



Mémoire d'Actuariat

Lundi 1er mars 2021

Par : Leo LOVISOLO

Titre : **Construction d'un générateur climatique permettant l'évaluation d'options basées sur un indice sécheresse en France**

Confidentialité : Non

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

Membres du jury :

Fabrice BENARD

Yi RONG

Pierre AILLIOT

Ying LI

Yann MILOE

Anthony DERIEN

Anthony NAHELOU

Entreprise :

SeaBird Conseil

Signature :

Directeur de mémoire :

Pierre THEROND

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

Signature responsable entreprise :

Signature candidat :

Résumé

La gestion des risques climatiques est un enjeu important pour un grand nombre d'acteurs économiques. Un agriculteur peut voir son chiffre d'affaire impacté par les aléas climatiques tels que la tempête, la grêle ou la sécheresse. Ces mêmes phénomènes naturels peuvent, au même titre, dégrader le bilan d'une compagnie d'assurance. C'est pour cette raison qu'en Europe, plusieurs autorités de contrôle ont récemment annoncé qu'elles auront des exigences croissantes sur la bonne gestion des risques climatiques faite par les compagnies d'assurance. L'échange de contrats financiers basés sur un indice météorologique permet alors de transférer ce risque à un autre parti et ainsi réduire son exposition aux aléas climatiques.

Aujourd'hui, les marchés organisés n'offrent qu'un choix très limité de ces produits financiers et la composition du marché de gré à gré est d'une opacité notable. Cela semble paradoxal dans un monde où l'information a tendance à être échangée de manière plus libre et transparente. La NASA a d'ailleurs œuvré dans ce sens en développant une base de données climatiques sous forme d'API, très complète et accessible à tous. Ce constat nous a amené à considérer le développement d'une méthode de tarification de dérivés climatiques directement basée sur cette API. Notre objectif est de construire un générateur climatique stochastique qui permettra l'évaluation de telles options.

L'expérience ayant montré que le péril sécheresse est l'un des plus coûteux en France, nous avons choisi de construire et de tarifier des options basées sur un indice sécheresse. Cette application concrète permettra de comparer la tarification faite par le générateur climatique avec d'autres méthodes plus traditionnelles. Bien que notre étude porte exclusivement sur le risque sécheresse en France, la méthode exposée pourra être appliquée à d'autres risques sous-jacents ou à une autre zone géographique. Après avoir détaillé le fonctionnement d'un dérivé climatique et mis en évidence les risques liés à la sécheresse, nous détaillerons l'ensemble de la démarche permettant la construction du générateur climatique. Enfin nous mettrons en contexte ces travaux au travers d'une application sur un portefeuille fictif de dérivés sécheresse.

Mots clés : dérivés climatiques, évaluation d'options, risque sécheresse, générateur climatique, CAT Bonds.

Abstract

Climate risk management is an important issue for a large number of economic players. A farmer can see his turnover impacted by climatic hazards such as storms, hail or drought. These same natural phenomena can, in the same way, degrade the balance sheet of an insurance company. For this reason, in Europe, several supervisory authorities have recently announced that they will be placing increasing demands on insurance companies to properly manage climate risks. The exchange of financial contracts based on a meteorological index then makes it possible to transfer this risk to another party and thus reduce its exposure to climatic hazards.

Today, organized markets offer only a very limited choice of these financial products and the composition of the OTC market is noticeably opaque. This seems paradoxical in a world where information tends to be exchanged more freely and transparently. NASA has in fact worked in this direction by developing a climate database in the form of an API, which is very complete and accessible to all. This observation led us to consider the development of a pricing method for weather derivatives directly based on this API. Our objective is to build a stochastic weather generator that will allow the evaluation of such options.

As experience has shown that the drought hazard is one of the costliest in France, we have chosen to build and price options based on a drought index. This concrete application will allow us to compare the pricing made by the weather generator with other more traditional methods. Although our study focuses exclusively on the drought risk in France, the exposed method can be applied to other underlying risks or to another geographical area. After having introduced weather derivatives and highlighted the risks related to drought, we will detail the entire approach used to build the weather generator. Finally, we will put this work into context through an application on a fictitious portfolio of drought derivatives.

Keywords : weather derivatives, option pricing, drought risk, weather generator, CAT Bonds.

Table des matières

Remerciements	6
Lexique	7
I Mise en contexte	9
1 Présentation du risque sécheresse en France	10
1.1 Définition	10
1.2 Impact	11
1.3 Prise en charge	13
1.4 Évolution du péril en France	15
1.5 Disparités spatiales	16
2 Présentation des dérivés climatiques	18
2.1 Définition	18
2.2 Indices sous-jacents fréquemment utilisés	20
2.3 Marché des dérivés climatiques	21
2.3.1 Premiers échanges	21
2.3.2 Un marché organisé historique : Le Chicago Mercantile Exchange	21
2.3.3 Prédominance du marché de gré à gré	22
II Théorie	23
3 Stratégies de couverture par des produits dérivés climatiques	24
3.1 Options simples	24
3.1.1 Option d'achat	24
3.1.2 Option de vente	25
3.1.3 Différents types d'exercices pour une option	25
3.2 Produits combinés	26
3.2.1 Écart vertical	26
4 Évaluation d'une option climatique	27
4.1 Méthode Burn	27
4.2 Modélisation actuarielle du sous-jacent	28
4.3 Générateur climatique stochastique	30

