouris.

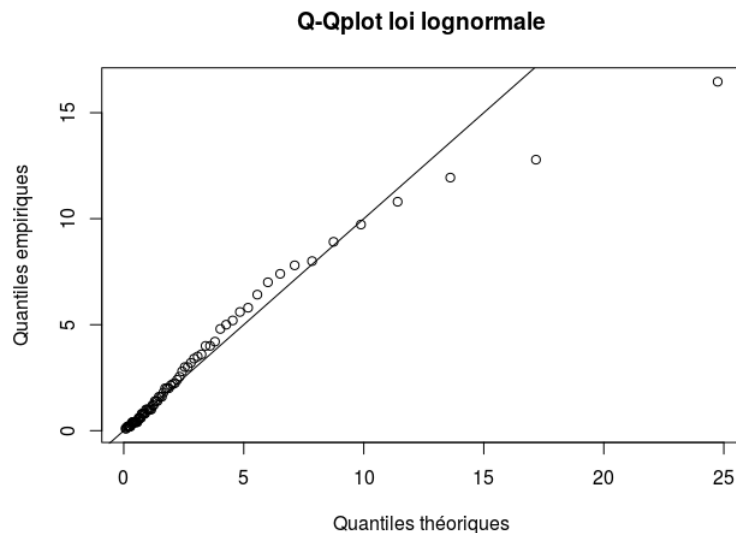


FIGURE 3.22 – QQ-plot de la loi log-normale pour les précipitations en janvier à PARIS MONTsouris.

La p-value du test de Kolmogorov-Smirnov est assez faible mais tout de même au-dessus de 5%. Nous cherchons à avoir la valeur la plus élevée possible, signifiant que l'échantillon observé a une forte probabilité de provenir de la distribution testée. En effet, on remarque sur le QQ-plot que les quantiles élevés sont surestimés par notre loi log-normale comparée à l'échantillon observé. Cela aura pour effet de modéliser des précipitations potentiellement plus extrêmes que l'observé. On préférera ceci à l'inverse par principe de prudence.

Il se trouve que pour PARIS MONTsouris la meilleure loi d'adéquation est toujours la loi log-normale. Ce n'est pas le cas pour toutes les stations, bien que la loi

log-normale ressorte dans une majorité de nos adéquations de loi. Ce résultat est en accord avec plusieurs études, comme notamment une étude sur les précipitations en Suisse [Fallot et Sautebin, 2018] spécifiant que la meilleure loi pour modéliser les pluies dans le pays est la loi log-normale, au même titre que la loi de Gumbel.

Finalement, pour PARIS MONTSOURIS voici nos résultats :

Mois	Loi retenue	$\mu = X$ et $\lambda = Y$
Janvier	Log-normale	$\mu = 0,310$ et $\sigma = 1,298$
Février	Log-normale	$\mu = 0,532$ et $\sigma = 1,256$
Mars	Log-normale	$\mu = 0,556$ et $\sigma = 1,261$
Avril	Log-normale	$\mu = 0,426$ et $\sigma = 1,322$
Mai	Log-normale	$\mu = 0,618$ et $\sigma = 1,522$
Juin	Log-normale	$\mu = 0,458$ et $\sigma = 1,422$
Juillet	Log-normale	$\mu = 0,430$ et $\sigma = 1,492$
Août	Log-normale	$\mu = 0,541$ et $\sigma = 1,521$
Septembre	Log-normale	$\mu = 0,430$ et $\sigma = 1,482$
Octobre	Log-normale	$\mu = 0,352$ et $\sigma = 1,403$
Novembre	Log-normale	$\mu = 0,420$ et $\sigma = 1,374$
Décembre	Log-normale	$\mu = 0,542$ et $\sigma = 1,366$

TABLE 3.3 – Lois optimales pour PARIS MONTSOURIS.

Résultats

Nous présentons tous nos résultats agrégés pour chaque station, à savoir la loi pour N_S^M ainsi que son/ses paramètre(s) puis la loi optimale pour Q_S^M avec ses paramètres associés. Du fait de la grandeur du tableau récapitulatif, nous l'avons placé en Annexe F.

3.2.4 Mesures de dépendances

Afin de vérifier si les précipitations relevées sur le bassin sont bien corrélées, nous avons évalué leurs dépendances au travers de calculs de corrélation. De la même façon que pour nos hauteurs d'eau entre stations, nous avons été contraints d'harmoniser nos plages de mesures afin d'obtenir des valeurs de précipitations pour chacune des stations sur une plage de temps commune t_0, \dots, t_n . Une fois ce travail d'interpolation réalisé, nous avons calculé les corrélations entre chacune des séries. Il existe différentes mesures de dépendances permettant de mettre en évidence des corrélations diverses. La plus connue est le coefficient de Pearson aussi appelé coefficient de corrélation linéaire. Cependant comme son nom l'indique il ne mesure que des dépendances linéaires et peut donc être proche de zéro entre deux variables alors qu'il existe un lien non linéaire entre ces deux variables. C'est pour cette raison qu'ont été introduits (τ) de Kendall et le rhô (ρ) de

Spearman, basés sur une corrélation des rangs sur un échantillon donné des deux variables considérées. Nous avons défini mathématiquement tous ces indicateurs en Annexe H.

Résultats

Nous avons calculé les corrélations avec chaque indicateur (Pearson, Spearman, Kendall) entre chacune des neuf stations et pour chaque mois de l'année. Nous avons donc récolté 26 tableaux de corrélations que nous mettrons en Annexe I pour une partie d'entre eux. Voici un exemple pour le coefficient de corrélation de Pearson sur la pluviométrie historique de janvier entre chacune de nos stations :

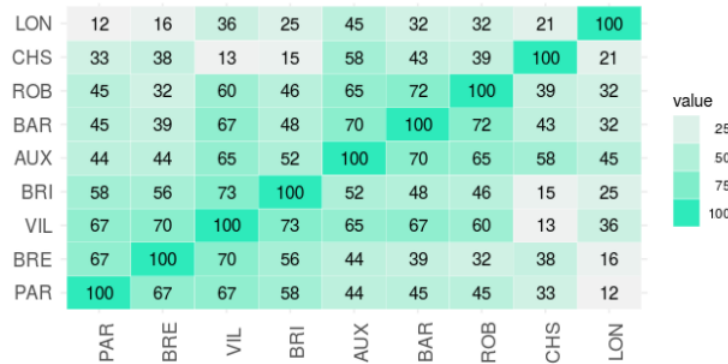


FIGURE 3.23 – Corrélations de Pearson entre les stations ISD sur le mois de Janvier (%).

On remarque que les corrélations sont très élevées entre la plupart des stations. Les corrélations les plus fortes relevées pour janvier sont entre Auxerre et Barberey par exemple. On observe bien la dépendance géographique car dans l'ensemble, les stations les plus proches les unes des autres sont les plus corrélées.

3.2.5 Copules multivariées

Les phénomènes météorologiques sont des phénomènes présentant une dynamique spatio-temporelle et sont régis par des lois très complexes impossibles à modéliser de façon triviale. Ils font toujours l'objet de nombreuses études et sont encore peu maîtrisés de nos jours. Nous savons que ces phénomènes se déplacent dans le temps et ce de façon non uniforme (pas toujours dans la même direction, ni à la même vitesse). Ainsi, les corrélations géographiques et temporelles ne sont pas fixes. La mise en place de modèles météorologiques complexes sort du cadre de ce mémoire et pourrait faire l'objet d'un mémoire entier. Dans un souci de compromis entre complexité et qualité des modèles, nous avons entrepris la modélisation individuelle de la pluviométrie de plusieurs stations couplée à une structure de dépendance afin de capter le maximum des caractéristiques

météorologiques. Nous avons ainsi entrepris une modélisation simplifiée des précipitations tout en essayant de refléter au maximum la réalité. Pour cela, nous avons étudié les historiques communs de nos stations pour chacun des mois de l'année et nous avons mis en place une structure de dépendance entre les variables aléatoires $P_{S_1}^M, P_{S_2}^M, \dots, P_{S_9}^M$. Ce travail est pertinent car nous avons pu observer de très fortes corrélations entre nos données historiques. Pour cela, nous avons besoin de trouver les relations décrites par les distributions conjointes. Cependant, retrouver la loi conjointe s'avère très difficile. Aussi, nous avons utilisé des copules couplées aux lois marginales des variables (travail réalisé dans la partie "Calibrage"). Cela nous permet ainsi de considérer de manière distincte la structure de dépendance décrite par la fonction de distribution conjointe et le comportement marginal de ces variables, qui comme nous avons pu le voir ne suivent pas toujours les mêmes lois. Nous avons défini tous les prérequis mathématiques nécessaires à la compréhension de cette partie en Annexe H. Par souci de simplification, nous avons décidé d'utiliser des copules gaussiennes, qui sont plus triviales à mettre en place, bien qu'elles ne constituent pas forcément les copules les plus adaptées, et un travail de recherche sur l'optimisation de l'adéquation de copules pourrait être réalisé afin de déterminer la famille de copules la plus adaptée.

Résultats

De la même façon que pour les parties précédentes, nous ne présenterons pas tous les résultats dans le corps de ce rapport par souci de concision. De plus, la dimension des copules paramétrées est de $d > 2$. Il est donc impossible de la représenter graphiquement. Nous nous limiterons donc à quelques graphiques en deux dimensions. La définition d'une copule gaussienne est en Annexe H, mais ce qu'il faut savoir est que son paramétrage est simple, il suffit d'avoir le tableau de corrélation de Pearson entre les différentes variables considérées. Ainsi, pour paramétrer la dépendance entre toutes nos stations nous utilisons le tableau 3.23, et pour paramétrer la dépendance entre un sous-ensemble de nos stations il suffira de prendre le sous-tableau correspondant. Nous pouvons faire de même avec les variables lagées afin de corrélérer la pluviométrie décalée géographiquement dans le temps.

Afin de simuler notre vecteur $Z = (p_{S_1,]T; T_{i+1}]}^M, \dots, p_{S_9,]T; T_{i+1}]}^M)$ de précipitations pour un intervalle $]T, T + 12h]$, nous commençons par simuler une réalisation $u = (u_1, \dots, u_9)$ de la copule gaussienne C que nous avons calibré. Nous obtenons alors notre vecteur par la transformation : $Z = (F_{S_1}^{-1}(u_1), \dots, F_{S_9}^{-1}(u_9))$ où les fonctions $F_{S_1}^{-1}$ sont les inverses des fonctions de répartition des lois des précipitations.

Nous pouvons maintenant simuler dans un mois donné, les précipitations conjointes aux neuf stations pour n'importe quelle demi-journée et en y affectant des lags si nécessaire.

3.3 Mise en place d'un modèle explicatif

3.3.1 Méthodologie

La dernière étape du module aléa est l'étape principale, qui consiste à déterminer la hauteur d'eau au niveau d'Austerlitz durant une année complète, qui représentera une simulation annuelle d'évolution du fleuve. Pour cela, nous allons caractériser la dépendance entre les précipitations tombées dans le bassin versant lagguées d'un temps adapté, correspondants aux valeurs trouvées lors de l'étude de la dynamique du fleuve. Cependant, expliquer la hauteur d'eau du fleuve directement avec les précipitations du bassin n'est pas pertinent, car le fait qu'il pleuve énormément à un instant t ne veut pas dire que le fleuve va être en crue quelques instants plus tard, mais veut dire que le fleuve va monter de niveau. Prenons un exemple concret. Supposons que nous sommes en août avec un niveau d'eau très bas du fait de l'absence de précipitations depuis des semaines. Si subitement des averses torrentielles s'abattent sur le bassin versant, la hauteur d'eau relevée à Austerlitz peut passer de 0,5 mètre à 3,5 mètres en quelques jours, ce qui témoigne d'une évolution très rapide (on rappelle que le niveau de la Seine à Austerlitz augmente rarement de beaucoup plus de 50cm par jour même en cas d'averses extrêmes). Ainsi, l'évolution est forte du fait des fortes pluies mais le fleuve n'est pas pour autant en crue. A l'inverse, lorsque le fleuve est déjà en période de crue (6 mètres par exemple) mais que les précipitations sont nulles dans tout le bassin à un instant T , cela ne veut pas dire qu'à l'instant $T+3$ jours ou $T+4$ jours le fleuve sera à un niveau bas. Ce sont les évolutions de la hauteur d'eau qui sont déterminées par les précipitations du bassin versant. Ainsi, notre variable réponse Y n'est pas la hauteur d'eau relevée à Austerlitz à un instant T mais la différence entre la hauteur d'eau à cet instant T par rapport à la hauteur d'eau relevée à $T-12h$. Représentons graphiquement la différence entre deux relevés consécutifs à AUSTERLITZ :

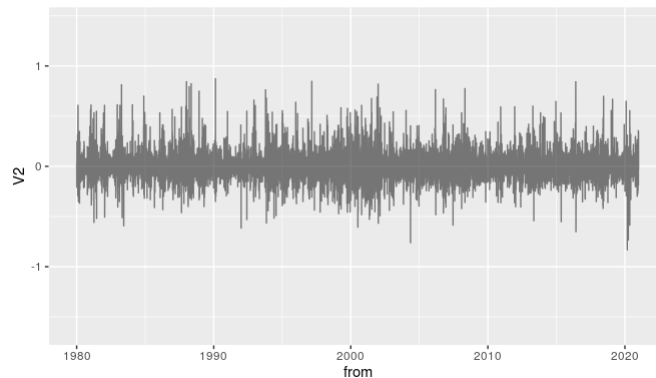


FIGURE 3.24 – Différence de hauteur d'eau entre deux relevés consécutifs à Austerlitz.

Ainsi, à partir de ce modèle, il nous suffit d'avoir un point de départ H_0 de la hauteur d'eau initiale, puis l'on calcule la hauteur du fleuve par demi-journée à partir de la valeur précédente et de la sortie de notre modèle. Pour cela, il y a la nécessité de partir d'une

valeur initiale H_0 pour chaque simulation. Nous avons fait le choix de la fixer à la moyenne historique du mois de janvier.

Ainsi, nous ne modélisons pas la hauteur d'eau de janvier à décembre mais de août à juillet, ce qui ne changera rien à nos calculs de pertes brutes. Le choix du mois de départ a été effectué en minimisant l'écart-type des valeurs mensuelles des hauteurs d'eau mensuelles historiques. L'idée est de partir d'un mois "stable" et très peu propice à l'apparition d'une crue afin de ne pas louper des potentielles crues sur ce mois-ci, rendu compliquée par une valeur initiale d'ordre de la moyenne annuelle du fleuve (autour de 1 mètre). Ce n'est pas un problème car nous ne sommes pas dans un cadre de prévision des crues mais dans un cadre d'estimation des périodes de retour et des pertes potentielles. Ainsi, nous avons décidé de prendre pour valeur initiale la moyenne historique du mois de janvier.

Maintenant que la méthodologie est complète, il nous reste à mettre en place nos modèles explicatifs. Nous avons mis en place deux types de modèles différents afin de caractériser la différence entre les hauteurs d'eau relevées à Austerlitz par pas de 12h avec la pluviométrie antérieure du bassin versant.

3.3.2 Modèle linéaire (LM)

Théorie

Le modèle linéaire (LM) est un modèle de prédiction facile à mettre en place car relativement simple. C'est un modèle qui permet d'exprimer une variable Y en fonction d'une combinaison linéaire de variables explicatives X . Il s'écrit mathématiquement :

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_p x_{i,p} + W_i$$

Avec :

- y_i : i^{ème} modalité de la variable à expliquer
- $x_{i,j}$: i^{ème} modalité de la variable X_j
- X_j : j^{ème} variable explicative avec $j \in \{1, \dots, p\}$
- β_j : coefficient de la j^{ème} variables explicatives
- W_i : variable aléatoire i.i.d centrée réduite

On peut également l'écrire sous forme matricielle afin de simplifier les notations :

$$Y = X\beta + W$$

Avec :

- Y : variable à expliquer
- X : matrice des variables $(X_j)_{j=1\dots p}$
- β : matrice des coefficients $(\beta_j)_{j=1\dots p}$
- W : vecteur des réalisations de W_i

La problématique est alors d'estimer le vecteur β qui caractérise le lien entre chacune des variables explicatives (les pluviométries) et la variable réponse (la hauteur d'eau à Austerlitz). Nous cherchons la valeur de β qui va optimiser la vraisemblance du modèle. Nous admettrons ici (la démonstration est en Annexe J) sa valeur comme étant :

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Modélisation

Nous avons donc mis en place un modèle linéaire afin d'expliquer l'évolution par demi-journée de la hauteur d'eau du fleuve en fonction des précipitations laggées relevées dans les différentes parties du bassin. Ainsi, pour chaque station, nous possédons les relevés bruts et laggés de 1 à 12 demi-journées. Dans un premier temps, et pour chaque station, nous avons retenu uniquement les lags correspondant à ceux observés pour les stations *HYDRO* voisines dans leur zone d'influence (cellule de Voronoï), pour obtenir un intervalle probable de lag optimal.

Station	PAR	BRE	VIL	BRI	AUX	BAR	ROB	CHS	LON
Intervalle lag	0	0 à 1	1 à 3	5 à 8	4 à 7	3 à 5	6 à 10	4 à 7	10 à 12

Intervalles de lags par station.

Ensuite, nous retirons par station la variable explicative la moins pertinente au sens de la p-value. Nous obtenons alors un modèle avec une variable explicative par station, dont voici la sortie **R** :

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.799e-02  7.986e-03  -2.252 0.025056 *
PAR          -1.191e-02  6.479e-03  -1.838 0.067040 .
BRE           1.213e-02  7.857e-03   1.543 0.123862
VILLag1      5.019e-03  1.292e-03   3.885 0.000126 ***
VILLag2      6.672e-04  1.145e-03   0.583 0.560584
VILLag3     -3.874e-03  1.538e-03  -2.519 0.012299 *
BRILag5     -1.230e-03  9.068e-04  -1.356 0.176112
BRILag6     -1.473e-03  1.410e-03  -1.045 0.296867
BRILag7     -3.230e-05  9.216e-04  -0.035 0.972065
BRILag8     -9.386e-04  1.417e-03  -0.662 0.508307
AUXLag4      2.432e-03  2.037e-03   1.194 0.233623
AUXLag5     -7.245e-05  1.411e-03  -0.051 0.959094
AUXLag6      1.243e-03  2.264e-03   0.549 0.583545
BARLag3     -5.069e-04  1.824e-03  -0.278 0.781298
BARLag4      1.425e-04  1.605e-03   0.089 0.929330
BARLag5      7.524e-04  1.988e-03   0.379 0.705268
ROBLag6     -8.940e-04  1.467e-03  -0.609 0.542728
ROBLag7      1.461e-03  1.765e-03   0.828 0.408335
ROBLag8      1.568e-03  1.376e-03   1.140 0.255269
ROBLag9      2.275e-03  1.521e-03   1.496 0.135706
ROBLag10     8.227e-04  1.349e-03   0.610 0.542556
CHSLag4      2.269e-03  9.950e-03   0.228 0.819764
CHSLag5      2.113e-04  9.452e-03   0.022 0.982175
CHSLag6      1.283e-02  9.486e-03   1.353 0.177199
CHSLag7     -2.341e-03  9.426e-03  -0.248 0.803997
LONLag10     1.217e-03  1.725e-03   0.706 0.480931
LONLag11     1.598e-03  1.587e-03   1.007 0.314769
LONLag12    -1.504e-03  1.415e-03  -1.063 0.288634
    
```

FIGURE 3.25 – Sortie **R** du premier LM.

On observe que certaines stations possèdent une p-value supérieure à 5%, ce qui nous conduit à devoir les retirer. Ce processus *backward* consiste à retirer les variables une à une jusqu'à tomber sur un modèle viable. Pour cela nous nous basons sur la p-value et on retire donc la station CHATILLON laggé de cinq demi-journées. *In fine*, après avoir retiré toutes les données des stations PARIS, BRIENNE, AUXERRE et LONGVIC, nous obtenons le modèle suivant :

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0079835  0.0020960  -3.799 0.000148 ***
BRE          0.0042848  0.0009003   4.760 2.04e-06 ***
VILLag1     0.0031574  0.0004162   7.587 4.39e-14 ***
BARLag3     0.0007295  0.0004214  -1.731 0.033550 *
ROBLag10    0.0005436  0.0003512   2.402 0.016368 *
---

```

FIGURE 3.26 – Sortie **R** du LM final.

On remarque que finalement, le modèle n'a pas conservé PARIS MONTSOURIS, au détriment de la station BRETTIGNY. Cela peut s'expliquer par le fait que BRETTIGNY est tout juste en amont de Paris et donc aura potentiellement plus tendance à mieux capter les précipitations influant sur la hauteur d'eau du fleuve, là où les pluies relevées à PARIS MONTSOURIS tombent aussi en aval de la station *HYDRO* d'Austerlitz. On a donc la confirmation que les pluies tombant en amont du bassin ne sont pas négligeables, bien au contraire. La p-value pour la station BARBEREY est à la limite du 5%, nous décidons de la conserver tout de même. On peut interpréter l'*Intercept* du modèle comme l'évolution moyenne du fleuve en l'absence de pluie, qui est d'environ -0,008 soit 0,8cm. On peut remarquer que les stations globalement les plus éloignées sont sorties de la modélisation hors ROBINSON. Ce résultat n'est finalement pas si étonnant que cela pour plusieurs raisons.

La première correspond à la corrélation entre des stations. La complexité du modèle vient du fait que l'on doit avoir des stations assez rapprochées afin de capter au maximum l'ensemble de la pluviométrie du bassin mais qu'elles ne doivent pas être trop proches sous peine d'effectuer des doubles comptages de pluie. En effet, deux stations trop proches risquent d'observer une pluviométrie très semblable et donc le modèle va avoir tendance à retirer une des deux variables car elles sont trop corrélées. Cela explique qu'entre PARIS MONTOURIS et BRETTIGNY, une seule des deux stations soit conservée par le modèle.