III	Cadre d'étude	31
5	Présentation de la base de données	32
5.1	Présentation générale	32
5.1.1	Source	32
5.1.2	Propriétés	32
5.1.3	Implémentation sous R	33
5.2	Précisions sur les variables utiles	34
5.2.1	Température maximale	34
5.2.2	Précipitation	34
5.3	Avantages et limites de cette API	35
5.3.1	Avantages	35
5.3.2	Limites	36
5.4	Pour aller plus loin	36
6	Structuration du dérivé sécheresse	37
6.1	Choix de l'indice sécheresse sous-jacent	37
6.1.1	Calcul du KBDI	38
6.1.2	Simulation sur le sol français	39
IV	Construction d'un générateur climatique	40
7	Modélisation des jours pluvieux	42
7.1	Chaîne de Markov à deux états	42
7.1.1	Définitions préliminaires	42
7.1.2	Matrice de transition	42
7.2	Estimation des coefficients	43
7.2.1	Protocole	43
7.3	Résultats obtenus	44
8	Modélisation de l'intensité des pluies	47
8.1	Rappels sur la loi gamma	47
8.1.1	Définition de la loi gamma	47
8.2	Estimation des paramètres de la loi gamma	48
8.3	Méthode des moments	48
8.3.1	Résultats obtenus	49
8.4	Ajustement de la loi gamma sur nos données de précipitations	50
9	Modélisation de la température	52
9.1	Introduction aux séries temporelles	52
9.2	Estimation de la tendance et de la saisonnalité	54
9.3	Processus de retour à la moyenne	59
9.3.1	Formulation	59
9.3.2	Estimation de la force de rappel	60
9.3.3	Modélisation du résidu ω	61
10	Construction du KBDI	63
10.1	Simulations	63
10.2	Cohérence du modèle	65

V	Application	66
11	Calcul de la prime pure de contrats basés sur le KBDI	67
11.1	Couverture considérée	67
11.2	Méthode Burning Cost	68
11.3	Modélisation du sous-jacent	69
11.4	Générateur climatique	69
11.5	Tarification de l'écart haussier	70
12	Agrégation du risque	71
12.1	Méthodologie	71
12.2	Construction de l'indice agrégé et calcul de la prime pure	72
12.3	Étude de la mutualisation du risque	73
	Conclusion	74
	Bibliographie	76
	Annexe	80

Remerciements

Tout d'abord, j'adresse mes remerciements aux enseignants de l'Euria pour leur investissement tout au long de mon cursus scolaire.

Je tiens à remercier vivement Pierre THEROND pour son accueil au sein de Sea-Bird, pour son accompagnement tout au long de la rédaction de ce mémoire et pour la confiance qu'il a accordé à mes travaux.

Je suis également reconnaissant envers Franck VERMET, Sonia GUELOU, Pierre AILLIOT, Mathieu CLAVIER, Elodie BUSSON, Antoine BADILLET et Laurent LAMBERT pour leur relecture, leurs conseils et leur précieuse contribution.

J'ai également une pensée pour Théo CÔME qui m'a aidé à comprendre le fonctionnement d'un générateur climatique et Saïda DERKAOUI qui m'a conseillé sur les parties les plus techniques du mémoire et je les en remercie.

Enfin, mes derniers remerciements vont à Bernard FINAS, David WHITEHEAD et Aurelien SCHWACHTGEN qui ont accepté de me transmettre l'expertise qu'ils ont acquis dans le domaine des risques climatiques.

Lexique

API : Application Programming Interface (Interface de programmation)

Cat Nat : Catastrophes Naturelles

CCR : Caisse Centrale de Réassurance

FFA : Fédération Française de l'Assurance

USGS : United States Geological Survey (Institut d'études géologiques des USA)

GIEC : Groupe d'experts Intergouvernemental sur l'Evolution du Climat

Première partie

Mise en contexte

Chapitre 1

Présentation du risque sécheresse en France

Nous allons commencer par définir les notions essentielles à la bonne compréhension de notre problématique. Dans ce premier chapitre, nous définirons le risque sécheresse et les nombreux impacts qu'il peut avoir. Nous étudierons également les caractéristiques de ce phénomène climatique en terme de sinistralité et comment ce risque est couvert en France.

1.1 Définition

Tout d'abord, il est important de comprendre que la sécheresse est une notion parfois complexe et pouvant être définie de différente manière suivant sa nature. Le Centre d'Information sur l'EAU [CIEAU] propose une définition pour chacun des trois types de sécheresse :

- Sécheresse météorologique :

"Correspond à une pluviométrie trop faible sur une durée prolongée".

- Sécheresse agricole :

"Lorsque le taux d'humidité dans les sols (à un mètre de profondeur) est trop faible pour assurer des bonnes conditions de cultivations. Ce type de sécheresse survient même en cas de précipitations normales, elle dépend de la nature des sols, des pratiques agricoles et du type de plantes cultivées. Par exemple, le riz et le maïs consomment beaucoup d'eau".

- Sécheresse hydrologique :

"Se manifeste lorsque le niveau des cours (eaux souterraines dans les nappes phréatiques et eaux de surface dans les rivières, lacs et fleuves) d'eau baisse significativement, on parle alors d'étiage. Un volume trop faible de précipitations est un des facteurs déclenchant ce type de sécheresse, mais également la perméabilité des sols contenant ces cours d'eau et qui conditionne la capacité d'infiltration et de ruissellement de l'eau".

Les températures durablement élevées et les faibles précipitations sont les principaux facteurs météorologiques qui engendrent la sécheresse. Pour la suite de notre raisonnement, nous retiendrons que la sécheresse est un phénomène explicable par ces deux variables et ne ferons plus de distinction entre les 3 types de sécheresse définis dans la littérature.

La définition de la sécheresse n'est pas la même pour tous les pays. En France, elle est définie par une absence de pluie prolongée. Aux États-Unis, elle est définie comme une baisse des précipitations par rapport à la moyenne historique des précipitations constatées dans cette zone géographique. Cette définition permet notamment de prendre en compte la différence de climat entre ses états.

1.2 Impact

Les conséquences de la sécheresse sont très nombreuses et parfois complexes. Il est difficile d'en faire une liste exhaustive. Nous proposons ici une énumération synthétique des principaux impacts de la sécheresse connus à ce jour.

Impact agricole

La sécheresse a un impact direct sur le rendement des récoltes et peut, dans certains cas, aller jusqu'à une destruction totale des cultures. Une réduction de la production agricole a de nombreuses conséquences : diminution des revenus pour les agriculteurs, augmentation des prix alimentaires, chômage, famine. Le bétail peut également mourir par manque de nourriture et entraîner une pénurie de viande qui s'additionnera au manque de récoltes.

Impact forestier

En période de sécheresse, les feux de forêt sont plus fréquents et plus ravageurs. Les blessures ou décès résultant de la proximité du feu et l'intoxication au monoxyde de carbone font partie des effets directs des incendies de forêts. Une augmentation des consultations médicales et des hospitalisations liées à des maladies respiratoires consécutives aux feux de forêt a été observée. Cela est notamment dû aux débris nocifs pouvant voyager dans l'air. Une exposition de quelques minutes ou quelques jours aux particules émises lors d'incendies de forêts peut engendrer des symptômes tels qu'une irritation des voies respiratoires et une difficulté à respirer [Bustinza, 2010].

Le monde agricole et forestier n'est pas le seul concerné par la sécheresse. Bien qu'indirectement touchées, la santé et l'habitation sont deux secteurs sensibles à ce phénomène climatique.

Impact sanitaire

En période de sécheresse, souvent liée à une période de fortes chaleurs, il est possible d'observer une hausse d'indemnisation liée à des malaises [Figaro, 2019]. Cependant l'impact sanitaire lié à la sécheresse peut prendre des formes plus complexes [STANKE et al., 2013] : lorsqu'une période de sécheresse est suivie d'une période de fortes pluies ou d'une inondation, plusieurs épidémies de l'encéphalite de Saint-Louis ont été observées aux États-Unis. Le virus du Nil occidental a aussi été lié à des conditions de sécheresse

en Californie, au Mississippi et au Colorado. Le chikungunya, l'encéphalite japonaise, et la fièvre de la vallée du Rift ont aussi été associés à la sécheresse. Des maladies comme la schistosomiase et certaines infections transmises par les tiques semblent également liées au temps sec, notamment en Afrique [INSPQ].

D'après l'institut national de santé publique du Québec, un temps sec précédé de faibles pluies multiplierait par trois le risque de maladies gastro-intestinales aiguës chez les personnes utilisant un puits domestique [LLOYD et al., 2007]. Il semble que de telles conditions météorologiques amènent à une baisse de la nappe phréatique et à une contamination des eaux souterraines par les eaux de surface, de même qu'une concentration des eaux usées peut contaminer les sources d'eau [FEBRIANI et al., 2010]. La sécheresse peut également contribuer au stress, à la fatigue, à la perturbation du sommeil et pourrait engendrer une augmentation de la consommation d'alcool. Des problèmes de dépression et de suicide accrus seraient également une conséquence de ce phénomène climatique [INSPQ].

Impact sur l'habitation

L'effet de sécheresse-réhydratation des sols est un phénomène peu connu des particuliers mais qui est pourtant très couteux pour les assureurs. C'est un phénomène que l'on observe sur les terrains argileux. Le sol rétrécit au moment de la sécheresse et se regonfle lors de la réhydratation. Cela crée un différentiel de support au niveau des fondations et conduit à des fissures dans le bâtiment, pouvant le rendre inhabitable ou même provoquer son écroulement. Fort heureusement, les immeubles collectifs, ayant souvent des fondations plus robustes, ne sont que très rarement touchés.