La seconde raison est que les stations les plus éloignées ont tendance à être moins explicatives au vu du modèle. Cela paraît logique, car s'il pleut une journée donnée à la source du bassin et nulle part ailleurs pendant les jours qui suivent, l'onde de crue aura tendance à être dissoute avant qu'elle atteigne le bassin parisien. On le voit dans la valeur des coefficients β_i estimés. Au plus la station est éloignée de Paris, au plus le coefficient et donc l'influence des pluies est faible.

On notera que la sortie de notre modèle influencera nos simulations pour nos précipitations. Ainsi, nous n'avons besoin de simuler que les précipitations tombant dans les bassins de BRETTIGNY, VILLAROCHE, BARBEREY et ROBINSON. La structure de dépendance qui sera mise en place entre ces stations sera donc entre les séries laggées,

nous utiliserons donc le tableau de corrélation de Pearson entre les données historiques de BRETTIGNY, VILLAROCHE laggée d'une demi-journée, BARBEREY laggée de trois demi-journées, et ROBINSON laggée de 10 demi-journées, conditionnellement au mois de l'année considéré. Voici le tableau correspondant pour le mois de janvier :

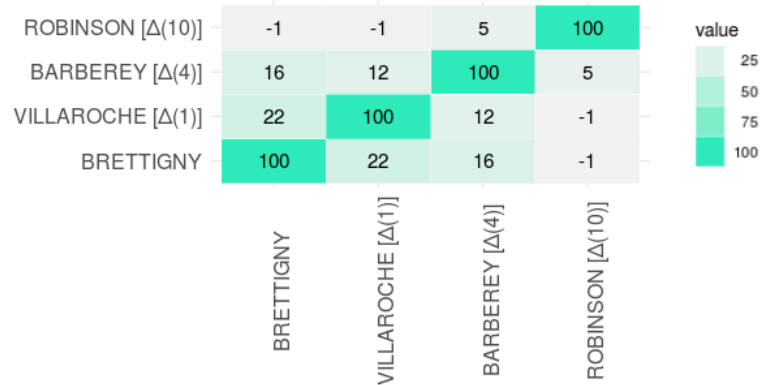


FIGURE 3.27 – Dépendance linéaire entre les stations du modèle LM avec lags.

On remarque alors que la dépendance entre ROBINSON et les autres stations est plutôt faible. Cela paraît logique car un lag de 10 correspond à cinq jours de décalage, ce qui semble beaucoup pour un épisode pluvieux qui généralement se déplace plus rapidement.

3.4 Mise en place des simulations de la hauteur d'eau

Nous possédons maintenant tous les outils afin de connaître la distribution des pertes annuelles correspondant à notre portefeuille d'assurés. Pour cela, nous allons effectuer 10 000 simulations. Pour chacune de ces simulations, il nous faut dans un premier temps simuler nos précipitations annuelles par station. Nous utilisons alors nos lois ainsi que la structure de dépendance déterminées plus tôt, et nous simulons les pluies par pas de 12h sur l'ensemble du bassin. Ensuite, nous appliquons notre modèle afin de déterminer la hauteur d'eau à la station d'Austerlitz par demi-journées sur l'ensemble de l'année. Enfin, nous utilisons notre modèle de vulnérabilité et notre fonction de perte pour calculer les pertes brutes totales pour chacune des simulations. Une fois ce travail réalisé, nous sommes en mesure de représenter la distribution des pertes brutes annuelles. Nous allons également effectuer un calcul de période de retour pour chacune des différentes hauteurs d'eau, en portant une attention particulière à la période de retour de la crue de 1910. Tout cela permet d'avoir une vision complète du risque de crue de la Seine.

3.4.1 Simulation de la pluviométrie annuelle

Nous allons présenter un exemple de simulation, puis nous présenterons ensuite directement la distribution des pertes brutes annuelles (simulations agrégées). Commençons

par simuler nos précipitations sur une année entière pour les stations de notre modèle via la méthodologie que nous avons explicitée plus tôt.

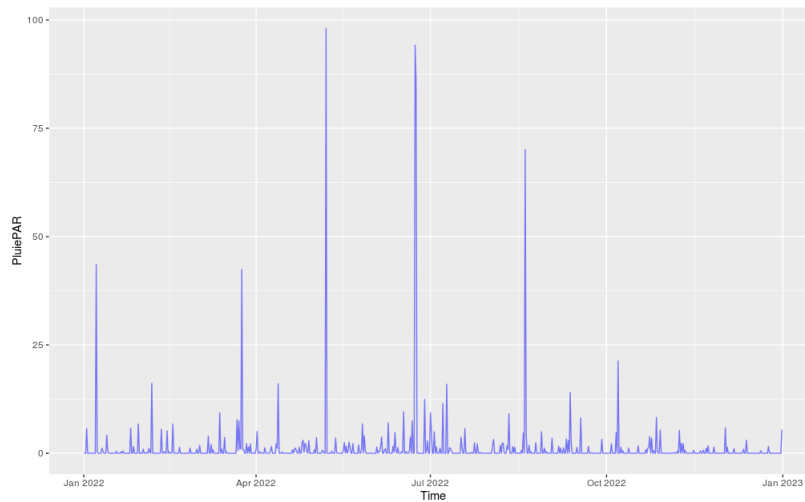


FIGURE 3.28 – Simulation annuelle des précipitations à BRETTIGNY.

On remarque pour cette simulation que l'on a obtenu peu de pluie durant l'automne. Néanmoins, on peut observer plusieurs pics de précipitations en mai et en juin, ce qui pourrait entraîner potentiellement une crue de la Seine. Observons la pluviométrie simulée avec une autre station du modèle (VILLAROCHE) :

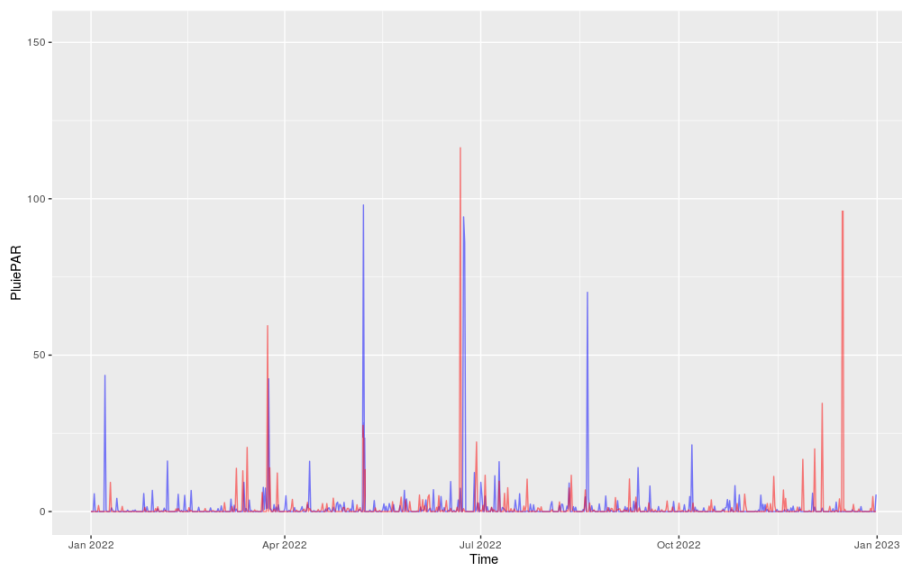


FIGURE 3.29 – Simulation annuelle simultanée des précipitations à BRETTIGNY et VILLAROCHE.

Le mois de juin a en effet observé un gros épisode de pluie sur tout le bassin lors de cette simulation. Il nous reste maintenant à observer la dynamique de la hauteur d'eau suite à ces précipitations et regarder si l'on observe une crue à cette période.

3.4.2 Étude de l'hydrométrie d'Austerlitz

Une fois que nous possédons l'intégralité de la pluviométrie annuelle des neuf stations modélisées, nous utilisons notre modèle explicatif afin d'en déduire la hauteur d'eau relevée à la station d'Austerlitz toujours par pas de 12h sur l'année simulée. Nous partons de la moyenne historique observée en janvier, puis nous calculons les évolutions de la hauteur d'eau toutes les 12h grâce à notre modèle explicatif. Afin de représenter au mieux la réalité, nous avons ajouté deux règles empêchant de modéliser des incohérences avec le monde réel, rendues possible par l'application de lois aléatoires. Nous voulons modéliser des événements rares mais pas improbables. Ainsi, nous avons limité la montée/descente du niveau de l'eau à 1 mètre sur une même demi-journée. Cette hypothèse est raisonnable au vu de l'historique observé. Nous savons que la Seine est un fleuve à crues lentes, jamais le fleuve ne peut monter ou descendre aussi rapidement. Si l'on regarde les précédents records, nous sommes autour de 0,75 mètres sur une demi-journée. La seconde hypothèse effectuée est que le fleuve ne peut jamais descendre sous -0,5m. Cela correspond environ à son niveau le plus bas historique et atteignable, le fleuve est alors totalement sec et le niveau d'eau ne peut pas descendre en dessous.

Pour illustrer la démarche, nous représentons une simulation ayant générée une crue de 6m10, soit la hauteur relevée lors de la crue de 2016.

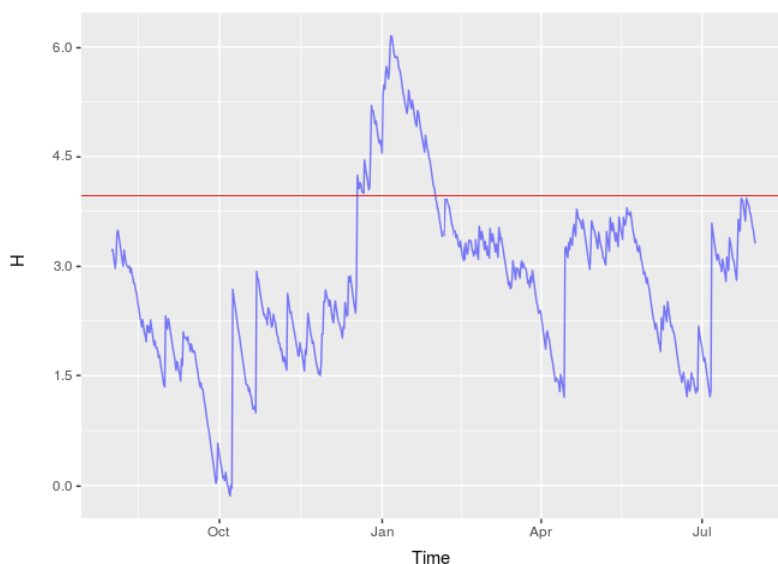


FIGURE 3.30 – Exemple de simulation annuelle de la hauteur d'eau à Austerlitz.

Nous pouvons également tracer quelques trajectoires sur le même graphique afin

d'observer des potentielles tendances :

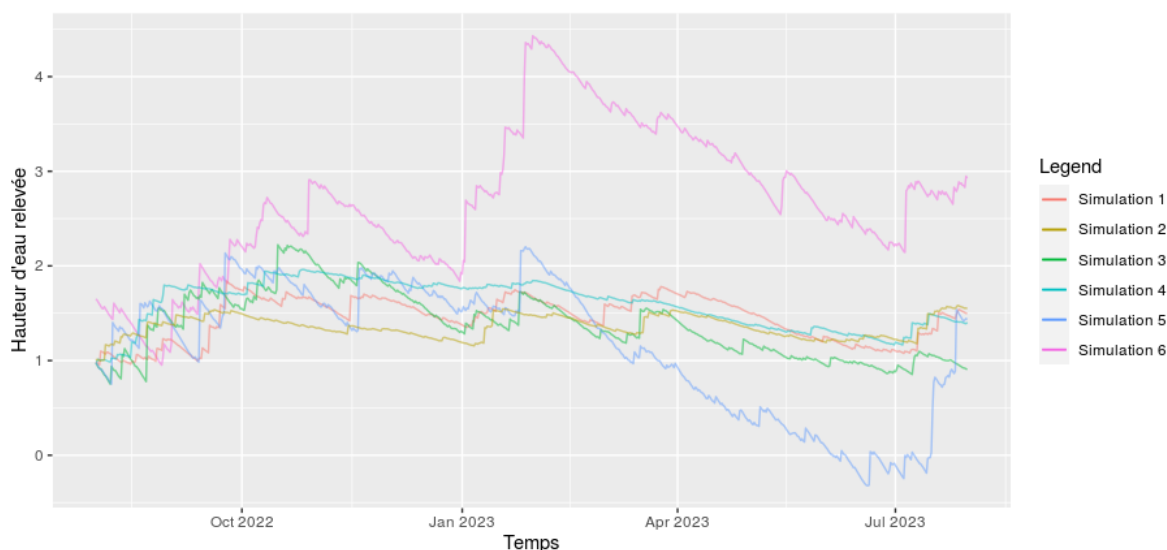


FIGURE 3.31 – 6 simulations annuelles de la hauteur d'eau à Austerlitz.

On remarque que les trajectoires sont bien différentes les unes des autres. On peut tout de même observer une tendance globale saisonnière, avec une hauteur d'eau généralement plus élevée en hiver qu'au début de l'été. On notera tout de même quelques limites à notre modèle, malgré les règles expliquées plus tôt. Le rôle des lacs réservoir et de leur gestion par l'EPTB n'est pas modélisé. Ainsi, lorsque le fleuve est en décrue et tend à atteindre l'étiage, l'EPTB libère l'eau des lacs, ce qui fait gonfler artificiellement la hauteur de l'eau. Il en est de même dans l'autre sens lors de l'apparition de crue. Cet effet est négligé dans nos simulations car nous savons qu'il possède un faible impact dans l'apparition des crues.

3.5 Étude des pertes brutes et des périodes de retour

3.5.1 Calcul des pertes brutes associées à chaque simulation

Maintenant que nous possédons la dynamique du fleuve sur l'année complète, nous calculons ainsi les pertes totales réalisées sur l'année. Pour cela, nous considérons les crues supérieures à 4 mètres relevées dans l'année, nous déterminons les expositions touchées grâce à notre module d'exposition et nous évaluons les pertes associées grâce à notre module de vulnérabilité et à nos taux de destruction dépendants de la durée de la crue comme développé dans le chapitre précédent. Pour l'exemple présenté ci-dessus, l'année simulée a connu une unique crue de 6 mètres environ sur une durée de trois jours au-dessus de la barre des 6 mètres et impliquant des pertes d'environ 40 millions d'euros.

3.5.2 Évaluation de la distribution des pertes brutes annuelles

Effectuer un grand nombre de simulations via notre modèle permet d'obtenir une distribution des pertes financières annuelles sur notre portefeuille assuré du fait de la spécificité de la trajectoire pour chaque simulation, ce qui demeure plus précis que de simples périodes de retour de crues avec leurs pertes associées. Nous sommes en mesure de calculer ainsi les courbes OEP (pertes maximales annuelles) et AEP (pertes totales annuelles) afin de visualiser la distribution de charges. Ce modèle pourra alors par exemple être utilisé pour challenger la formule standard dans le cadre de la directive Solvabilité II, mettre en place des scénarios de stress en vue de la production d'un ORSA, ou pour optimiser le programme de réassurance de traités non proportionnels d'une société d'assurance. Pour le portefeuille que nous possédons, voici la distribution des pertes totales annuelles que l'on obtient :

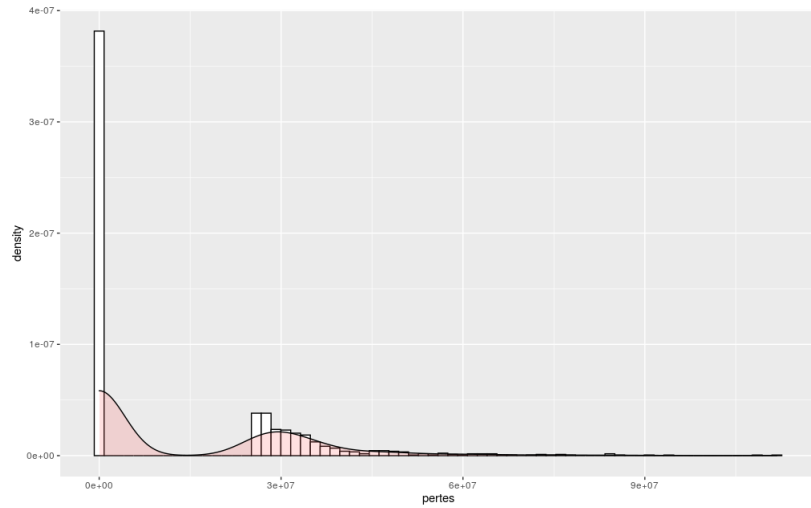


FIGURE 3.32 – Distribution des pertes annuelles pour notre portefeuille assuré (10 000 simulations).

On observe de façon logique une forte représentation de zéro synonyme d'une absence de crue. Nous avons considéré qu'en dessous de 4 mètres, il n'y aurait aucun impact financier. Ces valeurs nulles correspondent aux années sans crue majeure de la Seine, c'est-à-dire aux années où la hauteur de la Seine à Austerlitz n'a pas dépassé 4 mètres. Nous pouvons représenter la distribution de charges en retirant les années sans pertes afin d'y voir plus clair.

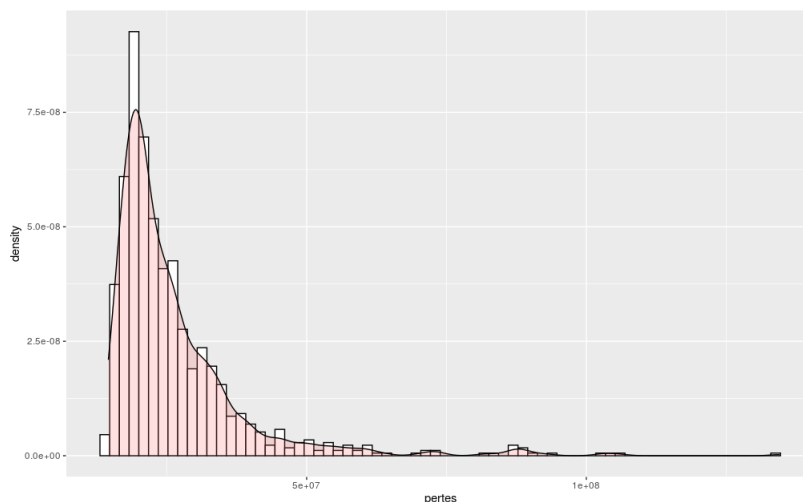


FIGURE 3.33 – Distribution des pertes annuelles pour notre portefeuille assuré pour les simulations ayant entraîné au moins une crue.

Cette vision n'est cependant pas optimale, car nous ne pouvons pas apprécier la probabilité de subir un certain montant de pertes. Il faut donc regarder la fonction de répartition empirique associée.

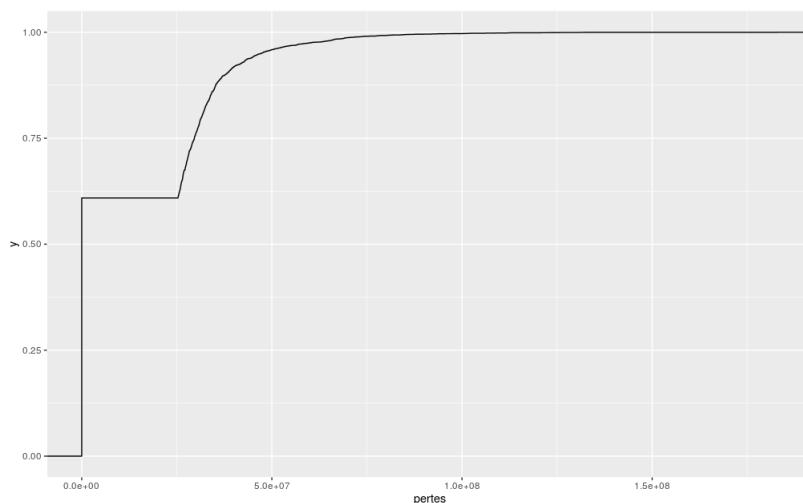


FIGURE 3.34 – Fonction de répartition empirique des pertes annuelles pour notre portefeuille assuré.

Maintenant que l'on connaît la fonction de répartition de nos pertes brutes annuelles, il est intéressant d'y associer les périodes de retour. Cela nous donne une vision du risque porté par l'assureur. Dans le cadre de Solvabilité II, un assureur doit être en mesure de faire face aux pertes associées à un événement bicentenaire. Ces périodes de retour se calculent simplement en inversant la fonction de survie. Voici ce que l'on obtient :

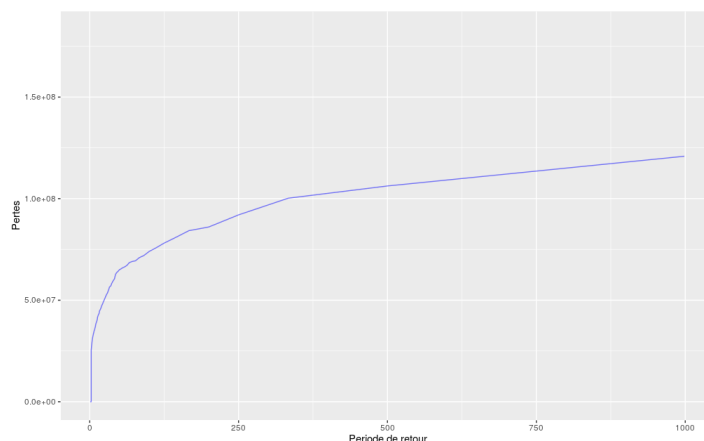


FIGURE 3.35 – Pertes brutes annuelles et périodes de retour.