En 2003, par exemple, un épisode de sécheresse particulièrement sévère s'est traduit en France par des dégâts liés au phénomène de sécheresse-réhydratation des sols appelé phénomène de subsidence, coûtant très cher aux compagnies d'assurance [Sénat]. "En 2003, l'épisode de sécheresse extrêmement sévère s'est traduit par des dégâts liés à la subsidence qui ont coûté 2,3 milliards d'euros aux assureurs" [AFP et FFA, 2019]. Notons que les dégâts causés par la sécheresse n'étaient pas visibles immédiatement. C'est en général six mois ou un an après les périodes de sécheresse que les dégâts occasionnés sont constatés.

A ce titre, l'article 68 de la loi ELAN du 23 novembre 2018 impose une "Prévention des risques de mouvement de terrain différentiel consécutif à la sécheresse et à la réhydratation des sols". Une étude géotechnique est notamment exigée "en cas de vente d'un terrain non bâti constructible" et "avant la conclusion de tout contrat ayant pour objet des travaux de construction ou la maîtrise d'œuvre d'un ou de plusieurs immeubles à usage d'habitation ou à usage professionnel et d'habitation ne comportant pas plus de deux logements".

La complexité de l'étude du phénomène de subsidence a fait l'objet d'un mémoire présenté devant l'institut des actuaires [SCHULTE, 2016].

Évolution de la sécheresse avec le réchauffement climatique

Lord d'une interview donnée à l'agence France-Presse, la FFA a déclaré en 2019 que "sur fond de réchauffement climatique, la sécheresse est le sinistre dont le coût risque de croître le plus rapidement ces prochaines années" [AFP et FFA, 2019].

Bien qu'il soit probable qu'un réchauffement de la planète occasionne, dans l'ensemble, plus de précipitations, les sécheresses pourraient devenir plus fréquentes ou s'aggraver à cause des changements climatiques [CCME, 2003]. De fait, des températures élevées associées à une saison chaude prolongée pourraient favoriser une perte d'humidité par évaporation. Cette perte dépasserait les gains réalisés par l'augmentation des précipitations. Selon une récente étude du Centre national américain de recherche atmosphérique, la sécheresse devrait s'accroître à l'échelle de la planète au cours du XXI^e siècle, sauf dans les régions situées à des latitudes élevées où le climat deviendrait plus humide [DAI, 2010].

Les sinistres causés par la sécheresse dans le monde sont de plus en plus fréquents et intenses. Il est fort probable que la demande de couverture contre ce risque climatique croisse fortement dans les années à venir. Nous allons maintenant voir quels sont les outils proposés aujourd'hui pour nous protéger contre les risques liés à la sécheresse.

1.3 Prise en charge

En France, avant la loi du 13 juillet 1982 qui encadre les périls liés aux catastrophes naturelles, l'aide publique était l'un des seuls recours possibles face à la sinistralité liée à la sécheresse. Mais cette aide s'avérait être souvent insuffisante. Depuis, le régime des catastrophes naturelles permet aux sinistrés d'être indemnisés conjointement par des compagnies d'assurance et par l'État. En France, le risque sécheresse est pris en compte dans le régime des catastrophes naturelles. Nous proposons dans cette partie une analyse synthétique des notions importantes de ce régime concernant la couverture qu'il apporte, la manière dont il est financé et ses limites. Pour comprendre son fonctionnement de manière plus fine, il est possible de consulter [CCR] et [FFA].

Couverture

- Condition d'indemnisation :

Pour être couvert par le régime des catastrophes naturelles, un arrêté interministériel doit reconnaître l'état de catastrophe naturelle dans le journal officiel et les biens doivent être assurés par une assurance dommage. Les biens assurés uniquement au titre de responsabilité civile ne sont pas pris en compte par le régime.

- Évènements couverts par le régime :

Le régime des catastrophes naturelles n'établit pas de liste précise des périls couverts et des seuils déterminant le caractère catastrophique d'un événement. N'excluant aucun phénomène dans sa garantie, il prend uniquement en compte le critère "intensité anormale d'un agent naturel", rendant son application flexible en contrepartie d'un manque d'exhaustivité souvent critiqué.

Par expérience, le risque sécheresse fait partie des deux risques les plus significatifs de ce régime. Avec le risque inondation ils représentent 91 % de la sinistralité répertoriée par le régime des catastrophes naturelles de 1982 à 2019.