Pour notre portefeuille assuré, on peut alors évaluer notre perte bicentenaire à environ 88 millions d'euros.

3.5.3 Évaluation des périodes de retour

Afin de quantifier le risque d'occurrence des crues de la Seine en région parisienne, il est intéressant de s'intéresser également aux périodes de retour, c'est-à-dire au temps statistique entre deux occurrences d'une même intensité donnée (hauteur d'eau). Ainsi, une période de retour bicentenaire correspond à un évènement se produisant en moyenne tous les 200 ans ou ayant une chance sur 200 de se produire chaque année. Nous avons tenté d'évaluer ces périodes de retour en effectuant 50 000 simulations de la hauteurs d'eau à Austerlitz sur une année et en comptabilisant les crues qui apparaissent. Nous comparons également nos résultats aux périodes de retour empiriques basés sur la plage de temps 1649-2016⁵ (figure 1.21). Voici les résultats que nous avons obtenus avec notre modèle LM :

Hauteur d'eau (m)	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9
PdR empirique ⁶ (a)	5	7	11	19	31	41	124 ⁶	186 ⁶	∞^6
PdR estimée ⁷ (a)	5	8	12	21	45	104	168	357	1250
Pertes estimées (M€)	31,5	35,2	39,7	47,8	63,4	74,5	84,3	101	126

TABLE 3.4 – Périodes de retour et pertes selon la hauteur d'eau.

5. On notera que les périodes de retour des crues supérieures à 8 mètres sont artificiellement diminuées par le fait de considérer une plage d'historique débutant en 1649 et non pas plus tôt. Cela est cependant nécessaire pour estimer les périodes de retour des crues moindres, car nous prenons la profondeur d'historique maximale. En effet, il est impossible de capturer avec précision des périodes de retour dont la valeur coïncide avec l'ordre de grandeur de la profondeur d'historique.

6. PdR = Période de retour. Les périodes de retour sont exprimées en années.

Période de retour	5	10	20	30	50	100	200	500	1000
Hauteur d'eau (m)	5,0	5,8	6,4	6,8	7,1	7,5	8,2	8,7	8,9
Pertes estimées (M€)	31,5	37,7	47,1	53,6	65,0	74,1	86,0	106	121

TABLE 3.5 – Hauteur d'eau et pertes selon les périodes de retour.

On peut remarquer que la période de retour de 200 ans est au niveau 8,2 mètres. Selon notre modèle, la thèse présentant la crue de 1910 comme centennale est réfutée. Afin de bien observer la relation hauteur d'eau / période de retour, nous pouvons tracer le graphique de la hauteur d'eau en fonction de la période de retour considérée :

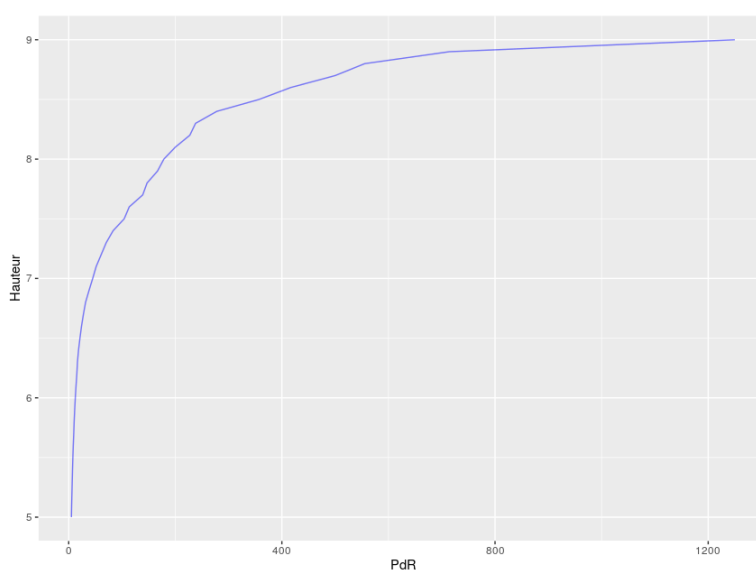


FIGURE 3.36 – Hauteur d'eau et périodes de retour.

3.5.4 Cartographie des périodes de retour

Une fois que nous avons nos périodes de retour, nous pouvons faire une cartographie des périodes de retour qui permet selon la zone dans laquelle se trouve l'exposition, de connaître la période de retour approximative d'une crue atteignant le bien assuré :

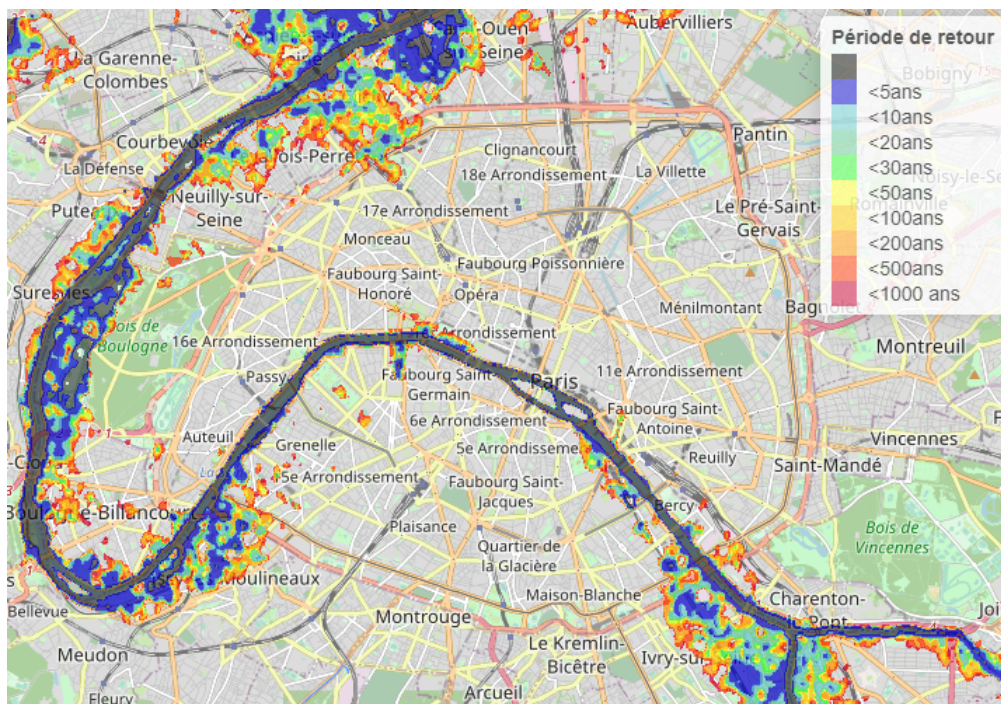


FIGURE 3.37 – Cartographie des zones sous risque pour les périodes de retour estimées.

3.6 Scénarios d'application du changement climatique

3.6.1 Contexte

Le GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat) a publié en ce mois d'août 2021 la première partie de son sixième rapport alarmant sur le climat [GIEC, 2021] approuvée par 234 auteurs et 195 gouvernements. Il constitue la plus importante contribution scientifique sur le climat et sera prise en compte lors de la COP26 en novembre 2021. L'organisme pointe très clairement du doigt l'origine humaine du dérèglement climatique et promet une augmentation de la fréquence et de l'intensité des phénomènes extrêmes. Chaque augmentation de température, aussi petite soit-elle, a une incidence sur l'apparition d'évènements catastrophiques. Pour ce qui est de l'inondation, les experts expliquent que les quantités de précipitations évolueront différemment selon les régions du globe, rendant les zones arides encore plus arides et les zones humides encore plus humides. Ils prévoient également une augmentation de la fréquence et de l'intensité des épisodes pluvieux extrêmes, augmentant donc de façon inéluctable le risque de crue. On peut citer les exemples récents de l'Allemagne ou de la Chine qui ont subis des inondations meurtrières et causant des dégâts considérables. Le WWA⁷ a publié un rapport à la suite des inondations survenues en Europe et plus particulièrement en Allemagne en juillet 2021 [WWA, 2021]. *"Le changement climatique a accru*

la probabilité, mais également l'intensité [des évènements de juillet]." rapporte la WWA. Selon les spécialistes, la survenue d'un tel épisode extrême sur ces régions a été rendue jusqu'à neuf fois plus probable à cause du réchauffement climatique.

Le GIEC a publié cinq nouveaux scénarios d'évolution du climat d'ici 2100, qui dépendent de la capacité des pays à diminuer drastiquement leurs émissions de CO₂ et ce le plus vite possible. Le plus optimiste des scénarios limiterait l'augmentation de la température à 1,5°C d'ici 2100, tandis que le scénario le plus pessimiste prévoit une évolution des températures grimant jusqu'à 4,5°C.

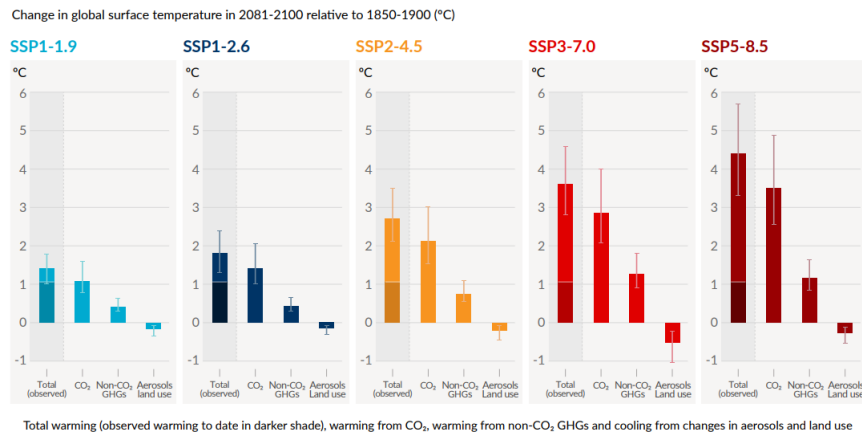


FIGURE 3.38 – Scénarios évalués par le GIEC.

Les précipitations seront également impactées, et ont déjà commencées à l'être. Les experts du GIEC prédisent une augmentation certaine du nombre d'inondations dans le monde du fait de l'augmentation des épisodes pluvieux extrêmes. Ce n'est pas tant l'augmentation des précipitations qui est le plus dangereux mais l'augmentation des épisodes de pluie à intensité extrême.

Nous avons tenté d'évaluer les impacts probables sur le risque de crue en nous appuyant sur les études réalisées par le GIEC ou le WWA. Ainsi, nous avons constitué trois scénarios que nous avons nommés de façon identique aux scénarios GIEC : SSP1-1.9, SSP2-4.5 et SSP5-8.5. Le premier représente le scénario optimiste, le deuxième le scénario moyen et enfin le dernier représente le scénario pessimiste. Nous avons proposé une règle d'évolution de nos lois de précipitations pour chacun des scénarios et nous avons évalué l'impact sur nos périodes de retour.

3.6.2 Intégration d'hypothèses GIEC

Notre modélisation de la hauteur d'eau de la Seine au travers des précipitations relevées en amont du bassin nous permet de pouvoir tester des scénarios en accord avec l'étude du GIEC. Ainsi, nous allons effectuer des chocs de fréquence et d'intensité sur

7. World Weather Attribution : organisme regroupant des centaines de climatologues de différentes institutions à travers le monde.

nos précipitations et en évaluer les conséquences. Nous avons mis en place trois scénarios afin d'évaluer l'impact du changement climatique sur le risque de crue de la Seine dans le bassin parisien. Nous avons mis en place des hypothèses sur l'ordre de grandeur des évolutions de fréquence et d'intensité mais l'intérêt de cette modélisation est qu'elle est complètement paramétrable et permet d'envisager n'importe quel scénario de choc sur les précipitations.

Scénario SSP1-1.9

Pour notre premier scénario SSP1-1.9, nous considérons que la moyenne des précipitations n'augmentera pas dans le futur, mais que les épisodes pluvieux seront plus sévères. Ainsi, nous considérons que l'espérance de notre loi des précipitations reste stable. Pour cela, nous augmentons la sévérité moyenne de +5% et nous baissions en parallèle la fréquence de l'inverse de ces +5%. On notera #1, #2 et #3 nos scénarios. On considérera alors pour toute station S et pour tout mois M , nous avons les lois modifiées :

$$N_S^M \#1 \sim \mathcal{B} \left(m^M, \frac{1}{(1+5\%)} \frac{\widehat{\lambda}_S^M}{m^M} \right)$$

$$N_S^M \#1 \sim \mathcal{P} \left(\frac{1}{(1+5\%)} \widehat{\lambda}_S^M \right)$$

$$N_S^M \#1 \sim \mathcal{NB} \left(m^M, \frac{1}{(1+5\%)} \frac{\widehat{\lambda}_S^M}{m^M} \right)$$

Pour la sévérité des précipitations, selon la loi régissant les pluies d'une station S pour un mois M , nous avons modifié les paramètres de cette loi de telle sorte à obtenir :

$$\mathbb{E}[Q_S^M \#1] = (1+5\%) \mathbb{E}[Q_S^M]$$

Scénario SSP2-4.5

Pour notre second scénario SSP2-4.5, nous considérons que les précipitations vont augmenter en fréquence et en sévérité dans le futur. Ainsi, nous considérons que l'espérance de chacune de nos lois (fréquence et sévérité) va augmenter de +5%. On considérera alors pour toute station S et pour tout mois M , nous avons les lois modifiées :

$$N_S^M \#2 \sim \mathcal{B} \left(m^M, (1+5\%) \frac{\widehat{\lambda}_S^M}{m^M} \right)$$

$$N_S^M \#2 \sim \mathcal{P} \left((1+5\%) \widehat{\lambda}_S^M \right)$$

$$N_S^M \#2 \sim \mathcal{NB} \left(m^M, (1+5\%) \frac{\widehat{\lambda}_S^M}{m^M} \right)$$

Pour la sévérité des précipitations, selon la loi régissant les pluies d'une station S pour un mois M , nous avons modifié les paramètres de cette loi de telle sorte à obtenir :

$$\mathbb{E}[Q_S^M \#2] = (1 + 5\%) \mathbb{E}[Q_S^M]$$

Scénario SSP5-8.5

Pour notre dernier scénario SSP5-8.5, nous considérons que les précipitations vont augmenter en fréquence et en sévérité comme pour SSP2-4.5 mais de façon deux fois plus importante. Ainsi, nous considérons que l'espérance de chacune de nos lois (fréquence et sévérité) va augmenter de +10%. On considérera alors pour toute station S et pour tout mois M , nous avons les lois modifiées :

$$N_S^M \#3 \sim \mathcal{B} \left(m^M, (1 + 10\%) \frac{\widehat{\lambda}_S^M}{m^M} \right)$$

$$N_S^M \#3 \sim \mathcal{P} \left((1 + 10\%) \widehat{\lambda}_S^M \right)$$

$$N_S^M \#3 \sim \mathcal{NB} \left(m^M, (1 + 10\%) \frac{\widehat{\lambda}_S^M}{m^M} \right)$$

Pour la sévérité des précipitations, selon la loi régissant les pluies d'une station S pour un mois M , nous avons modifié les paramètres de cette loi de telle sorte à obtenir :

$$\mathbb{E}[Q_S^M \#3] = (1 + 10\%) \mathbb{E}[Q_S^M]$$

3.6.3 Résultats

Nous avons réévalué les périodes de retour des différences hauteurs d'eau avec les lois choquées selon les critères définis pour chacun des scénarios. En appliquant n'importe lequel de ces scénarios, on observe un réel impact sur le nombre de crues observées. Les périodes de retour sont affectées, et notamment les périodes de retour des événements extrêmes. On observe une diminution de plus en plus considérable à mesure que le scénario est plus pessimiste. Voici l'impact estimé de chacun des scénarios sur les périodes de retours :

Hauteur d'eau (m)	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9
PdR actuelle	5	8	12	21	45	104	168	357	1250
PdR SSP1-1.9 (+1,5°C)	5	8	11	19	39	94	146	327	1030
PdR SSP2-4.5 (+2,5°C)	4	7	10	16	31	72	126	271	786
PdR SSP5-8.5 (+4,5°C)	3	6	8	11	24	41	86	127	487

TABLE 3.6 – Périodes de retour sous différents des scénarios de changement climatique.

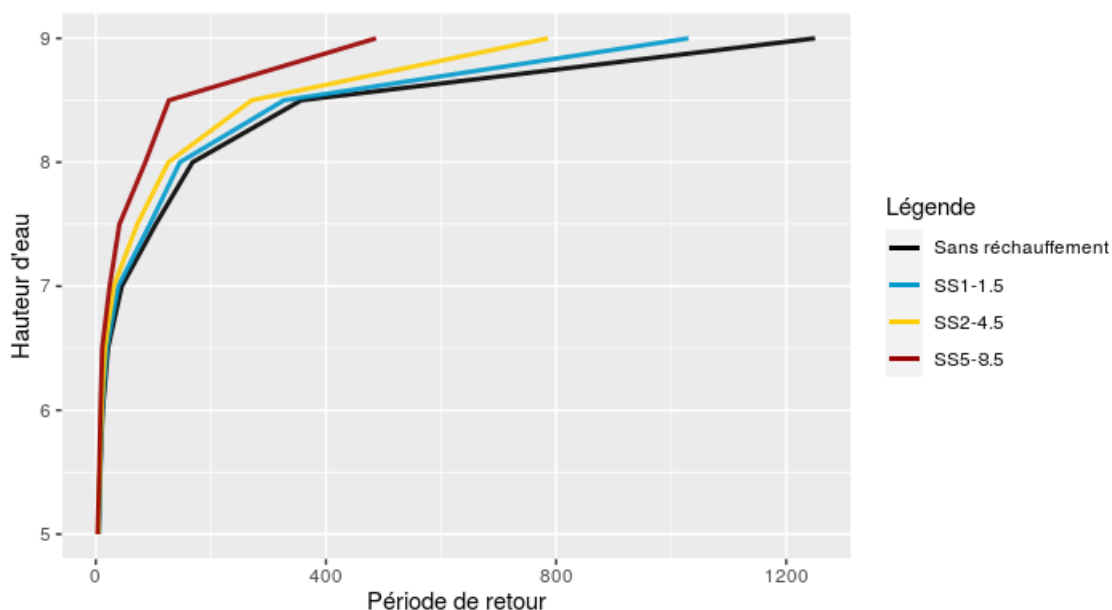


FIGURE 3.39 – Évolution des périodes de retour selon le scénario de réchauffement climatique.

Scénario SSP1-1.9

Voici le détail de l'impact du scénario SSP1-1.9 sur nos indicateurs :

Hauteur d'eau (m)	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9
Probabilité d'atteinte	0,2	0,13	0,09	0,05	0,03	0,01	7e-3	3e-3	1e-3
PdR estimée (a)	5	8	11	19	39	94	146	327	1030
Quantile pertes (M€)	31,5	35,2	39,7	47,8	63,4	74,5	84,3	101	126

TABLE 3.7 – Impacts du scénario SSP1-1.9.

On peut ainsi tracer l'évolution de la cartographie des zones sous risque par périodes de retours :

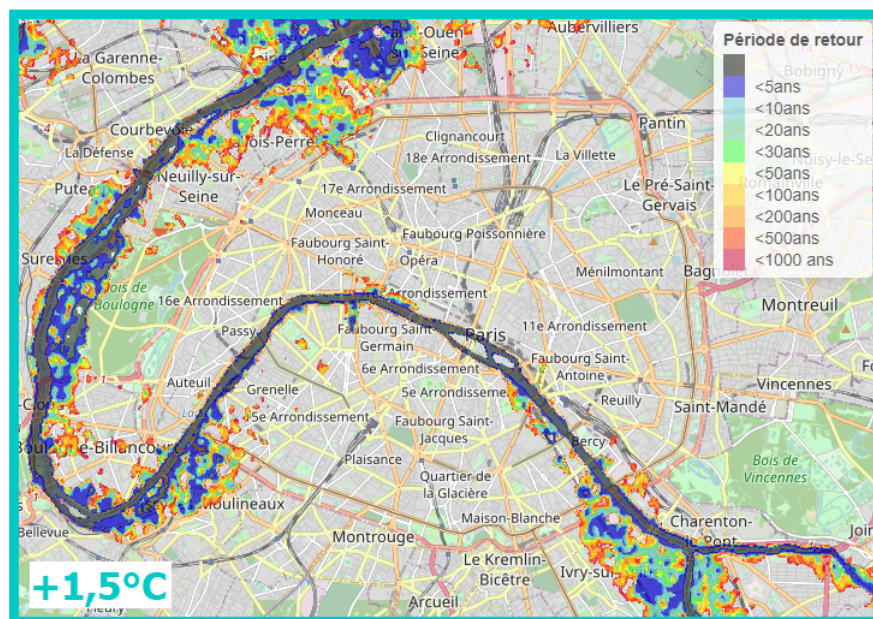


FIGURE 3.40 – Cartographie des zones sous risque pour les périodes de retour estimées sous SSP1-1.5.

Scénario SSP2-4.5

Nous faisons de même pour le deuxième scénario :

Hauteur d'eau (m)	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9
Probabilité d'atteinte	0,25	0,14	0,1	0,06	0,03	0,01	8e-3	4e-3	1e-3
PdR estimée (a)	4	7	10	16	31	72	126	271	786
Quantile pertes (M€)	31,5	35,2	39,7	47,8	63,4	74,5	84,3	101	126

TABLE 3.8 – Impacts du scénario SSP2-4.5.