FIGURE 1.1 – CCR : Répartition par péril de la sinistralité Cat Nat non-auto cumulée de 1982 à 2019

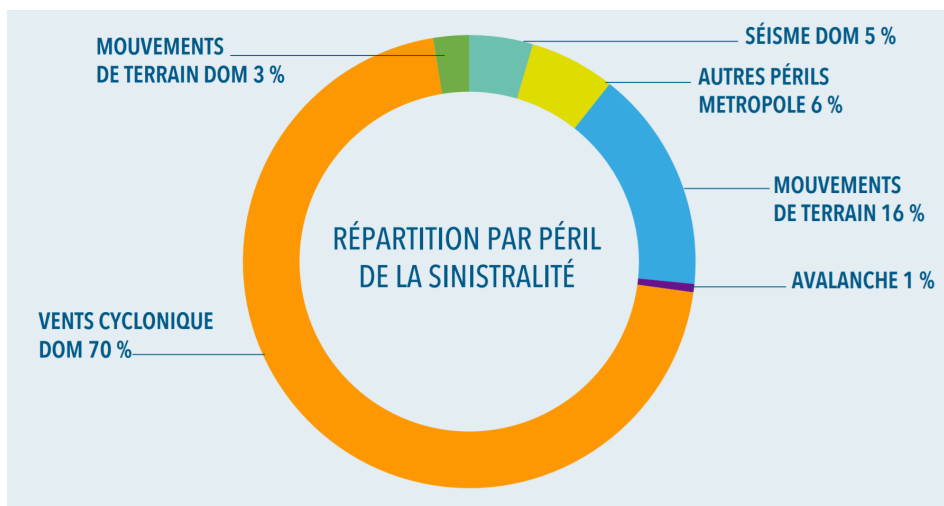


FIGURE 1.2 – CCR : Répartition par péril de la sinistralité Cat Nat non-auto cumulée de 1982 à 2019 pour la catégorie AUTRES

Notons que la sinistralité causée par les tempêtes, la grêle ou la neige n'est pas couverte par le régime des catastrophes naturelles mais dans dans une garantie spécifique appelée garantie TGN.

- Dommages couverts par le régime :

À la suite d'une catastrophe naturelle, les dommages causés directement sur des biens matériels couverts dans le cadre d'un contrat d'assurance sont pris en charge. Prenant en charge les dommages aux bâtiments et au mobilier, la couverture Cat Nat agit également sur l'intégralité du processus de remise en état : du travail de déblayage et nettoyage, jusqu'aux frais de remise en état du bien comme les frais d'études géotechniques ou honoraires d'autres intervenants.

Financement

Ce n'est pas une prime forfaitaire mais un taux de prime intégré dans les contrats d'assurance qui permet le financement du régime des catastrophes naturelles. La loi du 13 juillet 1982 stipule que le taux de prime est "un taux unique défini par arrêté pour chaque catégorie de contrat. Ce taux est appliqué au montant de la prime ou cotisation principale ou au montant des capitaux assurés, selon la catégorie de contrat". Ce taux diffère selon trois catégories : biens et construction, vol et incendie automobile et dommage automobile.

Type de risque	Taux de prime
Construction	12 %
Biens particuliers	12 %
Biens professionnels	12 %
Vol incendie auto	6 %
Dommages auto	0.5 %

Source : CCR

L'application de franchises étant un élément important pour le financement du régime des catastrophes naturelles, l'État impose ces montants et interdit leur rachat.

<i>Franchise minimale (en €)</i>	Non Sécheresse	Sécheresse
Damage non professionnel	380	1520
Damage professionnel	min (10 %, 1140)	3050
Pertes d'exploitation	min (3j ouvrés, 1140)	3050

Source : CCR

Pour le risque de subsidence, la franchise est quatre fois plus importante que la franchise de base pour les biens à usage non professionnel. Le montant des franchises peut également être revu à la hausse si les communes sont touchées à plusieurs reprises sans qu'elles aient mis en place un Plan de Prévention des Risques Naturels.

1.4 Évolution du péril en France

Le graphique ci-dessous confirme l'importance de la sécheresse dans le régime des catastrophes naturelles. En étudiant le détail des périls pour chaque année, nous pouvons constater que le péril sécheresse est fréquemment plus important que n'importe quel autre péril pris en charge par le régime des catastrophes naturelles. Ce fut notamment le cas en 2011, où 71 % de la sinistralité était dû à la sécheresse. Nous pouvons également remarquer qu'à partir de 2016, le péril sécheresse est presque systématiquement plus important que le risque inondation.

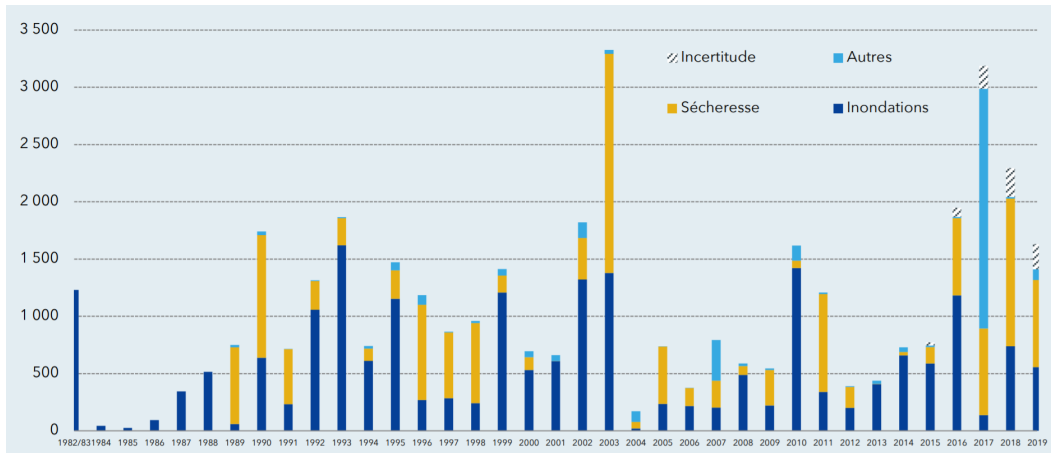


FIGURE 1.3 – CCR : Sinistralité Cat Nat non-auto de 1982 à 2019 (M€)