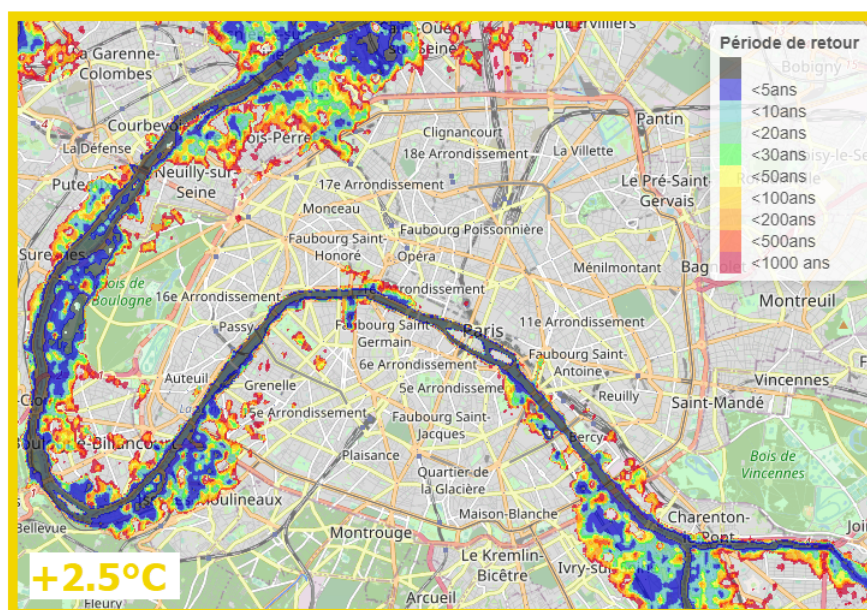


FIGURE 3.41 – Cartographie des zones sous risque pour les périodes de retour estimées sous SSP2-4.5.

Scénario SSP5-8.5

Hauteur d'eau (m)	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9
Probabilité d'atteinte	0,33	0,17	0,13	0,09	0,05	0,02	0,01	8e-3	2e-3
PdR estimée (a)	3	6	8	11	24	41	86	127	487
Quantile pertes (M€)	31,5	35,2	39,7	47,8	63,4	74,5	84,3	101	126

TABLE 3.9 – Impacts du scénario SSP5-8.5.

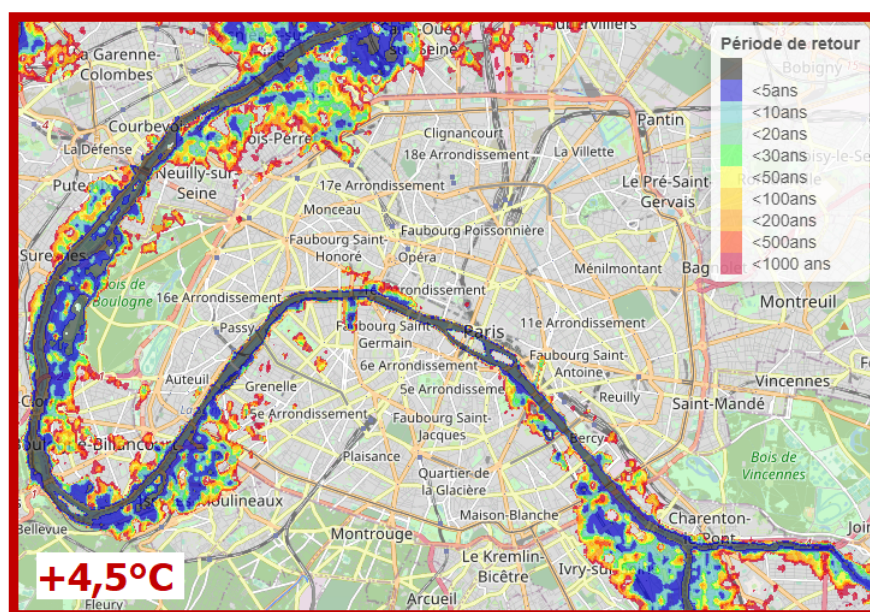


FIGURE 3.42 – Cartographie des zones sous risque pour les périodes de retour estimées sous SSP5-8.5.

Pour conclure, une augmentation de $4,5^{\circ}\text{C}$ par rapport à l'ère pré-industrielle verrait les crues de 1910 et de 2016 se reproduire deux à trois fois plus souvent, dans le cadre de ces scénarios.

Conclusion

Le travail réalisé dans ce mémoire avait pour objectif de mettre en place une modélisation générale du risque de crue de la Seine dans le bassin parisien.

Nous avons dans un premier temps réalisé une cartographie de la topographie parisienne. Ainsi, via le géocodage de nos expositions nous avons pu obtenir l'altitude de chacun des biens assurés. Ensuite, à l'aide de notre cartographie de la Seine et de la réalisation de cartes d'empreinte de crue, nous avons pu déterminer de façon précise quelles expositions de notre portefeuille étaient touchées par une crue d'intensité donnée.

L'étude d'un historique de sinistres inondation couplée à des hypothèses formulées par avis d'experts nous ont permis de calibrer des taux de destruction moyens dépendant à la fois des caractéristiques du bien assuré, de l'intensité de la crue mais aussi de la durée d'immersion du bâtiment afin de modéliser au mieux les dégâts causés par une inondation.

Enfin, la mise en place d'un modèle d'évaluation de l'évolution de la hauteur d'eau du fleuve via l'étude des précipitations du bassin versant sur les jours antérieurs nous a permis de simuler des trajectoires annuelles de hauteurs d'eau afin de calculer les périodes de retour associées à chaque intensité de crue ainsi que les pertes brutes induites sur notre portefeuille assuré.

Les applications de ce modèle sont diverses : modification des conditions contractuelles et tarifaires, optimisation de réassurance, étude des risques humains, étude de l'impact du dérèglement climatique, etc. Nous avons d'ailleurs développé quelques scénarios à la fin de ce mémoire en nous basant sur le rapport du GIEC. Ces scénarios permettent d'illustrer les multiples utilisations qu'une modélisation fine du risque de crue peut avoir comme avantage dans le futur pour les assureurs.

Il est à noter que si nous nous sommes intéressés au risque de crue de la Seine sur le bassin parisien à raison des données que nous avons à disposition, le modèle est cependant répliquable à n'importe quel bassin fluvial du globe.

Nous pouvons néanmoins lister quelques limites à notre étude. La modélisation de la pluviométrie a été réalisée sous différentes simplifications, du fait de la complexité de

leur loi évolutive réelle. En effet, la météorologie et notamment la pluviométrie est une science régie par des lois bien plus complexes que celles que nous avons mises en place. Les phénomènes météorologiques sont dynamiques à la fois spatialement et temporellement. C'est pourquoi une modélisation plus robuste viendrait améliorer considérablement ces travaux, mais cela sortait totalement du cadre de ce mémoire.

De plus, des travaux ultérieurs pourraient ainsi s'intéresser à la modélisation de gel des sols, de gestion des lacs réserve ou de taux de remplissage des nappes phréatiques afin d'améliorer le pouvoir explicatif du modèle.

Nous pouvons aussi faire état d'une autre piste d'amélioration par rapport aux taux de destruction. Nous avons essayé d'améliorer nos taux de destruction au moyen de quelques résultats provenant d'études scientifiques, mais la collecte d'un grand catalogue de données et une meilleure identification de la sinistralité chez les assureurs pourrait permettre d'affiner cette partie de la modélisation.

Malgré les limites relevées, nous avons mis en place un modèle opérationnel et permettant de nous donner une vision scientifique sur des indicateurs encore considérés relativement incertains comme les périodes de retour des crues historiques. Notre modèle permet également pour n'importe quel assureur de calculer un SCR adapté à son portefeuille en modèle interne, et cela ne représente qu'une des nombreuses applications possibles de cette étude.

Table des figures

1	Cartographie de nos expositions sur le bassin parisien.	vi
2	Carte topographique du bassin parisien (base de données <i>Copernicus</i>). . .	vii
3	Reproduction de la carte d’empreinte de la crue de 1910.	vii
4	Pourcentage d’exposition dans l’empreinte de crue selon la hauteur d’eau.	viii
5	Taux de destruction selon la hauteur d’immersion et l’année de construc- tion du bien assuré (pour une maison et une durée d’immersion de 3 jours).	ix
6	Taux de destruction retenus adaptés aux caractéristiques de la maison. . .	ix
7	Taux de destruction retenus adaptés aux caractéristiques de l’appartement (1er étage).	ix
8	Évaluation des pertes brutes selon la dynamique de crue (hauteur/durée).	x
9	Cartographie des stations <i>ISD</i> étudiées.	xi
10	Exemple de simulations annuelles de la hauteur d’eau à Austerlitz.	xii
11	Cartographie des zones sous risque pour les périodes de retour estimées. .	xii
12	Évolution des périodes de retour selon le scénario considéré.	xiii
13	Évolution des périodes de retour selon le scénario considéré.	xiv
14	Mapping of our exposures in the Paris basin.	xvi
15	Topographic map of the Paris basin (<i>Copernicus</i> data).	xvi
16	Footprint map of the 1910 flood.	xvii
17	Percentage of exposure in the flood footprint by water level.	xvii
18	Rate of destruction according to the height of immersion and the year of construction of the insured property (for a house and an immersion period of 3 days).	xviii
19	Selected destruction rates adapted to the characteristics of the house. . . .	xix
20	Selected destruction rates adapted to the characteristics of the flat (1st floor).	xix
21	Assessment of gross losses according to flood dynamics (height/duration).	xix
22	Mapping of considered <i>ISD</i> stations.	xx
23	Examples of annual simulations of the water level at Austerlitz.	xxi
24	Mapping of areas at risk with estimated return periods.	xxii
25	Evolution of return periods according to the scenario considered.	xxii
26	Evolution of return periods according to the scenario considered.	xxiii
1.1	Répartition des catastrophes naturelles par typologie (1990 à 2007). <i>Source</i> : <i>EM-DAT - The Emergency Events Database</i>	4

1.2	Evolution du nombre de catastrophes naturelles depuis 1950. <i>Source : EM-DAT - The Emergency Events Database</i>	4
1.3	Evolution du coût des catastrophes naturelles depuis 1950. <i>Source : EM-DAT - The Emergency Events Database</i>	4
1.4	Indice de risque CAT NAT par pays. <i>Source : catnat.net</i>	5
1.5	Evolution du coût des CAT NAT en France depuis 1982. <i>Source : CCR</i>	5
1.6	Répartition des CAT NAT en France. <i>Source : CCR</i>	6
1.7	Répartition des CAT NAT en France (AUTRES). <i>Source : CCR</i>	6
1.8	Exposition de la population française aux aléas climatiques (2016). <i>Sources : MTES, 2017 ; Insee, 2014 ; IGN, 2016. Traitements : SDES, 2019.</i>	6
1.9	Dispositif d'indemnisation des assurés. <i>Source : Ministère de l'intérieur</i>	7
1.10	Inondation par ruissellement. <i>Source : Hydro Eau France</i>	9
1.11	Inondation par submersion marine. <i>Source : Hydro Eau France</i>	9
1.12	Inondation par débordement de cours d'eau. <i>Source : Hydro Eau France</i>	10
1.13	Inondation par remontée de nappes. <i>Source : Hydro Eau France</i>	10
1.14	Carte des zones inondables. <i>Source : Ministère de l'écologie</i>	12
1.15	Bassin de la Seine. <i>Source : GIP Seine-Aval.</i>	14
1.16	Principaux affluents de la Seine.	14
1.17	Débits moyens des fleuves français.	15
1.18	Débits moyens annuels à Austerlitz. <i>Source : DRIEE</i>	15
1.19	Pluviométrie en France.	16
1.20	Pluviométries mensuelles sur le bassin de la Seine. <i>Source : Climate data</i>	17
1.21	Historique des crues de la Seine (1649-2016). <i>Source : Oliver BERRUYER.</i>	18
1.22	Photos d'époque lors de la crue de 1910.	19
1.23	Précipitations précédent la crue de 1910. <i>Source : Météo France.</i>	19
1.24	Hydrogramme de la crue de 1910. <i>Source : Episeine.</i>	20
1.25	Pluviométrie moyenne historique en mai à Paris. <i>Source : ECA.</i>	20
1.26	Crue de le Seine de mai 2016. <i>Source : Kilohn limahn.</i>	21
1.27	Carte des quatre lacs réservoirs de la Seine. <i>Source : Nord Nord West.</i>	22
1.28	Ancienne cartographie de la Seine à Paris.	23
1.29	Zouave du pont de l'alma et seuils d'alerte. <i>Source : DRIEE</i>	23
1.30	Zouave du pont de l'alma se noyant lors de la crue de 1910.	23
1.31	Infrastructures historiques en bords de Seine.	24
1.32	Impacts de la crue de 2016 à Paris. <i>Source : AFP.</i>	25
1.33	Démarche d'un modèle CAT NAT.	26
1.34	Démarche du module exposition.	27
1.35	Pluies simultanées et crues concomitantes.	28
1.36	Évaluation des périodes de retour par simulations.	29
1.37	Démarche du module aléa.	30
1.38	Exemple de courbe de vulnérabilité.	31
1.39	Démarche du module vulnérabilité.	32
2.1	Zone d'exposition considérée pour notre étude.	35

2.2	Carte des plus hautes eaux connues de le Seine sur le bassin parisien. <i>Source : DRIEE.</i>	36
2.3	Carte topographique de Paris. <i>Source : ThePromenader.</i>	37
2.4	Satellites d'observation de la Terre. <i>Source : ESA (European Space Agency).</i>	38
2.5	Vue d'ensemble des données Copernicus. <i>Source : Copernicus.</i>	38
2.6	Topographie : Zoom sur le bassin parisien. <i>Données : Copernicus.</i>	39
2.7	Topographie : Zoom sur Paris. <i>Données : Copernicus.</i>	40
2.8	Topographie : Zoom sur le pont d'Austerlitz. <i>Données : Copernicus.</i>	40
2.9	Cartographie de la Seine selon les seuils. <i>Données : Copernicus.</i>	42
2.10	Cartographie seuillée par zones.	43
2.11	Vue en éclaté des zones de différenciation.	43
2.12	Cartographie de la Seine (Modélisation).	44
2.13	Cartographie de la Seine (Google Maps).	44
2.14	Empreinte de la crue de 1910 (Modélisation).	44
2.15	Empreinte de la crue de 1910 (Geoportail).	44
2.16	Obtention des altitudes de nos expositions.	45
2.17	Parts d'appartements et de maisons dans la base.	46
2.18	Parts d'exposition selon l'étage de l'appartement.	46
2.19	Parts d'exposition selon la présence de dépendances dans l'habitation.	47
2.20	Parts d'exposition selon l'année de construction de l'habitation.	47
2.21	Parts d'exposition selon la superficie de l'habitation.	48
2.22	Parts d'exposition selon le nombre de pièces de l'habitation.	48
2.23	Parts d'exposition par département.	49
2.24	Parts d'exposition selon le capital assuré.	49
2.25	Expositions sur le bassin parisien.	50
2.26	Agrégation des expositions : Méthode du plus proche voisin (Rectangles).	51
2.27	Expositions sur le bassin parisien (Maille "Copernic").	51
2.28	Expositions sur le bassin parisien (Maille "Copernic") - Zoom.	52
2.29	Agrégation des expositions : Méthode du plus proche voisin (Hexagones).	52
2.30	Expositions sur le bassin parisien (Maille hexagonale).	53
2.31	Pourcentage d'exposition dans l'empreinte de crue selon la hauteur d'eau.	54
2.32	Cartographie d'une crue de 7 mètres et part d'expositions touchées.	55
2.33	Cartographie d'une crue de 9 mètres et expositions touchées	55
2.34	Cartographie des expositions dans l'empreinte de crue de 2016 (Maille à l'adresse).	56
2.35	Cartographie des expositions dans l'empreinte de crue de 2016 (Maille hexagonale).	57
2.36	Cartographie des expositions dans l'empreinte de crue de 1910 (Maille à l'adresse).	57
2.37	Cartographie des expositions dans l'empreinte de crue de 1910 (Maille hexagonale).	58
2.38	Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (Appartement ou Maison).	59

2.39	Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (selon étage).	59
2.40	Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (selon année de construction).	60
2.41	Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (selon département).	61
2.42	Sommes assurées et coût total des sinistres par type d'habitation.	62
2.43	Sommes assurées et coût total des sinistres par étage.	63
2.44	Sommes assurées et coût total des sinistres par année de construction.	63
2.45	Sommes assurées et coût total des sinistres par superficie.	64
2.46	Courbe de vulnérabilité pour une maison quelconque.	66
2.47	Courbe univariée de vulnérabilité pour les maisons selon l'année de construction (durée d'immersion fixée à " $\leq 2j$ ").	67
2.48	Courbe univariée de vulnérabilité pour les maisons selon la durée d'immersion (durée d'immersion fixée à " $\leq 2j$ ").	67
2.49	Taux de destruction selon la hauteur d'immersion et l'année de construction du bien assuré (pour une maison et une durée d'immersion de trois jours.	68
2.50	Taux de destruction retenus adaptés aux caractéristiques de la maison.	68
2.51	Courbe de vulnérabilité pour un appartement au rez-de-chaussée.	69
2.52	Courbe de vulnérabilité pour un appartement selon l'étage.	69
2.53	Taux de destruction retenus adaptés aux caractéristiques de l'appartement et à la crue (Appartement RDC).	70
2.54	Taux de destruction retenus adaptés aux caractéristiques de l'appartement et à la crue (Appartement 1er étage).	70
2.55	Taux de destruction retenus adaptés aux caractéristiques de l'appartement et à la crue (Appartement 2ème étage).	70
2.56	Étude de la perte brute pour une exposition du portefeuille et une crue donnée.	71
2.57	Étude de la perte brute pour une exposition du portefeuille et une crue donnée.	72
2.58	Étude de la perte brute selon la dynamique de crue (hauteur et durée d'immersion).	73
3.1	Cartographie des stations étudiées.	77
3.2	Réseau constitué des stations <i>HYDRO</i> étudiées.	78
3.3	Historique des hauteurs d'eau relevées à Austerlitz depuis 1885.	79
3.4	Différents aperçus des données hydrométriques.	80
3.5	Relevés hydrométriques à Austerlitz et à Arcis entre janvier et avril 2020.	81
3.6	Interpolation des relevés d'eau du 1er janvier 2020 à Austerlitz.	82
3.7	Méthode des moindres carrés.	83
3.8	Dynamique des ondes de crue du bassin versant de la Seine.	85
3.9	Stations ISD sur le bassin de la Seine.	86

3.10	Cartographie des stations ISD étudiées.	87
3.11	Historique des pluies relevées à PARIS MONTSOURIS.	88
3.12	Précipitations à PARIS MONTSOURIS lors de la crue de 2016.	88
3.13	Historique des pluies relevées à PARIS MONTSOURIS entre 1983 et 2021 après interpolation bi-hebdomadaire	90
3.14	Moyennes historiques des pluies relevées par demi-journée à PARIS MONT- SOURIS par mois.	90
3.15	Écart-types historiques des pluies relevées par demi-journée à PARIS MONTSOURIS par mois.	91
3.16	Histogramme et densité empirique de la quantité de pluie tombée lors des demi-journées pluvieuses de janvier à PARIS MONTSOURIS.	92
3.17	Nombre de demi-journées de pluie durant le mois de janvier par année à Paris-Montsouris.	94
3.18	Adéquation d'une loi logistique sur les précipitations de janvier à PARIS MONTSOURIS.	94
3.19	Adéquation d'une loi log-normale sur les précipitations de janvier à PARIS MONTSOURIS.	95
3.20	Tableau des log-vraisemblances des adéquations de lois pour PARIS MONT- SOURIS.	95
3.21	Test de Kolmogorov-Smirnov pour les précipitations en janvier à PARIS MONTSOURIS.	96
3.22	QQ-plot de la loi log-normale pour les précipitations en janvier à PARIS MONTSOURIS.	96
3.23	Corrélations de Pearson entre les stations ISD sur le mois de Janvier (%).	98
3.24	Différence de hauteur d'eau entre deux relevés consécutifs à Austerlitz.	100
3.25	Sortie \mathbf{R} du premier LM.	102
3.26	Sortie \mathbf{R} du LM final.	103
3.27	Dépendance linéaire entre les stations du modèle LM avec lags.	104
3.28	Simulation annuelle des précipitations à BRETTIGNY.	105
3.29	Simulation annuelle simultanée des précipitations à BRETTIGNY et VIL- LAROCHE.	105
3.30	Exemple de simulation annuelle de la hauteur d'eau à Austerlitz.	106
3.31	6 simulations annuelles de la hauteur d'eau à Austerlitz.	107
3.32	Distribution des pertes annuelles pour notre portefeuille assuré (10 000 simulations).	108
3.33	Distribution des pertes annuelles pour notre portefeuille assuré pour les simulations ayant entraîné au moins une crue.	109
3.34	Fonction de répartition empirique des pertes annuelles pour notre porte- feuille assuré.	109
3.35	Pertes brutes annuelles et périodes de retour.	110
3.36	Hauteur d'eau et périodes de retour.	111
3.37	Cartographie des zones sous risque pour les périodes de retour estimées.	112
3.38	Scénarios évalués par le GIEC.	113

3.39	Évolution des périodes de retour selon le scénario de réchauffement climatique.	116
3.40	Cartographie des zones sous risque pour les périodes de retour estimées sous SSP1-1.5.	117
3.41	Cartographie des zones sous risque pour les périodes de retour estimées sous SSP2-4.5.	118
3.42	Cartographie des zones sous risque pour les périodes de retour estimées sous SSP5-8.5.	119
A.1	Sommes assurées et coût total des sinistres par pièces.	131
A.2	Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (Selon nombre de pièces).	132
B.1	Résumé des étapes du module aléa.	133
C.1	Disparité plages de temps et relevés.	135
F.1	Adéquations de loi pour PARIS MONTSOURIS.	141
F.2	Adéquations de loi pour BRETTIGNY.	142
F.3	Adéquations de loi pour VILLAROCHE.	143
F.4	Adéquations de loi pour BRIENNE.	144
F.5	Adéquations de loi pour AUX.	145
F.6	Adéquations de loi pour BARBEREY.	146
F.7	Adéquations de loi pour ROBINSON.	147
F.8	Adéquations de loi pour CHATILLON.	148
F.9	Adéquations de loi pour LONGVIC.	149
H.1	Copule d'indépendance. <i>Source : Matteo Zandi.</i>	158
H.2	Copule gaussienne. <i>Source : Matteo Zandi.</i>	159
I.1	Calcul des corrélations de Pearson entre les stations (mois de février). . .	161
I.2	Calcul des corrélations de Pearson entre les stations (mois de mars). . . .	162
I.3	Calcul des corrélations de Pearson entre les stations (mois de avril). . . .	162
I.4	Calcul des corrélations de Pearson entre les stations (mois de mai). . . .	163
I.5	Calcul des corrélations de Pearson entre les stations (mois de juin). . . .	163
I.6	Calcul des corrélations de Pearson entre les stations (mois de juillet). . . .	164
I.7	Calcul des corrélations de Pearson entre les stations (mois de août). . . .	164
I.8	Calcul des corrélations de Pearson entre les stations (mois de septembre). .	165
I.9	Calcul des corrélations de Pearson entre les stations (mois de octobre). . .	165
I.10	Calcul des corrélations de Pearson entre les stations (mois de novembre). .	166
I.11	Calcul des corrélations de Pearson entre les stations (mois de décembre). .	166
I.12	Observation de la structure de dépendance de la pluviométrie entre les stations.	167

Liste des tableaux

1	Périodes de retour et pertes selon la hauteur d'eau.	xii
2	Return periods and losses by water level.	xxi
1.1	Répertoire des catastrophes naturelles	3
1.2	Articulation du régime d'indemnisation CAT NAT en France.	8
1.3	Répertoire des variables de la BDD MRH.	27
3.1	Caractéristiques des stations <i>HYDRO</i> considérées.	79
3.2	Vitesse de propagation d'une onde de crue vers Austerlitz.	84
3.3	Lois optimales pour PARIS MONTsouris.	97
3.4	Périodes de retour et pertes selon la hauteur d'eau.	110
3.5	Hauteur d'eau et pertes selon les périodes de retour.	111
3.6	Périodes de retour sous différents des scénarios de changement climatique.	115
3.7	Impacts du scénario SSP1-1.9.	116
3.8	Impacts du scénario SSP2-4.5.	117
3.9	Impacts du scénario SSP5-8.5.	118
F.1	Tableau récapitulatif des adéquations de loi de la pluviométrie (1/3).	150
F.2	Tableau récapitulatif des adéquations de loi de la pluviométrie (2/3).	150
F.3	Tableau récapitulatif des adéquations de loi de la pluviométrie (3/3).	151

Annexe A

Analyses sur le nombre de pièces du bien assuré

Sommes assurées et coût total des sinistres par pièces

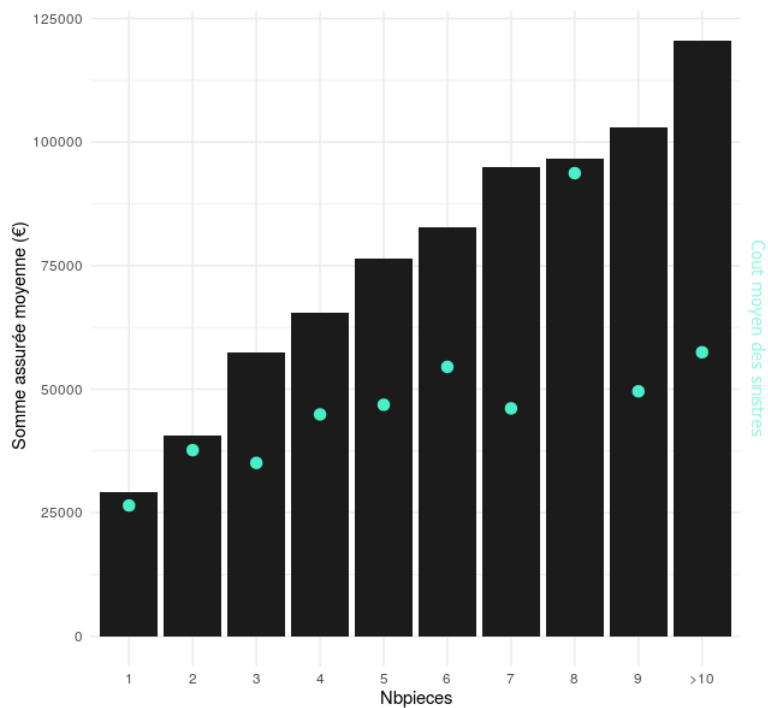


FIGURE A.1 – Sommes assurées et coût total des sinistres par pièces.

Part d'expositions touchées selon hauteur d'eau et le nombre de pièces

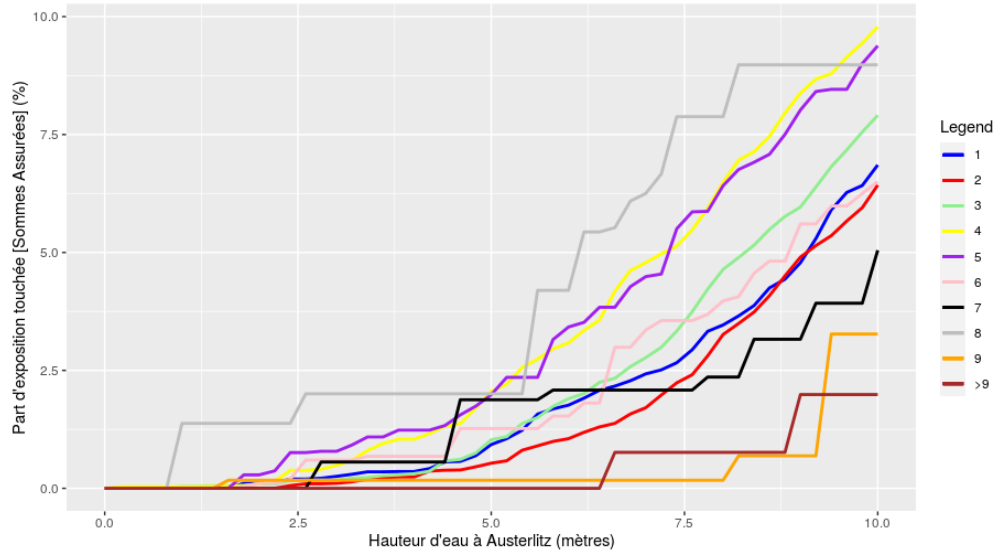


FIGURE A.2 – Analyse bivariée : Part d'expositions touchées selon hauteur d'eau (Selon nombre de pièces).

Annexe B

Résumé schématique des étapes du module aléa

Modélisation



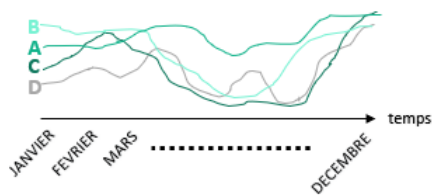
eaufrance



- Etude de la dynamique du fleuve afin de connaître les lags temporels spatiaux dans le bassin versant
- Etude des précipitations du bassin et calibrage de lois et de structures de dépendance sur ces dernières
- Mise en place d'un modèle explicatif de la hauteur d'eau à Austerlitz selon les précipitations apparentes dans le bassin à J-1, J-2, ...

Simulation

Simulation des précipitations journalières pour chaque station sur une année complète



Détermination de la hauteur d'eau journalière à Austerlitz consécutive aux précipitations

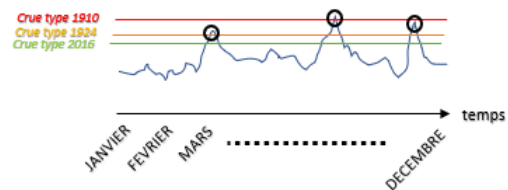


FIGURE B.1 – Résumé des étapes du module aléa.

Annexe C

Interpolation et agrégation des données ISD.

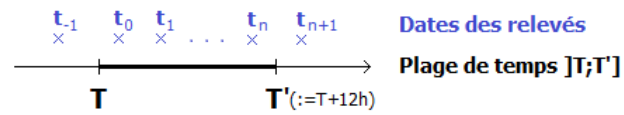


FIGURE C.1 – Disparité plages de temps et relevés.

Annexe D

Lois testées pour la simulation des précipitations

Voici les fonctions de densité des lois que nous avons testées :

Loi de Weibull : Soient $k > 0$ paramètre de forme et $\lambda > 0$ paramètre d'échelle.

$$f(x; k, \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} \quad (\text{D.1})$$

Loi Log-normale : Soient μ réel paramètre d'espérance et $\lambda > 0$ paramètre d'écart-type du logarithme de la variable.

$$f_Y(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \quad (\text{D.2})$$

Loi de Pareto : Soient $x_m > 0$ paramètre de location et $k > 0$ paramètre de forme.

$$f(x; k, x_m) = k \frac{x_m^k}{x^{k+1}} \text{ pour } x \geq x_m \quad (\text{D.3})$$

Loi Gamma : Soient $k > 0$ paramètre de forme et $\theta > 0$ paramètre d'échelle.

$$f(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\Gamma(k)\theta^k} \quad (\text{D.4})$$

Loi de Cauchy : Soient x_0 paramètre de position et $a > 0$ paramètre d'échelle.

$$f(x; x_0, a) = \frac{1}{\pi a \left[1 + \left(\frac{x-x_0}{a}\right)^2\right]} \quad (\text{D.5})$$

Loi de Gumbel : Soient μ paramètre de position et $\beta > 0$ paramètre d'échelle.

$$F'(x; \mu, \beta) = \exp\left(-e^{-\frac{x-\mu}{\beta}}\right) \quad (\text{D.6})$$

Loi Exponentielle : Soit $\lambda > 0$ paramètre d'intensité.

$$f(x) = \lambda e^{-\lambda x} \quad (\text{D.7})$$

Loi Logistique : Soient μ réel et $s > 0$.

$$f(x; \mu, s) = \frac{e^{-\frac{x-\mu}{s}}}{s \left(1 + e^{-\frac{x-\mu}{s}}\right)^2} \quad (\text{D.8})$$

Loi Log-logistique : Soient $\alpha > 0$ paramètre d'échelle et $\beta > 0$ paramètre de forme.

$$f(x; \alpha, \beta) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{[1 + (x/\alpha)^\beta]^2} \quad (\text{D.9})$$

Loi de Burr : Soient $c > 0$ et $k > 0$.

$$f(x; c, k) = ck \frac{x^{c-1}}{(1 + x^c)^{k+1}} \quad (\text{D.10})$$

Annexe E

Méthode du maximum de vraisemblance pour la loi Gamma

Méthode du maximum de vraisemblance pour la loi Gamma :

Fonction de vraisemblance : Soit X une variable aléatoire suivant une loi continue et décrite par la densité de probabilité f dépendant d'un paramètre θ (θ peut être un vecteur de paramètres). La vraisemblance est une fonction de θ , étant donné une réalisation x de la variable aléatoire X , qui s'écrit :

$$L(\theta|x) = f_{\theta}(x) = L(\theta|x) = f_{\theta}(x)$$

Maximum de vraisemblance pour la loi Gamma : La loi Gamma est une distribution à deux paramètres que l'on note α et β . Soit une observation x , la vraisemblance s'écrit alors :

$$L(\alpha, \beta|x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x).$$

Cherchons la valeur de β maximisant cette vraisemblance pour l'observation x . La formule étant complexe, on peut simplifier le calcul en passant au logarithme de la fonction de vraisemblance, ce qui ne changera pas notre résultat du fait de la stricte croissance de la fonction logarithme :

$$\log L(\alpha, \beta|x) = \alpha \log \beta - \log \Gamma(\alpha) + (\alpha - 1) \log x - \beta x$$

On calcule ensuite sa dérivée partielle par rapport à β , et on cherche à l'annuler pour trouver la valeur de β qui maximise la log-vraisemblance :

$$\frac{\partial \log L(\alpha, \beta|x)}{\partial \beta} = \frac{\alpha}{\beta} - x$$

La log-vraisemblance de l'ensemble des observations est la somme des log-vraisemblances de chaque observation individuelle (car les observations sont considérées indépendantes).

La dérivée partielle de l'ensemble des observations est donc la somme des dérivées partielles des observations individuelles :

$$\frac{\partial \log L(\alpha, \beta | x_1, x_2, \dots, x_n)}{\partial \beta} = \sum_{i=1}^n \frac{\alpha}{\beta} - x_i = n \frac{\alpha}{\beta} - \sum_{i=1}^n x_i$$

On se retrouve donc avec l'équation :

$$n \frac{\alpha}{\beta} - \sum_{i=1}^n x_i = 0$$

qui a pour solution :

$$\hat{\beta} = \alpha \frac{n}{\sum_{i=1}^n x_i}$$

$\hat{\beta}$ est donc l'estimateur du maximum de vraisemblance.

Annexe F

Résultats d'adéquation des données ISD

Adéquations optimales des lois par station et par mois

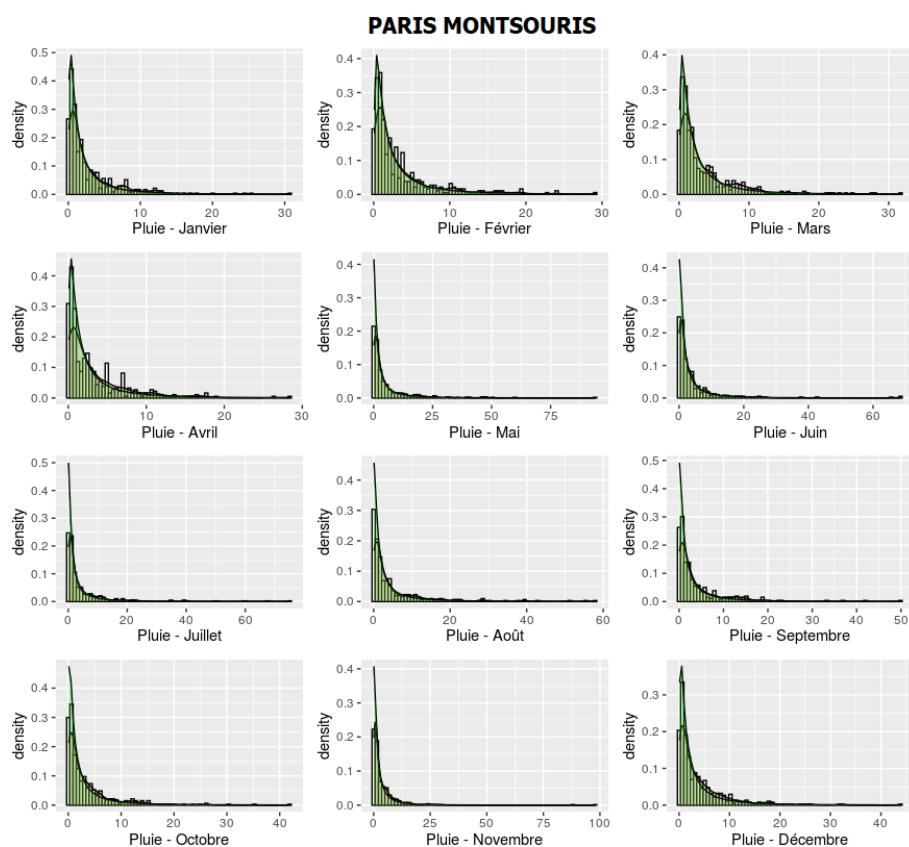


FIGURE F.1 – Adéquations de loi pour PARIS MONTSOURIS.

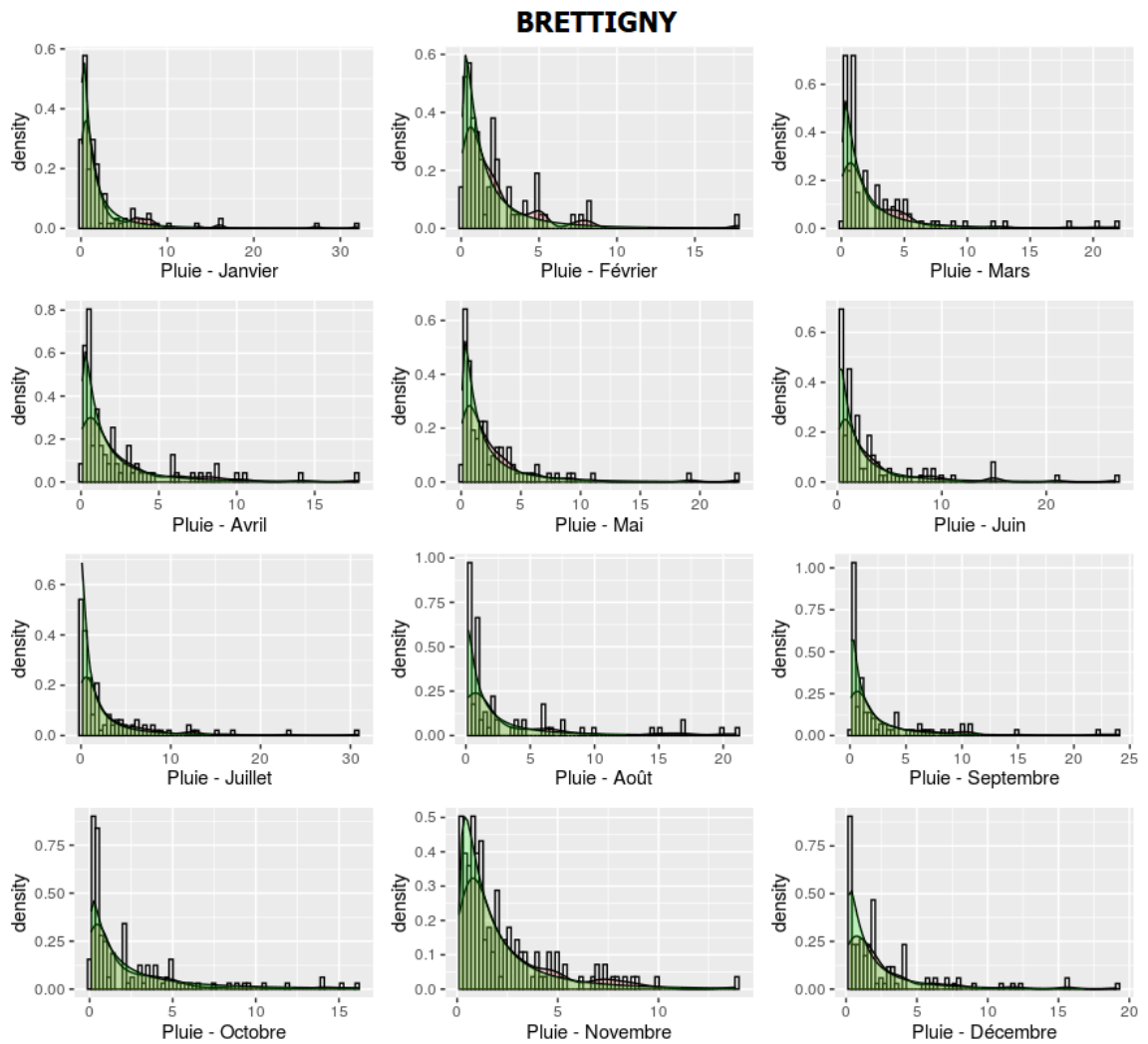


FIGURE F.2 – Adéquations de loi pour BRETTIGNY.

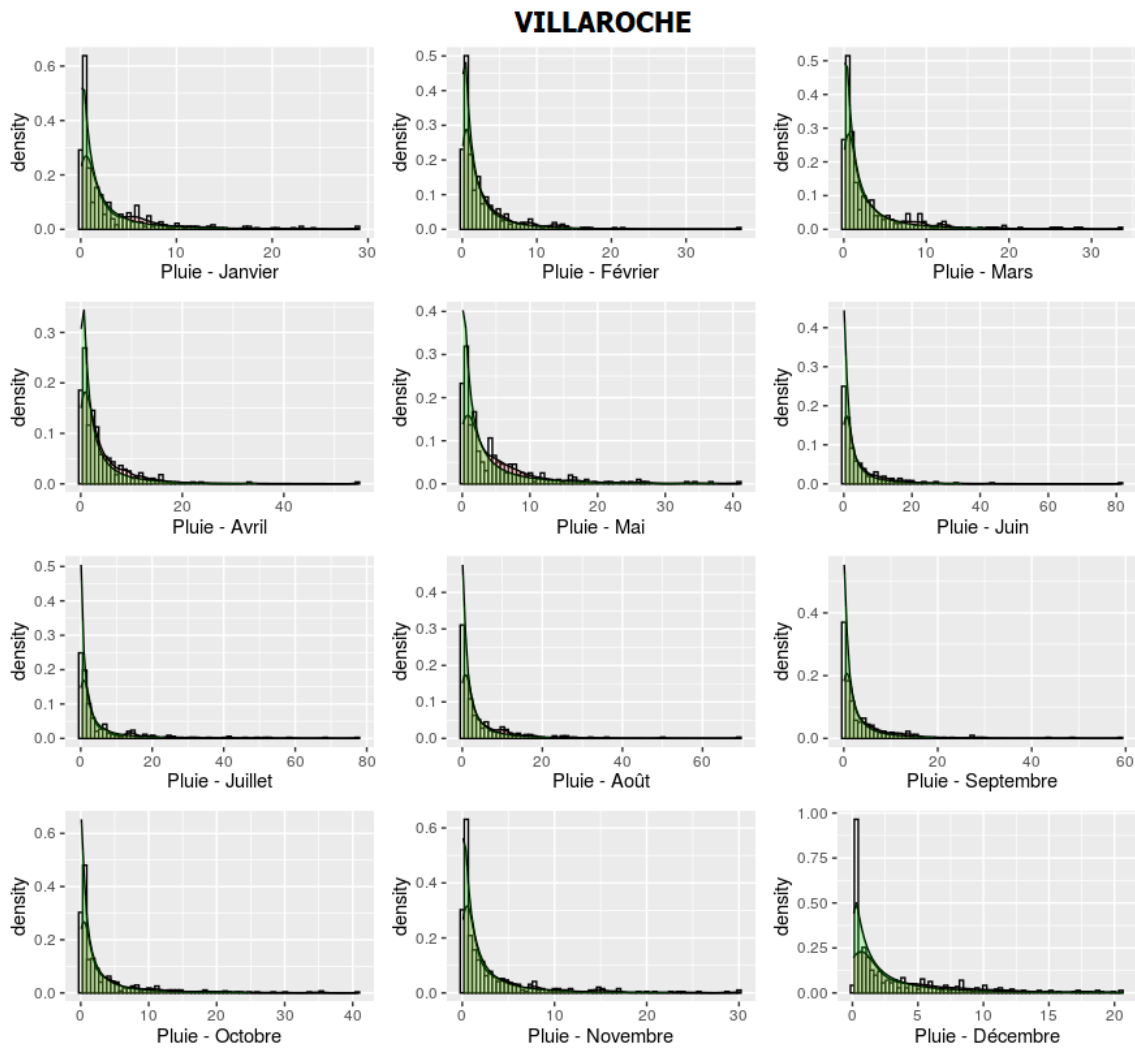


FIGURE F.3 – Adéquations de loi pour VILLAROCHE.