La sinistralité causée par la sécheresse en 2003 au titre du régime Cat Nat est comparable à celle causée par l’ouragan Irma survenu en 2017. Si ces deux phénomènes venaient à survenir durant la même année, il serait envisageable d’atteindre une sinistralité de 5 milliards d’euros.

1.5 Disparités spatiales

La cartographie de la sinistralité historique à échelle communale de 1995 à 2016 liée à la sécheresse met en évidence des disparités certaines. En effet, la sinistralité semble être concentrée sur certaines régions du sud-ouest et du centre de la France et ceci de façon très disparate.

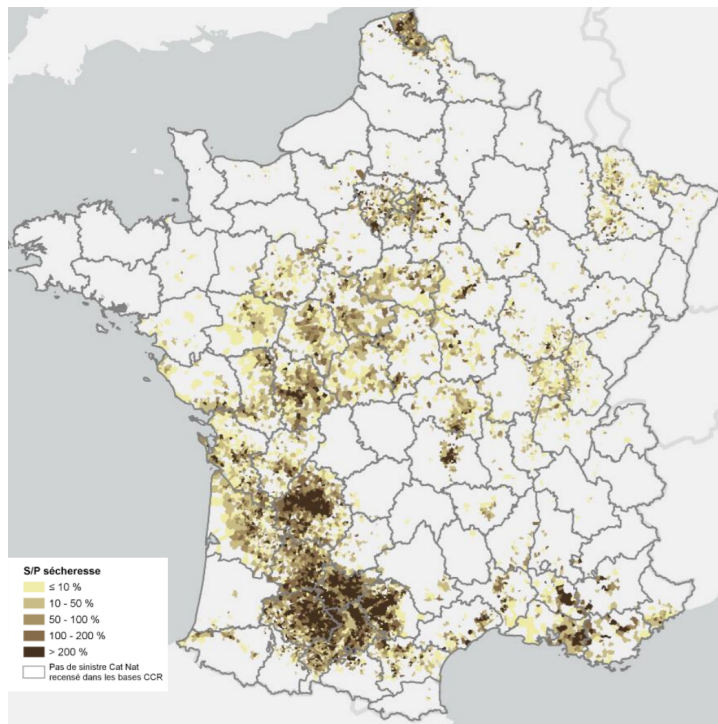


FIGURE 1.4 – CCR : S/P sécheresse en France par commune de 1995 à 2016

L'information fournie par cette cartographie n'est cependant que partielle et nous pourrions croire à tort que le péril sécheresse épargne totalement certaines régions du nord-ouest. La raison à cela est que la sinistralité est en majorité dû à l'effet de sécheresse-réhydratation des sols. C'est pour quoi la sinistralité est essentiellement visible dans les régions possédant de nombreuses terres argileuses [Voir carte en annexe].

L'effet de sécheresse-réhydratation des sols n'étant pas la seule conséquence de la sécheresse, il est important de garder à l'esprit que ce phénomène climatique impacte, de diverses manières, l'ensemble du territoire français.

La cartographie ci-après à été réalisée par la Caisse Centrale de Réassurance. Elle nous permet de constater que la sécheresse a davantage touché le nord-est que le sud-ouest de la France durant l'été 2019. Nous pouvons également constater que l'hydratation des sols de la région sud-ouest a été importante durant l'automne 2019. Ceci illustre en partie le phénomène de sécheresse-réhydratation des sols, qui est présent dans cette région et engendre la forte sinistralité constatée précédemment.