Tableaux récapitulatifs des lois et des paramètres retenus par station et par mois

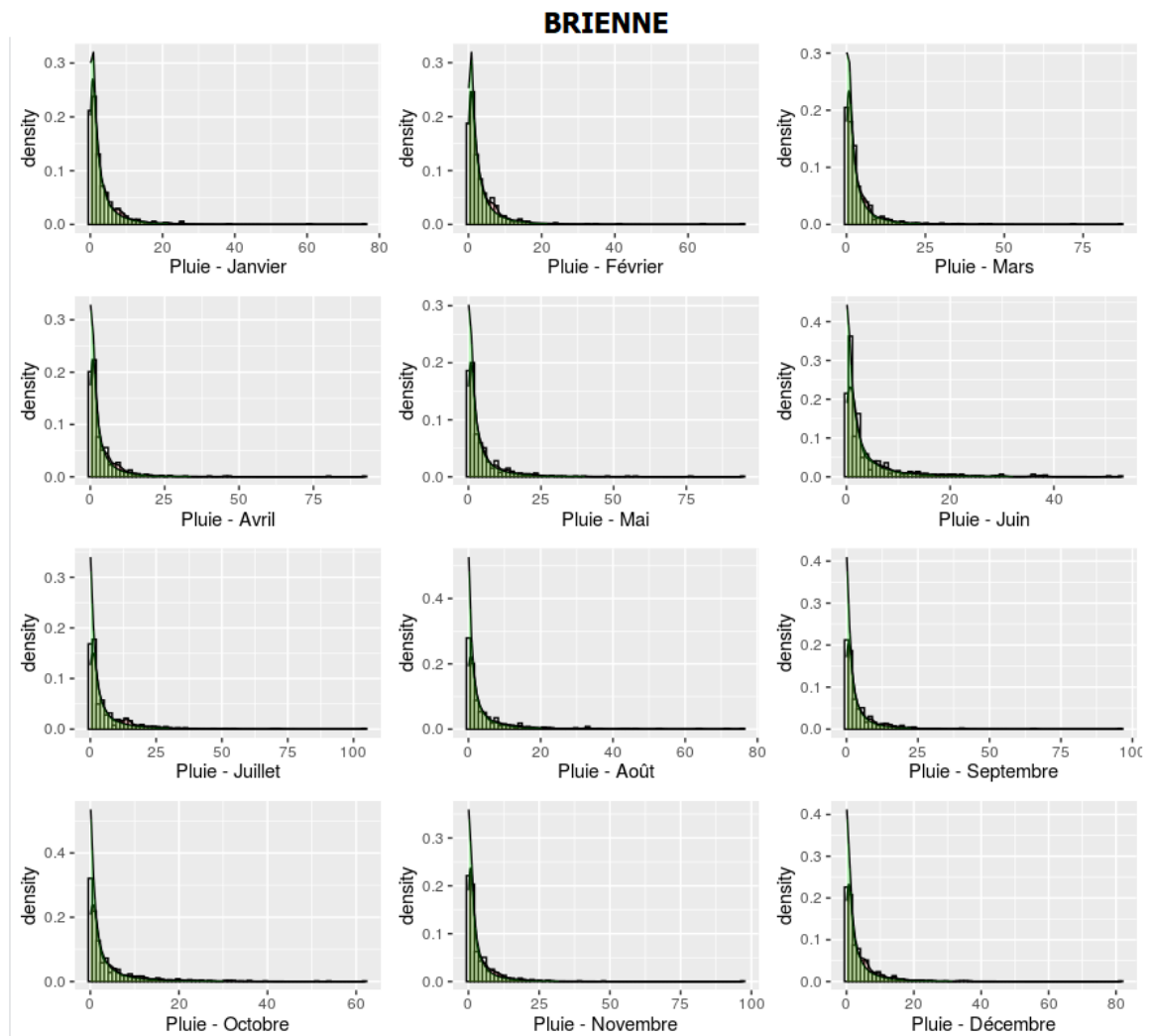


FIGURE F.4 – Adéquations de loi pour BRIENNE.

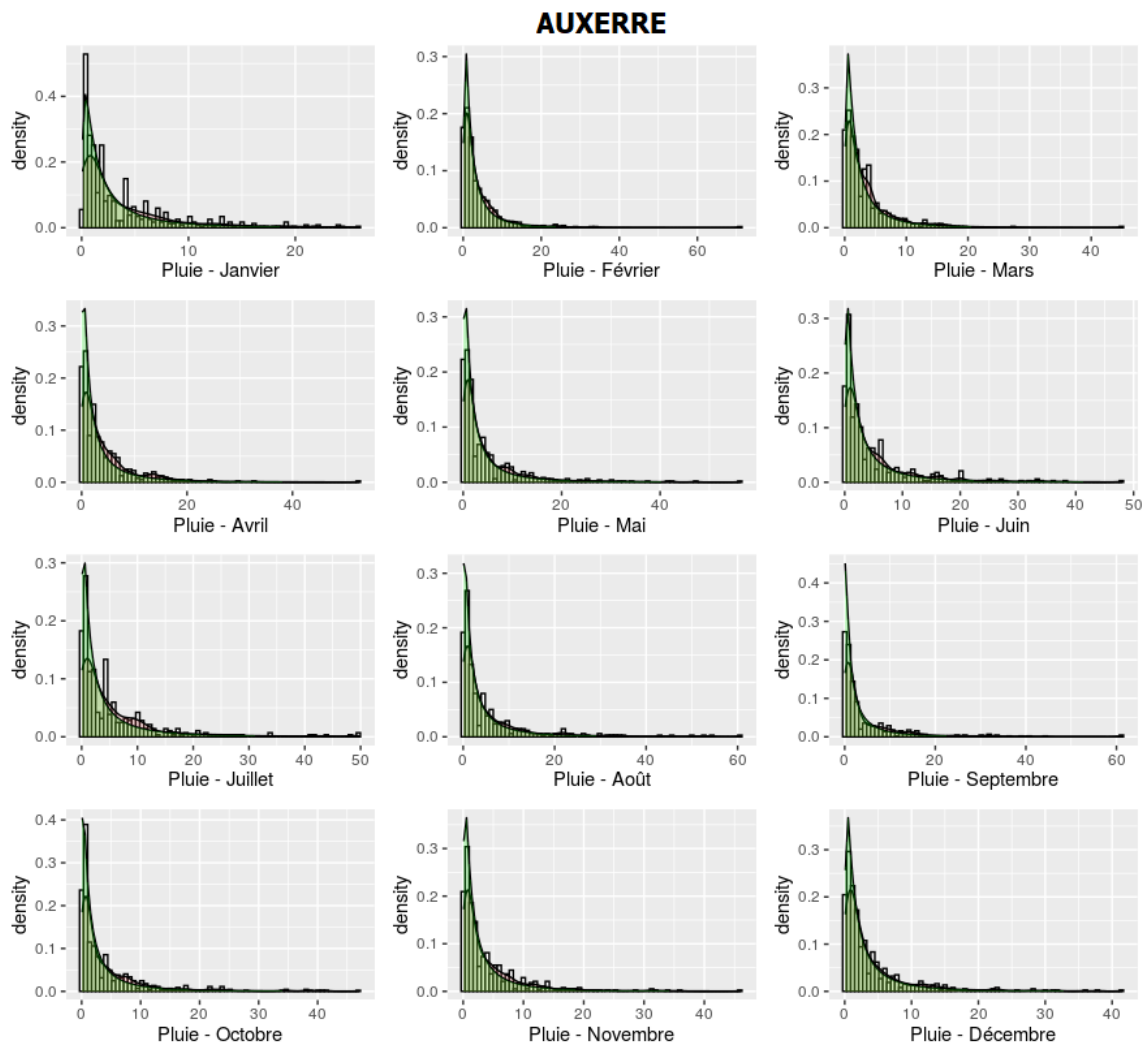


FIGURE F.5 – Adéquations de loi pour AUX.

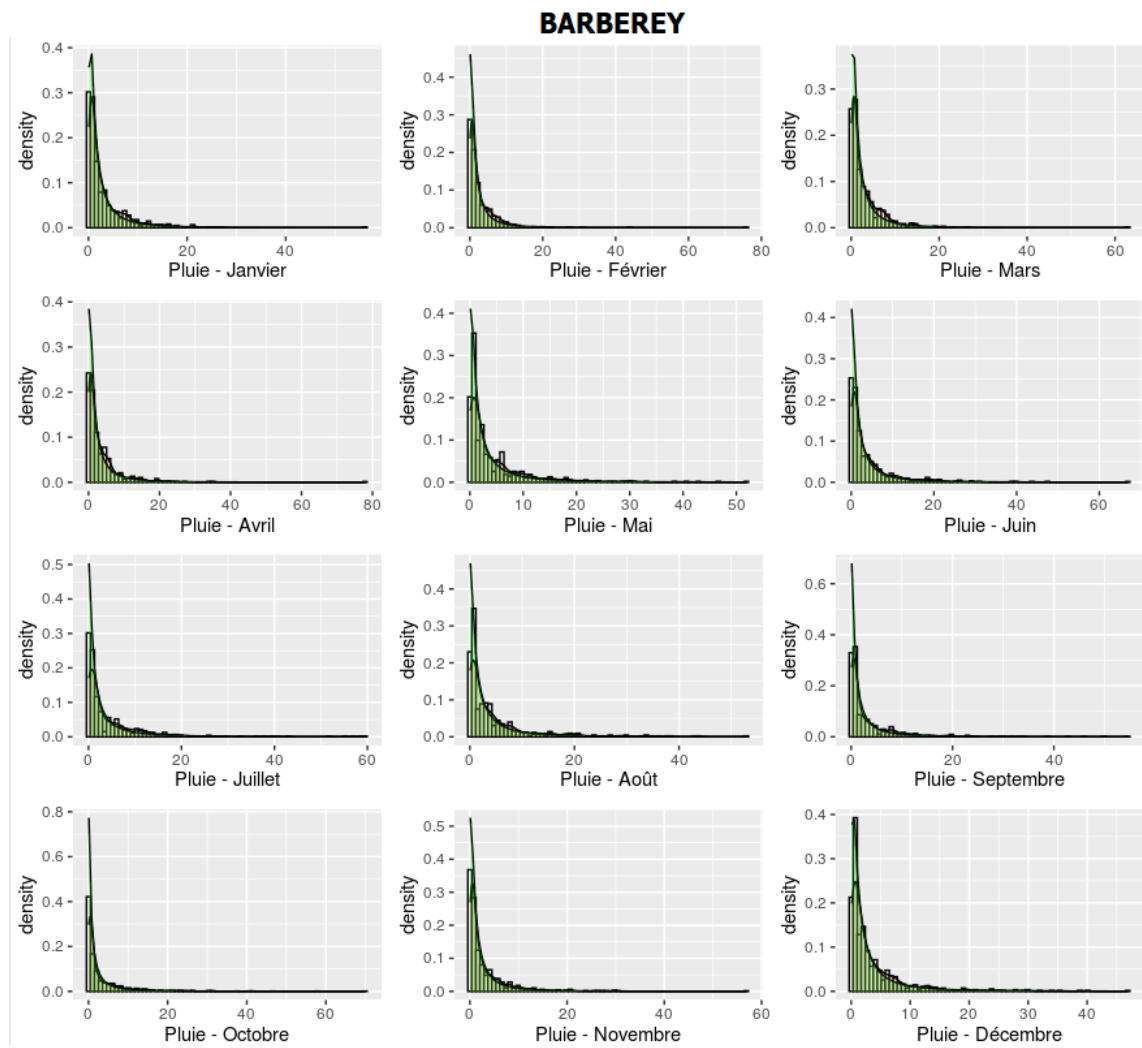


FIGURE F.6 – Adéquations de loi pour BARBEREY.

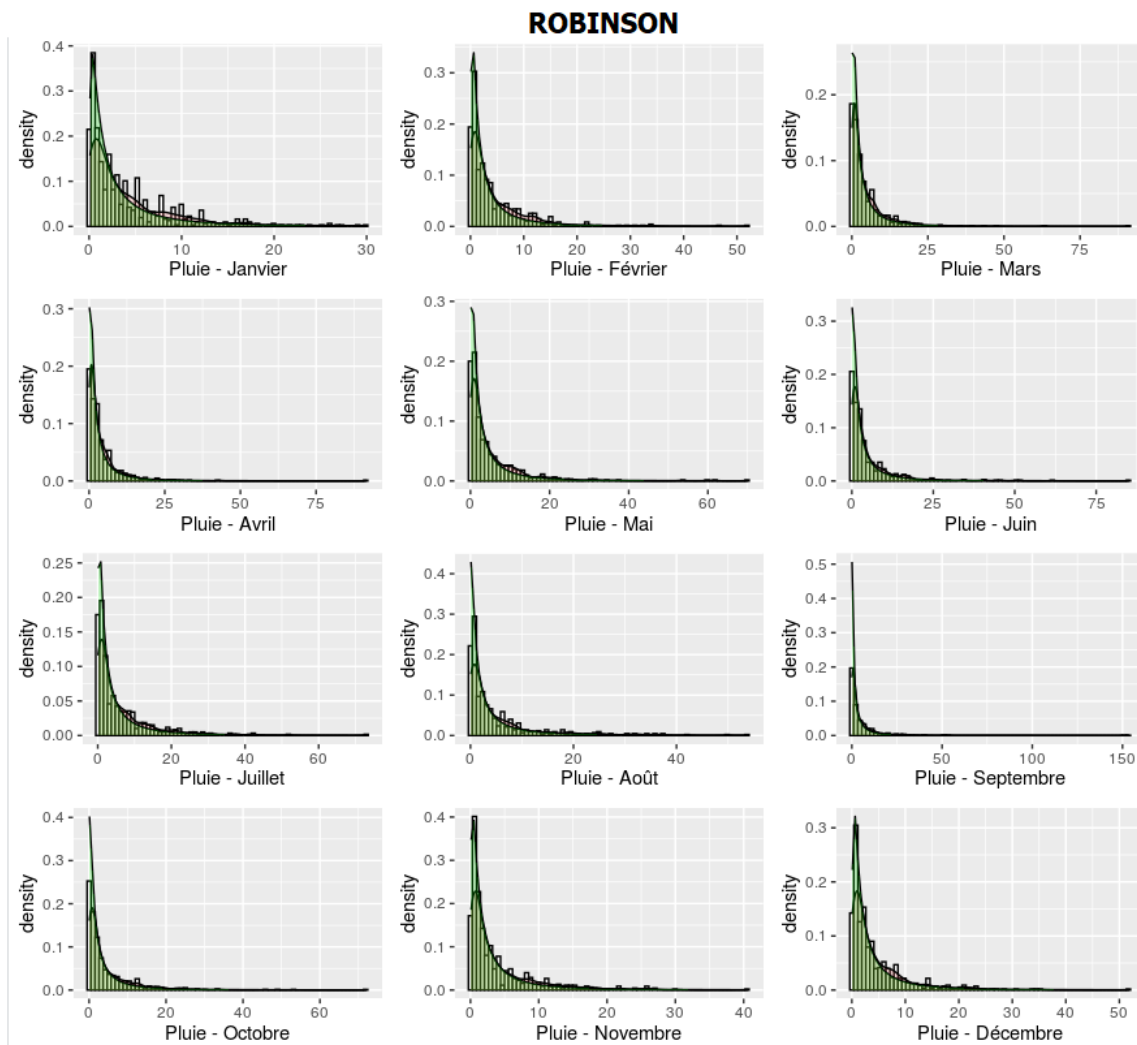


FIGURE F.7 – Adéquations de loi pour ROBINSON.

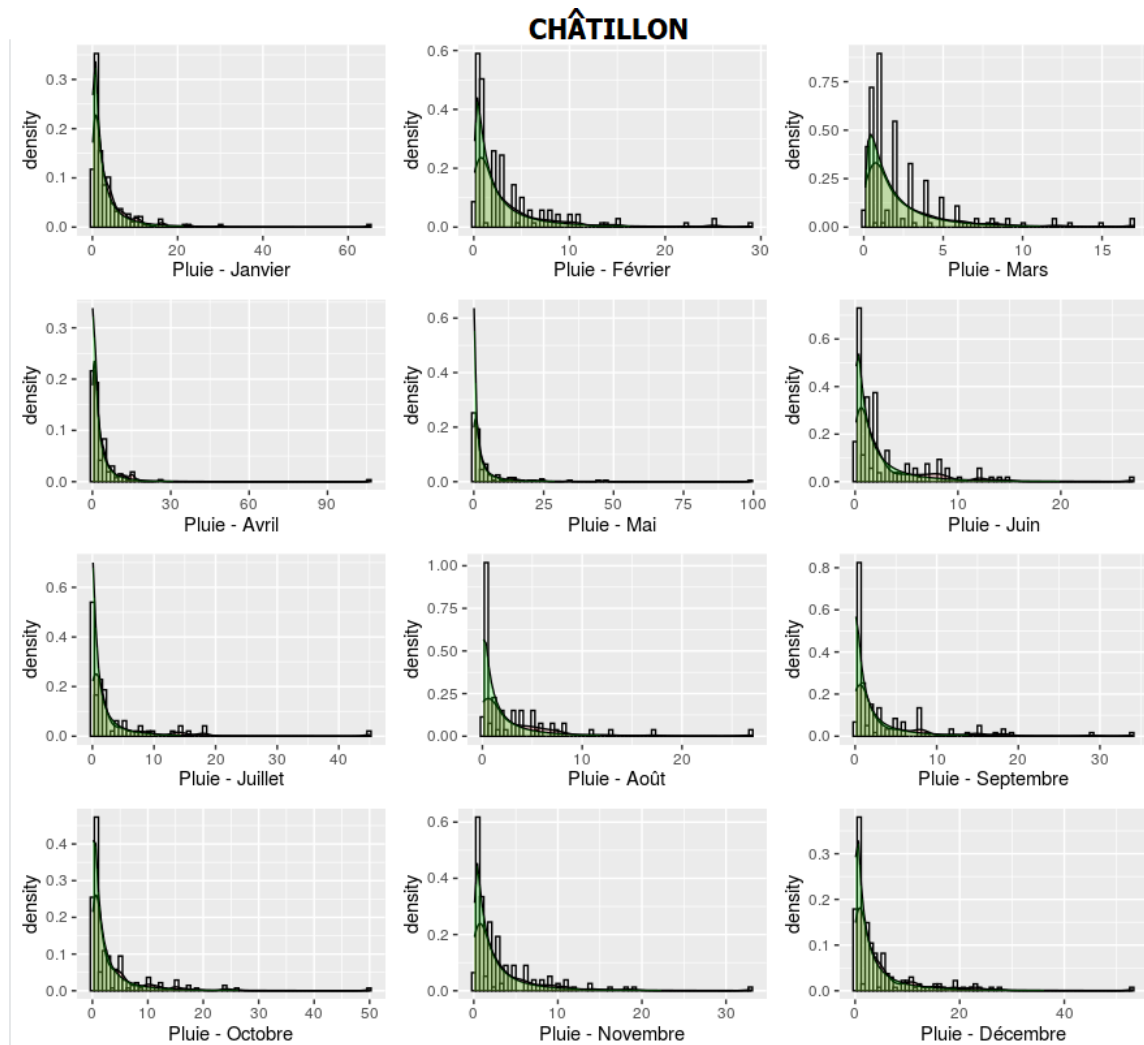


FIGURE F.8 – Adéquations de loi pour CHATILLON.

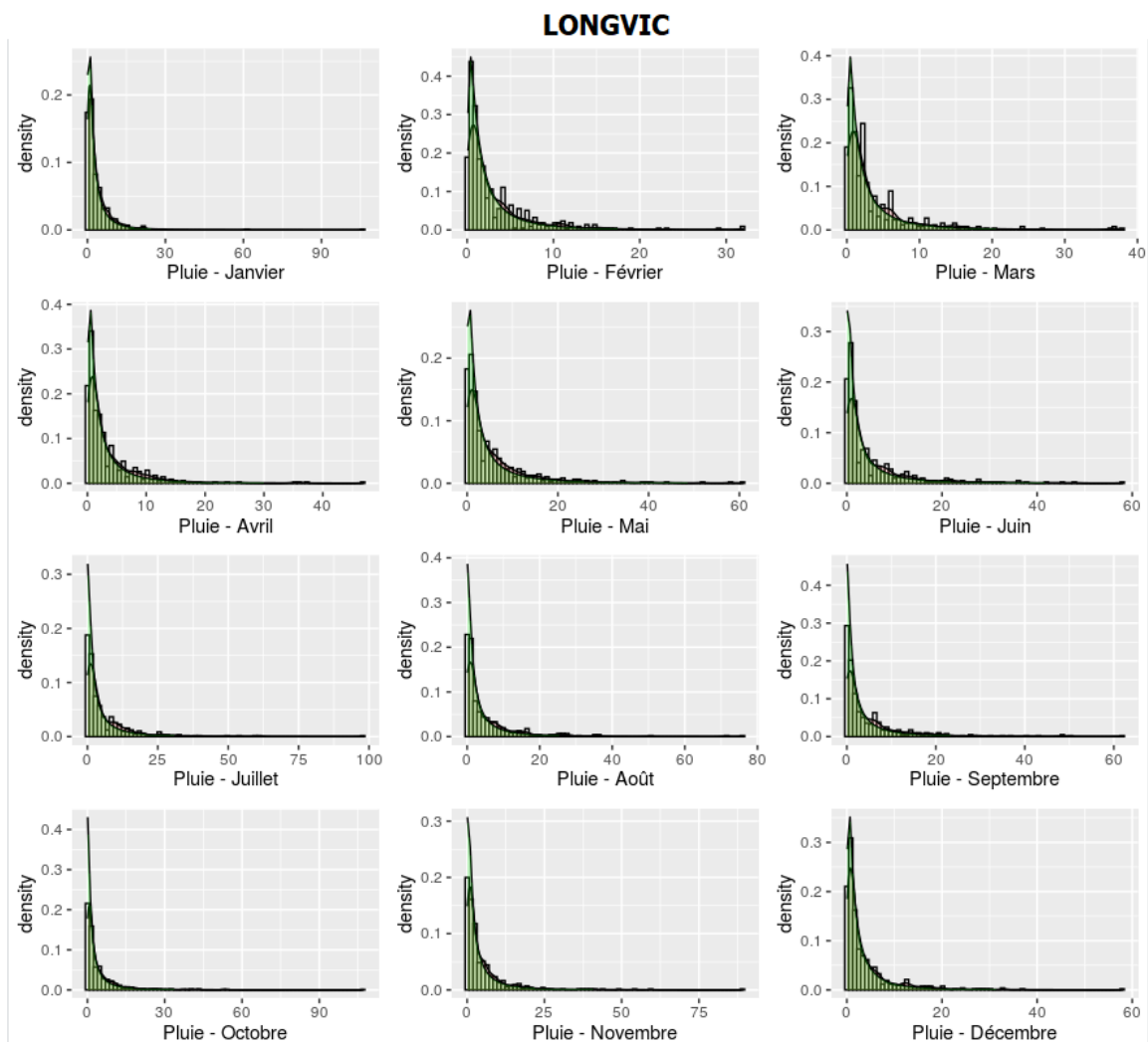


FIGURE F.9 – Adéquations de loi pour LONGVIC.

	PAR	BRE	VIL
Janvier	$Q_S^M \sim \mathcal{LN}(0.31, 1.30)$	$Q_S^M \sim \mathcal{WI}(0.93, 0.62)$	$Q_S^M \sim \mathcal{WI}(0.93, 0.62)$
Février	$Q_S^M \sim \mathcal{LN}(0.53, 1.26)$	$Q_S^M \sim \mathcal{LN}(0.11, 1.15)$	$Q_S^M \sim \mathcal{LN}(0.11, 1.15)$
Mars	$Q_S^M \sim \mathcal{LN}(0.56, 1.26)$	$Q_S^M \sim \mathcal{LN}(0.25, 1.21)$	$Q_S^M \sim \mathcal{WI}(0.92, 0.71)$
Avril	$Q_S^M \sim \mathcal{LN}(0.42, 1.32)$	$Q_S^M \sim \mathcal{LN}(0.13, 1.24)$	$Q_S^M \sim \mathcal{LN}(0.13, 1.24)$
Mai	$Q_S^M \sim \mathcal{LN}(0.62, 1.52)$	$Q_S^M \sim \mathcal{LN}(0.26, 1.20)$	$Q_S^M \sim \mathcal{LN}(0.26, 1.20)$
Juin	$Q_S^M \sim \mathcal{LN}(0.46, 1.42)$	$Q_S^M \sim \mathcal{LN}(0.38, 1.26)$	$Q_S^M \sim \mathcal{LN}(0.38, 1.26)$
Juillet	$Q_S^M \sim \mathcal{LN}(0.43, 1.49)$	$Q_S^M \sim \mathcal{WI}(0.79, 0.55)$	$Q_S^M \sim \mathcal{LN}(0.70, 2.39)$
Août	$Q_S^M \sim \mathcal{LN}(0.54, 1.52)$	$Q_S^M \sim \mathcal{WI}(0.86, 0.64)$	$Q_S^M \sim \mathcal{LN}(0.72, 2.56)$
Septembre	$Q_S^M \sim \mathcal{LN}(0.42, 1.48)$	$Q_S^M \sim \mathcal{WI}(0.85, 0.61)$	$Q_S^M \sim \mathcal{LN}(0.74, 2.37)$
Octobre	$Q_S^M \sim \mathcal{LN}(0.35, 1.40)$	$Q_S^M \sim \mathcal{W}(0.77, 1.91)$	$Q_S^M \sim \mathcal{BU}(2.45, 0.95, 0.2)$
Novembre	$Q_S^M \sim \mathcal{LN}(0.42, 1.37)$	$Q_S^M \sim \mathcal{LN}(0.27, 1.09)$	$Q_S^M \sim \mathcal{WI}(0.93, 0.75)$
Décembre	$Q_S^M \sim \mathcal{LN}(0.54, 1.37)$	$Q_S^M \sim \mathcal{LN}(0.26, 1.20)$	$Q_S^M \sim \mathcal{LN}(0.26, 1.20)$

TABLE F.1 – Tableau récapitulatif des adéquations de loi de la pluviométrie (1/3).

	BRI	AUX	BAR
Janvier	$Q_S^M \sim \mathcal{LN}(0.47, 1.28)$	$Q_S^M \sim \mathcal{LN}(0.13, 1.26)$	$Q_S^M \sim \mathcal{LN}(0.13, 1.26)$
Février	$Q_S^M \sim \mathcal{LN}(0.53, 1.26)$	$Q_S^M \sim \mathcal{LN}(0.12, 1.16)$	$Q_S^M \sim \mathcal{LN}(0.12, 1.16)$
Mars	$Q_S^M \sim \mathcal{LN}(0.54, 1.32)$	$Q_S^M \sim \mathcal{LN}(0.26, 1.22)$	$Q_S^M \sim \mathcal{LN}(0.26, 1.22)$
Avril	$Q_S^M \sim \mathcal{LN}(0.52, 1.34)$	$Q_S^M \sim \mathcal{BU}(0.97, 1.37, 0.9)$	$Q_S^M \sim \mathcal{LN}(0.13, 1.24)$
Mai	$Q_S^M \sim \mathcal{LN}(0.66, 1.40)$	$Q_S^M \sim \mathcal{LN}(0.26, 1.20)$	$Q_S^M \sim \mathcal{LN}(0.27, 1.20)$
Juin	$Q_S^M \sim \mathcal{LN}(0.47, 1.45)$	$Q_S^M \sim \mathcal{LN}(0.39, 1.27)$	$Q_S^M \sim \mathcal{LN}(0.39, 1.27)$
Juillet	$Q_S^M \sim \mathcal{LN}(0.80, 1.54)$	$Q_S^M \sim \mathcal{LN}(0.12, 1.46)$	$Q_S^M \sim \mathcal{LN}(0.13, 1.46)$
Août	$Q_S^M \sim \mathcal{LN}(0.45, 1.54)$	$Q_S^M \sim \mathcal{LN}(0.23, 1.38)$	$Q_S^M \sim \mathcal{LN}(0.23, 1.38)$
Septembre	$Q_S^M \sim \mathcal{LN}(0.54, 1.46)$	$Q_S^M \sim \mathcal{LN}(0.17, 1.36)$	$Q_S^M \sim \mathcal{WI}(0.86, 0.61)$
Octobre	$Q_S^M \sim \mathcal{LN}(0.40, 1.51)$	$Q_S^M \sim \mathcal{LN}(0.35, 1.40)$	$Q_S^M \sim \mathcal{WI}(0.78, 0.71)$
Novembre	$Q_S^M \sim \mathcal{LN}(0.48, 1.36)$	$Q_S^M \sim \mathcal{LN}(0.28, 1.09)$	$Q_S^M \sim \mathcal{LN}(0.28, 1.09)$
Décembre	$Q_S^M \sim \mathcal{LN}(0.46, 1.40)$	$Q_S^M \sim \mathcal{LN}(0.27, 1.20)$	$Q_S^M \sim \mathcal{LN}(0.27, 1.20)$

TABLE F.2 – Tableau récapitulatif des adéquations de loi de la pluviométrie (2/3).

	ROB	CHS	LON
Janvier	$Q_S^M \sim \mathcal{LN}(0.64, 1.36)$	$Q_S^M \sim \mathcal{LN}(0.54, 1.29)$	$Q_S^M \sim \mathcal{LN}(0.59, 1.26)$
Février	$Q_S^M \sim \mathcal{LN}(0.62, 1.38)$	$Q_S^M \sim \mathcal{LN}(0.46, 1.27)$	$Q_S^M \sim \mathcal{LN}(0.41, 1.25)$
Mars	$Q_S^M \sim \mathcal{LN}(0.69, 1.36)$	$Q_S^M \sim \mathcal{LN}(0.32, 1.09)$	$Q_S^M \sim \mathcal{LN}(0.54, 1.30)$
Avril	$Q_S^M \sim \mathcal{LN}(0.60, 1.37)$	$Q_S^M \sim \mathcal{LN}(0.45, 1.32)$	$Q_S^M \sim \mathcal{LN}(0.49, 1.31)$
Mai	$Q_S^M \sim \mathcal{LN}(0.74, 1.43)$	$Q_S^M \sim \mathcal{LN}(0.34, 1.60)$	$Q_S^M \sim \mathcal{LN}(0.86, 1.44)$
Juin	$Q_S^M \sim \mathcal{LN}(0.68, 1.44)$	$Q_S^M \sim \mathcal{LN}(0.22, 1.32)$	$Q_S^M \sim \mathcal{LN}(0.69, 1.47)$
Juillet	$Q_S^M \sim \mathcal{LN}(0.89, 1.45)$	$Q_S^M \sim \mathcal{WI}(0.82, 0.56)$	$Q_S^M \sim \mathcal{LN}(0.81, 1.52)$
Août	$Q_S^M \sim \mathcal{LN}(0.62, 1.54)$	$Q_S^M \sim \mathcal{WI}(0.84, 0.62)$	$Q_S^M \sim \mathcal{LN}(0.64, 1.50)$
Septembre	$Q_S^M \sim \mathcal{LN}(0.49, 1.55)$	$Q_S^M \sim \mathcal{BU}(0.13, 4.58, 4.0)$	$Q_S^M \sim \mathcal{LN}(0.58, 1.55)$
Octobre	$Q_S^M \sim \mathcal{LN}(0.57, 1.47)$	$Q_S^M \sim \mathcal{LN}(0.37, 1.34)$	$Q_S^M \sim \mathcal{LN}(0.52, 1.47)$
Novembre	$Q_S^M \sim \mathcal{LN}(0.52, 1.37)$	$Q_S^M \sim \mathcal{LN}(0.40, 1.26)$	$Q_S^M \sim \mathcal{LN}(0.67, 1.41)$
Décembre	$Q_S^M \sim \mathcal{LN}(0.73, 1.33)$	$Q_S^M \sim \mathcal{LN}(0.67, 1.39)$	$Q_S^M \sim \mathcal{LN}(0.53, 1.30)$

TABLE F.3 – Tableau récapitulatif des adéquations de loi de la pluviométrie (3/3).

Annexe G

Test de Kolmogorov-Smirnov

Le test de Kolmogorov-Smirnov est un test statistique mesurant la qualité de l'adéquation de loin en comparant les fonctions de répartition théorique et empirique via une quantité D_n appelée *statistique de Kolmogorov*. Étant donnée un échantillon de taille n et sa fonction de répartition empirique $F_n(x)$, on va tester l'hypothèse H_0 : L'échantillon provient de la loi de fonction de répartition $F(x)$ où $F(x)$ est défini à l'avance. Pour nous, F représente la fonction de répartition de la loi optimale obtenue par la méthode du maximum de vraisemblance. On a :

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \quad (\text{G.1})$$

D correspond à l'écart maximum en un point x entre la fonction de répartition théorique et la fonction de répartition empirique. Une valeur élevée de D indique que la distribution de l'échantillon s'éloigne sensiblement de la distribution de référence $F(x)$, et donc qu'il est donc peu probable que H_0 soit réalisé. Mathématiquement, on a pour toute constante $c > 0$:

$$P\left(D_n > \frac{c}{\sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} \alpha(c) = 2 \sum_{r=1}^{+\infty} (-1)^{r-1} \exp(-2r^2 c^2) \quad (\text{G.2})$$

Ainsi, si $D_n > \frac{c}{\sqrt{n}}$, on rejette H_0 . Il existe une correspondance entre le seuil critique α que l'on souhaite retenir et la constante c . Par exemple, pour un seuil critique = 0,05%, on a $c = 1,36$.

Annexe H

Théorie sur les copules

Mesures de dépendance

Corrélation de Pearson : soit X, Y deux variables aléatoires réelles, d'écart-types respectifs σ_X et σ_Y . La corrélation entre X et Y est alors donnée par :

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

où $\text{Cov}(X, Y)$ désigne la covariance entre X et Y .

Exemple : on considère cinq réalisations de deux variables aléatoires réelles X et Y .

Variables	Observations
X	10 30 50 40 20
Y	7 9 8 10 6

On obtient alors pour ces deux échantillons une corrélation linéaire de :

$$\hat{r}_p = 0.6$$

où \hat{r}_p est l'estimation de $\text{Cor}(X, Y)$.

Cette estimation empirique de la corrélation est obtenue comme suit :

$$\hat{r}_p = \hat{\sigma}_{XY} \hat{\sigma}_X \hat{\sigma}_Y$$

avec

$$\hat{\sigma}_{XY} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

tel que

$$\hat{\sigma}_X = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}, \quad \hat{\sigma}_Y = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$$

et

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

Comme nous l'avons précisé plus tôt, la mesure de corrélation linéaire n'est pas vraiment appropriée pour l'étude de variables non normales. On utilise alors d'autres mesures d'associations comme le tau (τ) de Kendall et le rhô (ρ) de Spearman.

Le tau de Kendall peut être considéré comme la différence entre la probabilité de concordance entre deux couples de variables aléatoires (X_1, X_2) et (Y_1, Y_2) et la probabilité de discordance entre ces mêmes deux couples.

Tau de Kendall : soit $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ un ensemble d'observations de variables jointes (X, Y) . Les paires d'observations (x_i, y_i) et (x_j, y_j) sont dites concordantes si :

$$x_i < x_j \text{ et } y_i < y_j \text{ ou si } x_i > x_j \text{ et } y_i > y_j$$

Elles sont dites discordantes si :

$$x_i < x_j \text{ et } y_i > y_j \text{ ou si } x_i > x_j \text{ et } y_i < y_j$$

Dans le cas où $x_i = x_j$ ou bien $y_i = y_j$, la paire n'est dite ni concordante ni discordante. Le tau de Kendall est alors défini comme :

$$\tau = \frac{(\text{nbre de paires concordantes}) - (\text{nbre de paires discordantes})}{\frac{1}{2} \times n \times (n - 1)}$$

où n est le nombre d'observations du couple (X, Y) .

Exemple : on considère deux variables aléatoires réelles X et Y , et on présente cinq observations de ces deux variables.

Variables	Observations
X	10 30 50 40 20
Y	7 9 8 10 6

On obtient alors pour ces deux échantillons un tau de Kendall de :

$$\tau = \frac{7-3}{\frac{1}{2} \times 5 \times 4} = 0.4$$

La 3^{ème} mesure d'association que nous avons utilisé est le rhô de Spearman. Le rhô de Spearman est, comme le tau de Kendall, une mesure d'association basée sur les probabilités de concordance et de discordance de deux couples de variables aléatoires. Avant de définir le rhô de Spearman, il est important de définir la variable aléatoire rg_X obtenue à partir de X .

On suppose que l'on a des réalisations (x_1, x_2, \dots, x_n) d'une variable aléatoire X . On obtient alors des réalisations de la variable aléatoire rg_X en donnant, pour chaque x_i , le rang de cette réalisation dans le vecteur des réalisations de X trié par ordre croissant :

Variables	Observations
X	10 30 50 40 20
rg_X	1 3 5 4 2

(10, 30, 50, 40, 20) sont des réalisations de X .

(1, 3, 5, 4, 2) sont des réalisations de rg_X .

Rh  de Spearman : le rh  de Spearman de deux variables al atoires X et Y est d fini comme la corr lation lin aire entre les rangs de X et Y , c'est- -dire :

$$\rho = Cor(rg_X, rg_Y) = \frac{Cov(rg_X, rg_Y)}{\sigma_{rg_X} \sigma_{rg_Y}}$$

Exemple : on consid re deux variables al atoires r elles X et Y , et on pr sente cinq observations de ces deux variables.

Variables	Observations
X	10 30 50 40 20
Y	7 9 8 10 6
rg_X	1 3 5 4 2
rg_Y	2 4 3 5 1

On obtient alors pour ces deux  chantillons un rh  de Spearman de :

$$\rho = \hat{r}_p(rg_X, rg_Y) = 0.6$$

Copules

Copule : Soit $I = [0, 1] \subset \mathbb{R}$. Une copule est une fonction $C : [0, 1]^n \implies [0, 1]$, $n \in \mathbb{N}$, d finie par les conditions suivantes :

1. C est attach e, c'est- -dire :

$$C(u_1, \dots, u_n) = 0 \text{ si une des composantes } u_i \in I, \forall i \in [1, n] \text{ est nulle.}$$

2. Les marges sont uniformes, c'est- -dire :

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \forall u_i \in I, \forall i \in [1, n]$$

3. C est une fonction n -croissante sur I^n . Pour $n=2$, cela signifie : pour tout $u_1, u_2, v_1, v_2 \in I$ tels que $u_1 \leq u_2$ et $v_1 \leq v_2$, on a

$$C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0$$

Théorème de Sklar : Soit n variables aléatoires (X_1, X_2, \dots, X_n) telles que, pour $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, la fonction de répartition conjointe est donnée par :

$$F(x_1, x_2, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

La fonction de répartition marginale de la i^e variable aléatoire, pour $i \in 1, 2, \dots, n$, est donnée par :

$$F_i(x_i) = \mathbb{P}(X_i \leq x_i).$$

On suppose n variables aléatoires X_1, \dots, X_n , et on note F_i la fonction de répartition de X_i . Si C est une copule, alors :

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

où F est une fonction de répartition multivariée de dimension n , dont les marges sont précisément F_1, \dots, F_n .

Réciproquement : si F est une fonction de répartition en dimension n , alors il existe une copule C telle que :

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

où les F_i sont les lois marginales de F .

Si ces lois marginales sont toutes continues, la copule C est alors unique, et donnée par la relation :

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)).$$

Dans ce cas, on pourra alors parler de la copule associée au vecteur aléatoire (X_1, \dots, X_n) .

Copule d'indépendance : Soit deux variables aléatoire X, Y indépendantes. Alors on a pour tout couple (x, y) :

$$\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x) \times \mathbb{P}(Y \leq y)$$

Donc on peut écrire la copule d'indépendance (figure 2.1) comme suit :

$$\forall (u, v) \in [0, 1]^2, C(u, v) = u \times v$$

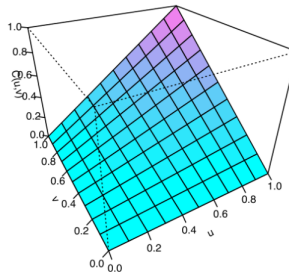


FIGURE H.1 – Copule d'indépendance. *Source : Matteo Zandi.*

Copule gaussienne : La copule gaussienne est sous-jacente à la distribution normale multi-variée. En effet, modéliser la structure de dépendance d'un échantillon par une copule gaussienne est cohérent avec la mesure de dépendance de Pearson (corrélation linéaire). La fonction de distribution de la copule gaussienne d-dimensionnelle s'écrit :

$$\forall (u_1, \dots, u_d) \in [0, 1]^d, C(u_1, \dots, u_d) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$$

où Φ^{-1} est l'inverse de la distribution normale centrée réduite uni-variée et $\phi_{\Sigma}(x) = \frac{\exp(-\frac{1}{2}x\Sigma^{-1}x')}{(2\pi)^{d/2}\det(\Sigma)^{1/2}}$ est la fonction de répartition de la loi normale multivariée centrée réduite de matrice de variance covariance Σ (matrice des corrélations de Pearson).