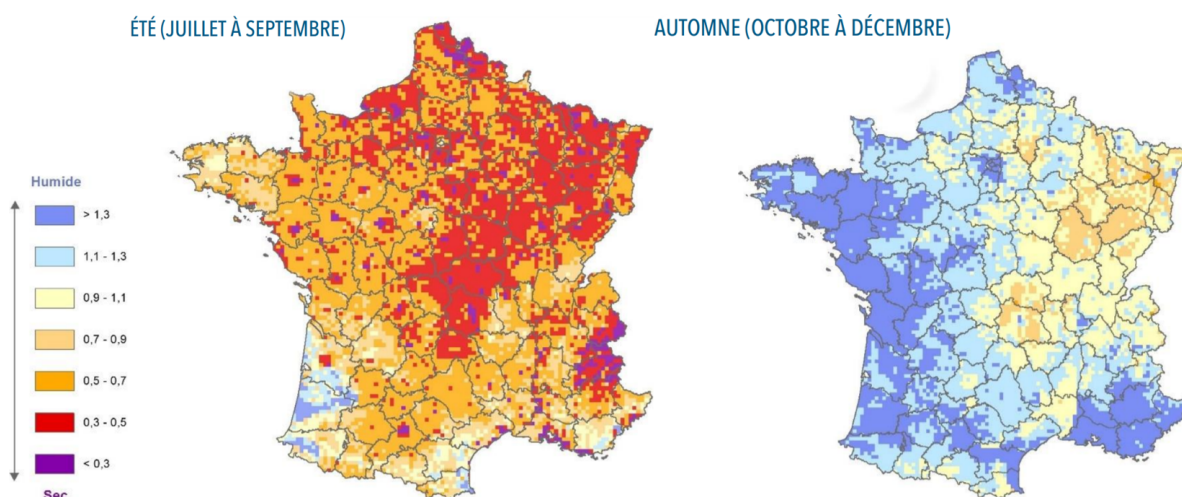


FIGURE 1.5 – CCR : sécheresse et humidité mesurée en 2019

Maintenant que nous avons vu que le risque sécheresse présente un risque conséquent en France, nous allons nous intéresser au transfert de ce risque entre un acteur météo-sensible et un investisseur. Ce transfert de risque peut notamment être réalisé par l'échange d'un produit financier appelé dérivé climatique. Le chapitre suivant sera destiné à la compréhension de son fonctionnement et de l'état du marché actuel.

Chapitre 2

Présentation des dérivés climatiques

2.1 Définition

Définition et Fonctionnement

Comme défini par Wx Risk Global, un dérivé climatique est un contrat qui permet à son détenteur de réduire un risque lié à un aléa climatique auquel il s'estime sensible [WxRG].

Tout d'abord, l'acheteur décrit l'indemnisation souhaitée comme fonction d'un indice climatique. Un indice climatique est une variable basée sur des relevés climatiques tels que la pluie ou la température, qui permettra de refléter de manière appropriée un risque particulier lié au climat. Plus tard, si l'indice climatique franchit le seuil à partir duquel l'acheteur s'est déclaré sensible, alors ce dernier recevra une compensation financière, souvent proportionnelle au dépassement du seuil fait par l'indice.

À travers cette transaction, c'est le vendeur du dérivé climatique qui sera porteur du risque lié à la valeur prise par l'indice et contre lequel l'autre partie souhaite se protéger. Lors de la vente du contrat, il touchera une prime proportionnelle à cet engagement.

Un indice doit vérifier certains critères pour jouer le rôle de sous-jacent d'un dérivé climatique :

- Son calcul doit être objectivement défini à partir de l'ensemble des variables climatiques sur lequel il est construit.
- Il doit être calculable dans le temps de manière fiable et transparente afin qu'aucun parti ne puisse remettre en cause sa valeur sur l'ensemble de la durée de vie du contrat.
- Aucun parti ne doit pouvoir influencer la valeur prise par l'indice.
- Le lieu où chacune des variables climatiques nécessaires au calcul de l'indice sont relevées doit être clairement identifié. Cela peut être, par exemple, une station météorologique ou des coordonnées d'un relevé satellitaire.

Les dérivés climatiques peuvent être basés sur n'importe quel indice sous-jacent dépendant de variables climatiques à condition que l'acheteur et le vendeur soient en accord mutuel sur sa définition et son utilisation.

Les indices les plus intuitifs peuvent être l'accumulation de pluie sur une période donnée ou le nombre de jours aux températures supérieures à un certain seuil. Les dérivés climatiques payent le plus souvent de manière linéaire à partir d'un seuil de déclenchement fixe et en définissant le prix d'une unité considérée. Certains contrats définissent également une limite de paiement.

Exemple illustratif

Un vendeur de glaces craint une diminution des ventes en raison d'un été anormalement froid. Il peut donc décider d'acheter un dérivé climatique pour compenser les mauvaises conditions climatiques de l'été à venir. Ainsi, si l'été suivant est anormalement froid, la perte financière engendrée sera en partie compensée par le dérivé climatique et si l'été suivant est chaud, il aura payé une prime du dérivé climatique qui sera compensée par un bon chiffre d'affaire.

Par exemple, les spécificités du contrat souscrit par le glacier pourraient être les suivantes :

- Indice :

Nombre de jours où la température maximale n'a pas dépassé 20°C au mois d'août.