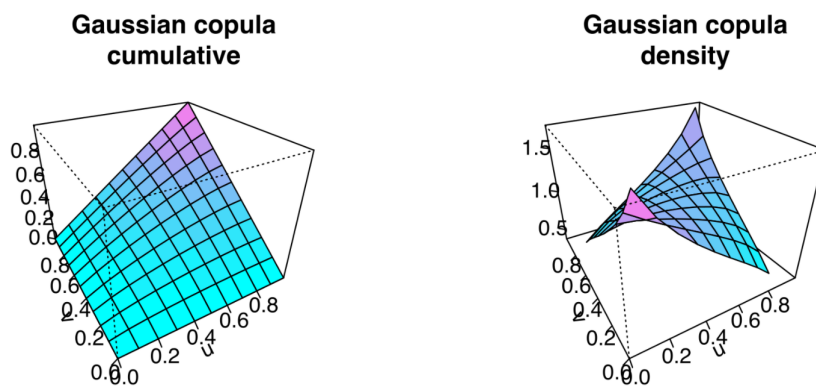


FIGURE H.2 – Copule gaussienne. *Source : Matteo Zandi.*

Annexe I

Dépendances entre les stations

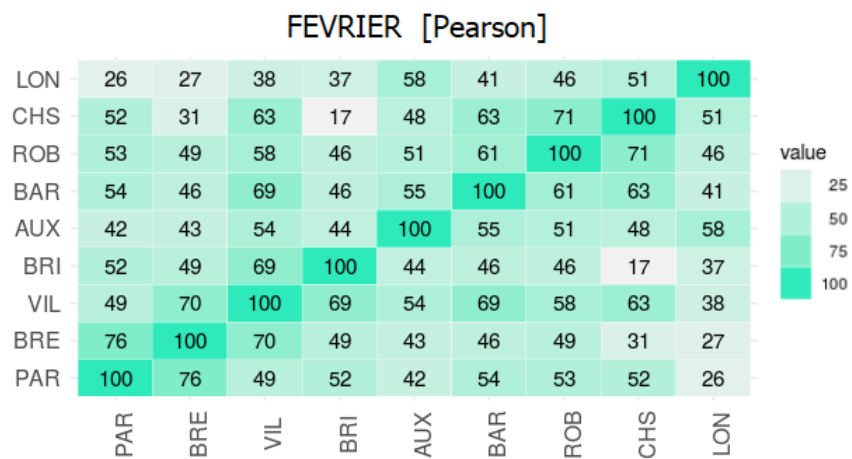


FIGURE I.1 – Calcul des corrélations de Pearson entre les stations (mois de février).

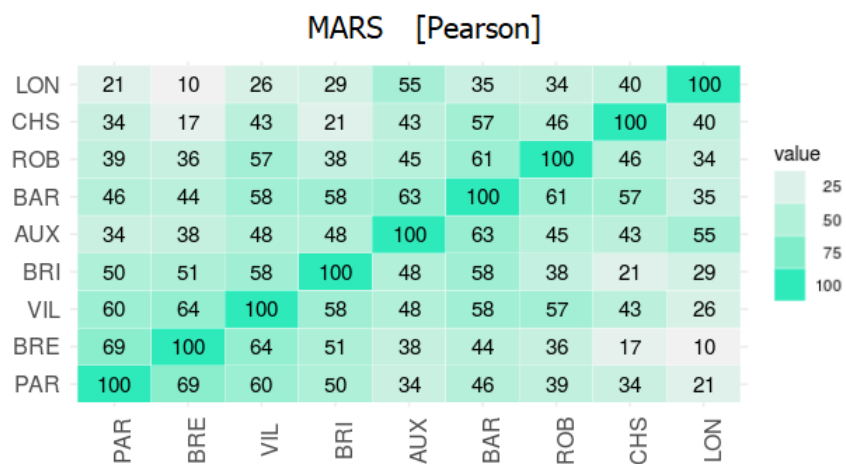


FIGURE I.2 – Calcul des corrélations de Pearson entre les stations (mois de mars).

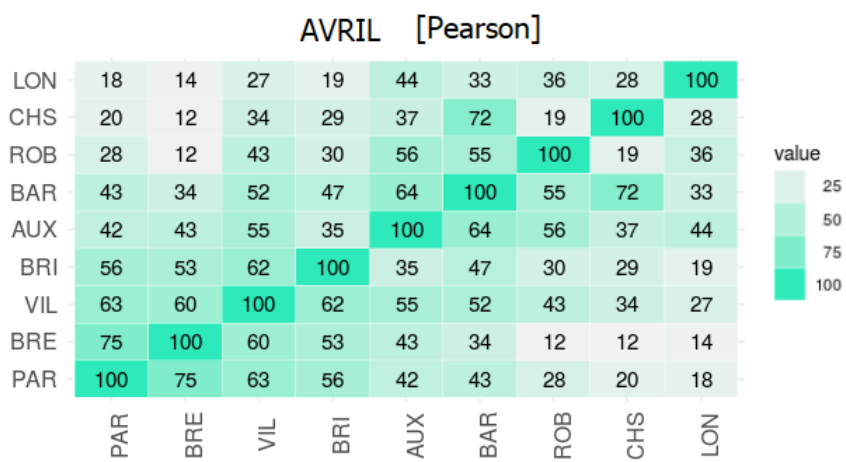


FIGURE I.3 – Calcul des corrélations de Pearson entre les stations (mois de avril).

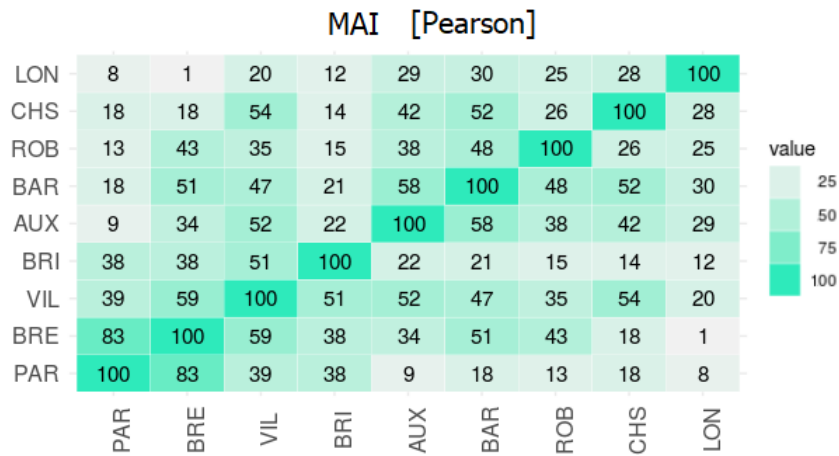


FIGURE I.4 – Calcul des corrélations de Pearson entre les stations (mois de mai).

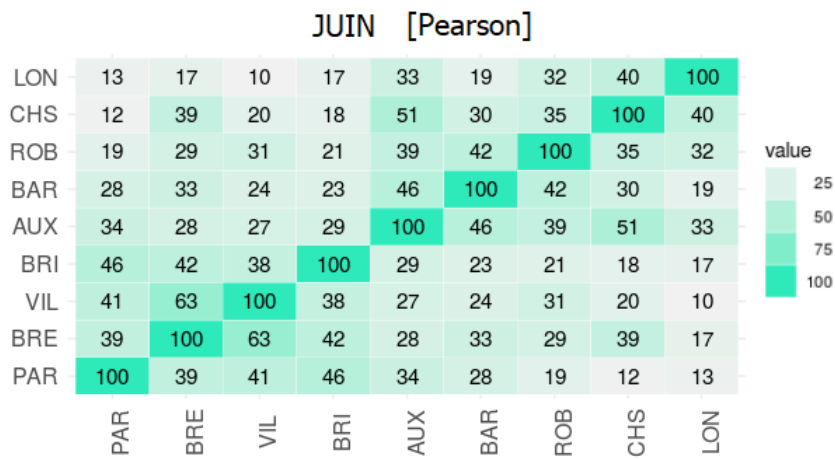


FIGURE I.5 – Calcul des corrélations de Pearson entre les stations (mois de juin).

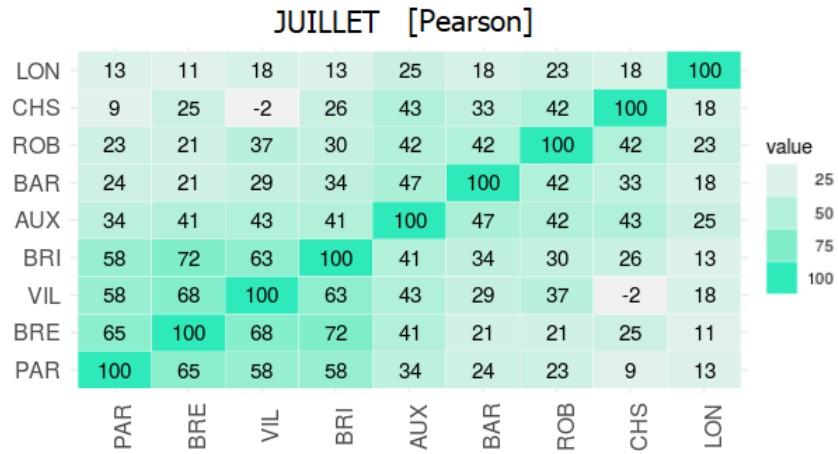


FIGURE I.6 – Calcul des corrélations de Pearson entre les stations (mois de juillet).

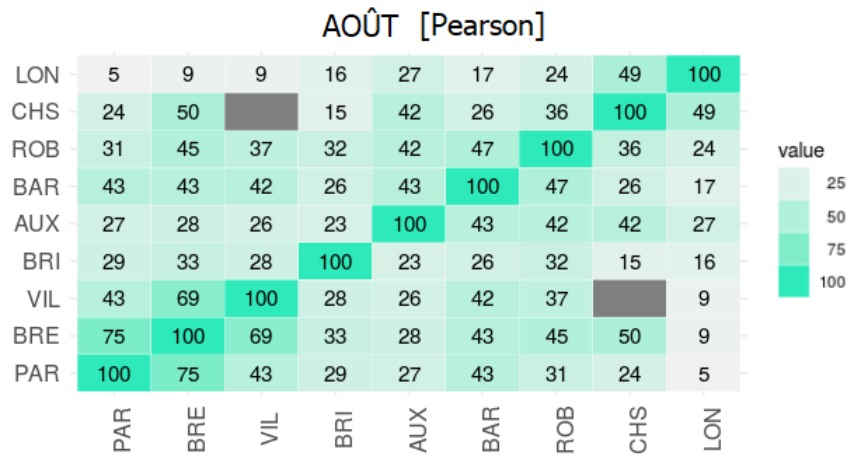


FIGURE I.7 – Calcul des corrélations de Pearson entre les stations (mois de août).

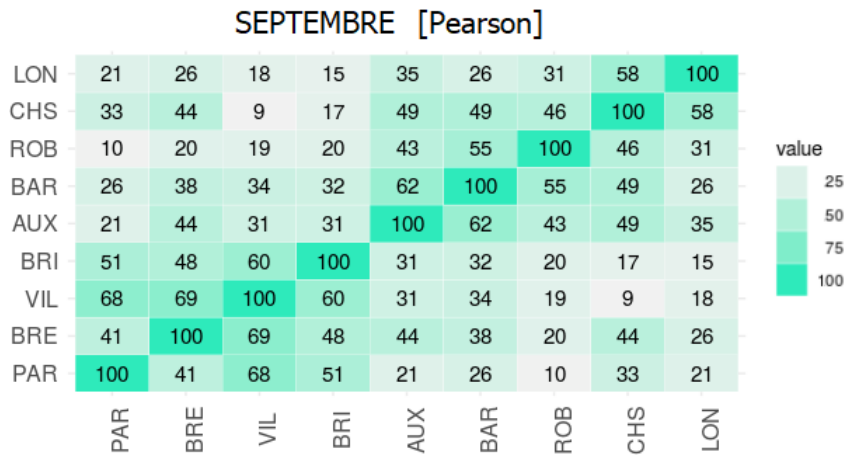


FIGURE I.8 – Calcul des corrélations de Pearson entre les stations (mois de septembre).

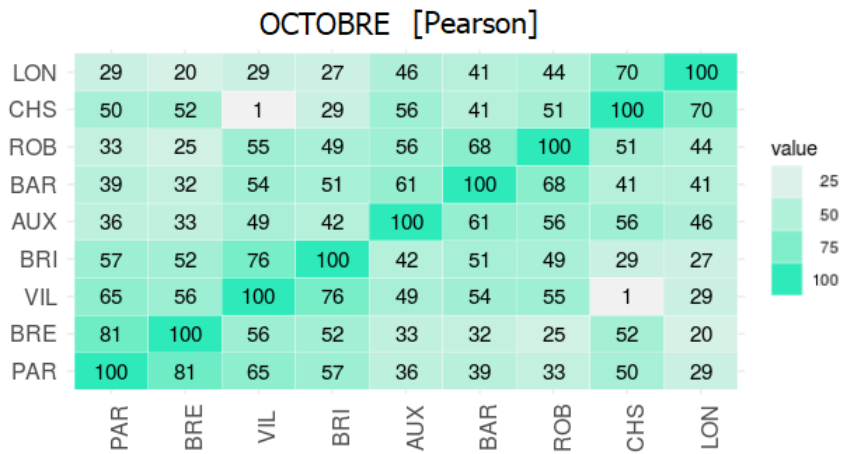


FIGURE I.9 – Calcul des corrélations de Pearson entre les stations (mois de octobre).

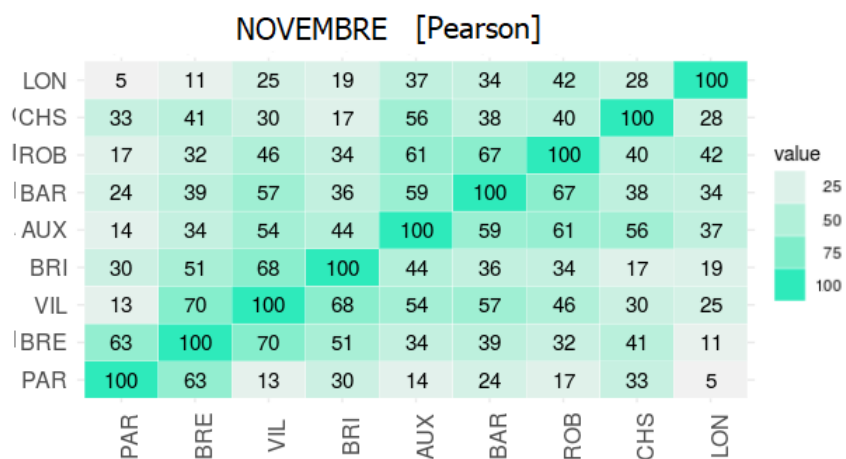


FIGURE I.10 – Calcul des corrélations de Pearson entre les stations (mois de novembre).

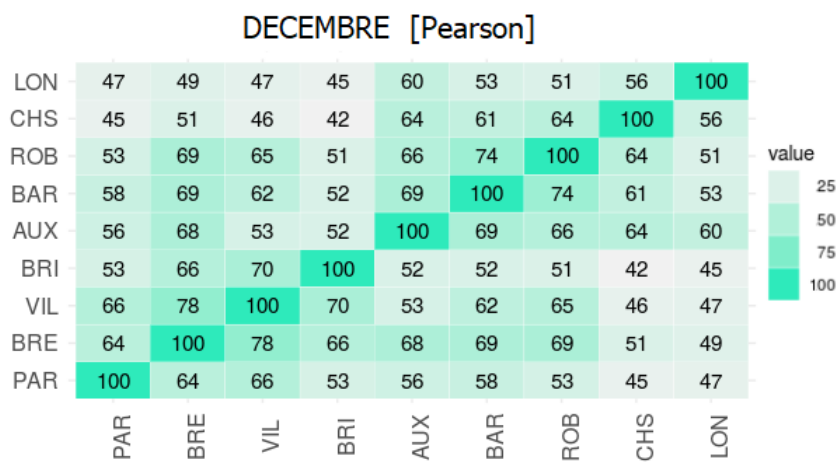


FIGURE I.11 – Calcul des corrélations de Pearson entre les stations (mois de décembre).

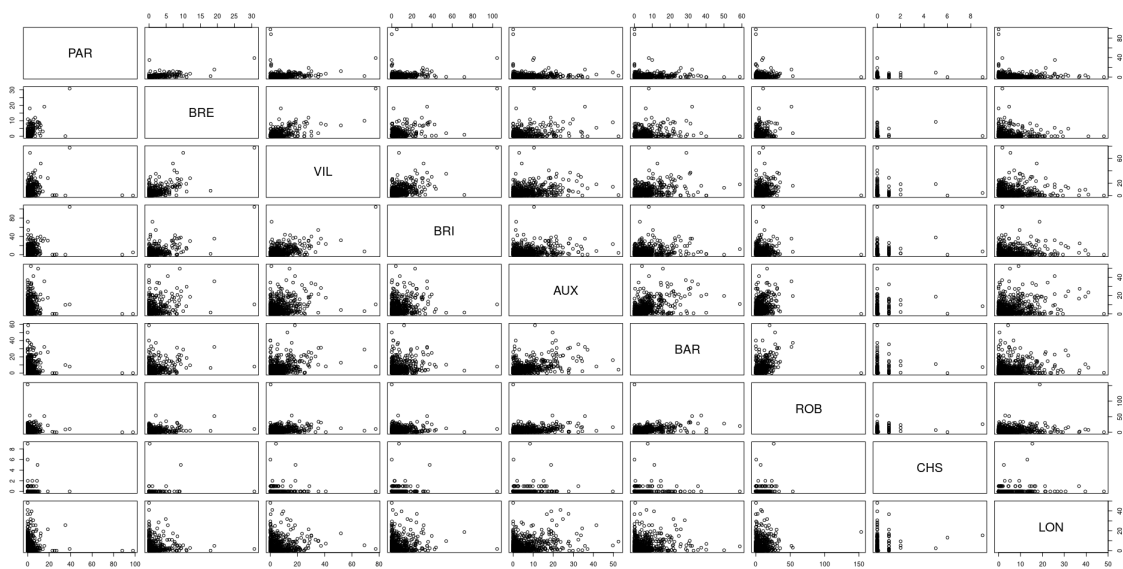


FIGURE I.12 – Observation de la structure de dépendance de la pluviométrie entre les stations.

Annexe J

Régression linéaire - Méthode des moindres carrés

Pour déterminer les coefficients β , on utilise la méthode des moindres carrés afin d'optimiser la régression. On cherche alors à minimiser l'erreur d'estimation en fonction de β . Par définition du modèle, $W = Y - X\beta$. On a alors :

$$\begin{aligned}\Delta(\beta) &= (Y - X\beta)'(Y - X\beta) \\ &= (Y' - \beta'X')(Y - X\beta) \\ &= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta \\ &= Y'Y - 2\beta'X'Y + \beta'X'X\beta\end{aligned}$$

Afin de minimiser cette valeur, on dérive en fonction de β et on cherche la valeur de β qui rend nulle cette quantité.

$$\begin{aligned}\frac{\delta(\Delta(\beta))}{\delta\beta} &= 0 \\ -2X'Y + 2\beta'X'X &= 0 \\ \Rightarrow \hat{\beta} &= (X'X)^{-1}X'Y\end{aligned}$$

Bibliographie

- [ACPR, 2021] ACPR (2021). Les principaux résultats de l'exercice pilote climatique 2020. *Rapport officiel*.
- [Berruyer, 2016] BERRUYER, O. (2016). Seine : une crue 2016 demi-millénaire. *Article pour Les Crises*.
- [Brun et Gache, 2013] BRUN, A. et GACHE, F. (2013). Risque inondation dans le grand paris : la résilience est-elle un concept opératoire? *Article*.
- [CCR, 2016a] CCR (2016a). Etude historique de la crue de 1910 & modélisation de scénarios de référence. *Rapport officiel*.
- [CCR, 2016b] CCR (2016b). Inondations de mai-juin 2016 en france. *Rapport officiel*.
- [CCR, 2018] CCR (2018). Retour sur les inondations de janvier et février 2018. *Rapport officiel*.
- [CCR, 2020] CCR (2020). Statistiques générales sur la sinistralité cat nat. *Rapport officiel*.
- [CEPRI, 2010] CEPRI (2010). Le bâtiment face à l'inondation. *Rapport officiel*.
- [D'Antin, 2014] D'ANTIN, H. (2014). Modélisation du risque de crue de la seine en région parisienne. *Mémoire d'actuariat*.
- [DRIEE, 2016] DRIEE (2016). Mission sur le fonctionnement hydrologique du bassin de la seine. *Rapport au premier ministre*.
- [DRIEE, 2017] DRIEE (2017). Métropole francilienne : une stratégie pour limiter les risques d'inondation 2016-2021. *Rapport officiel*.
- [El Hassani, 2017] EL HASSANI, H. (2017). Modélisation stochastique des inondations en france et applications en réassurance. *Mémoire d'actuariat*.
- [Fallot et Sautebin, 2018] FALLOT, J.-M. et SAUTEBIN, T. (2018). Evaluation des différentes lois statistiques pour l'estimation des précipitations journalières extrêmes en suisse. *Étude à l'université de Lausanne*.
- [Fujiki, 2017] FUJIKI, K. (2017). Etude prospective des impacts sociaux d'une inondation majeure en région ile-de-france. *Thèse de doctorat*.
- [GIEC, 2021] GIEC (2021). Climate change 2021 - the physical science basis. *Rapport officiel*.

- [Goubet, 1981] GOUBET, A. (1981). Les crues dans le bassin de la seine du 17e au début du 19e siècle. *Article pour La Houille Blanche*.
- [MTES, 2017a] MTES (2017a). Construire en zone inondable et réhabiliter après inondation. *Projet d'un document technique unifié (DTU)*.
- [MTES, 2017b] MTES (2017b). Inondations de mai et juin 2016 dans les bassins moyens de la seine et de la loire. *Rapport officiel*.
- [MTES, 2020] MTES (2020). Risques climatiques : six français sur dix sont d'ores et déjà concernés. *Data Lab*.
- [Vivian, 1977] VIVIAN, H. (1977). Averses extensives et crues concomitantes dans l'arc alpin. *Thèse d'état*.
- [WWA, 2021] WWA (2021). Rapid attribution of heavy rainfall events leading to the severe flooding in western europe during july 2021. *Rapport officiel*.