- Seuil de déclenchement :

5 jours

- Valeur d'une unité d'indice :

300 €

- Plafond de paiement :

6 000 €

Le graphique ci-dessous représente le montant que payera le dérivé climatique en fonction du nombre de jours où la température maximale n'a pas dépassé le seuil de 20 ° C :

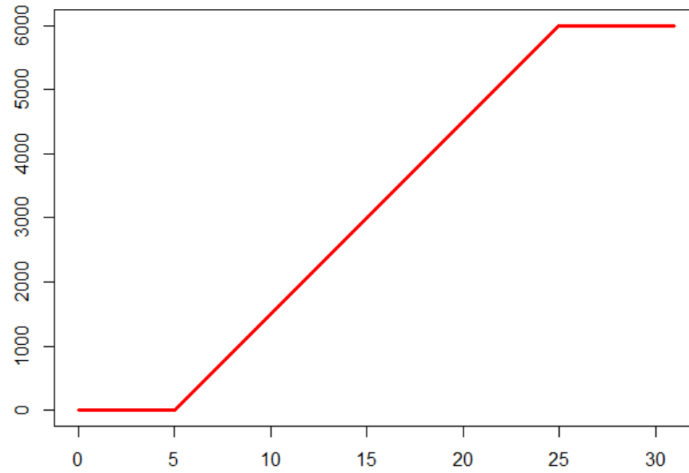


FIGURE 2.1 – Paiement en fonction du nombre de jours passés en dessous du seuil

2.2 Indices sous-jacents fréquemment utilisés

L'indice sous-jacent le plus souvent utilisé dans les contrats dérivés climatiques est l'indice Heating Degree Day et le Cooling Degree Day. Ces indices reflètent l'écart cumulé sur n jours entre la température moyenne T_i de chaque jour et une température T_{ref} de référence.

Un indice HDD sur une période de n jours est calculé avec la formule suivante :

$$HDD_n = \sum_{i=1}^n \max(0; T_{ref} - T_i)$$

Nous remarquons que plus la température est basse durant la période observée, plus l'indice HDD est élevé. Par exemple, une entreprise pourra, grâce à un contrat basé sur cet indice, bénéficier d'un paiement qui compensera la dépense en électricité nécessaire pour maintenir ses locaux à la bonne température. Dans le cas où la température est toujours supérieure à la température de référence, il n'a pas besoin de se chauffer et cette option ne lui rapporte rien.

La température de référence est fixée à 18.3°C pour les dérivés proposés par le Chicago Mercantile Exchange car c'est généralement à partir de ce seuil que les particuliers allument leur chauffage. Axa Re définit, quant à lui, plusieurs températures de références permettant d'offrir des produits adaptés aux compagnies vendant des boissons et dont la consommation baisse à partir d'une température donnée.

Sur le même modèle, l'indice CDD est calculé de la manière suivante :

$$CDD_n = \sum_{i=1}^n \max(0; T_i - T_{ref})$$

Ce sont au contraire les acteurs sensibles à une hausse des températures qui sont intéressés par l'achat d'options basées sur l'indice CDD. D'une manière similaire à celle exposée pour l'indice HDD, une compagnie qui vend de la crème glacée pourrait ainsi compenser une perte de son chiffre d'affaires causée par un été trop frais.

Exemple d'un indice HDD calculé sur une semaine en prenant une température de référence de 15 ° C :

	Jour 1	Jour 2	Jour 3	Jour 4	Jour 5
Température moyenne mesurée	12	16	19	15	13
Écart avec la température de référence	-3	1	4	0	-2

La valeur de l'indice HDD est ici de $0 + 1 + 4 + 0 + 0 = 5$

De tels indices peuvent notamment être intégrés dans des contrats à terme. Après avoir défini la valeur P_{unit} d'une unité d'indice, ce contrat payera à son terme le montant défini par :

Paiement à terme = P_{unit} x Valeur de l'indice HDD

2.3 Marché des dérivés climatiques

2.3.1 Premiers échanges

Le premier contrat dérivé climatique a été échangé en 1997 [CONSIDINE]. Ce premier contrat portait sur la température qu'il ferait durant l'hiver suivant. À la suite de cette initiative, le Chicago Mercantile Exchange prend conscience de l'intérêt pouvant être porté à ce type de produits financiers et introduit rapidement des options climatiques sur la marché. Le marché européen a mis plus de temps à proposer une offre. En 2001, Euronext est le premier à proposer une gamme de produits dérivés à partir de données climatiques fournies par Météo France, dont les bases de données sont considérées parmi les plus fiables d'Europe. L'année suivante, la société Britanque London International Financial Futures Exchange proposera à son tour des options basées sur la température mesurée dans la capitale de l'Angleterre, de la France et de l'Allemagne.

Nous allons maintenant détailler les différents marchés dans lesquels les dérivés climatiques peuvent être échangés et décrire comment le marché a évolué jusqu'à nos jours.

2.3.2 Un marché organisé historique : Le Chicago Mercantile Exchange

Le marché organisé est principalement animé par le Chicago Mercantile Exchange. Après avoir connu une croissance importante en 2005 et 2008, l'offre de dérivés climatiques sur le marché à terme du Chicago Mercantile Exchange semble stagner depuis quelques années. Aujourd'hui ce marché propose exclusivement des dérivés basés sur la température et uniquement disponibles dans 12 grandes villes. Le schéma ci-dessous permet de constater que le marché est concentré aux États-Unis et qu'il n'y a pas d'offre en Europe en dehors de Londres et d'Amsterdam.

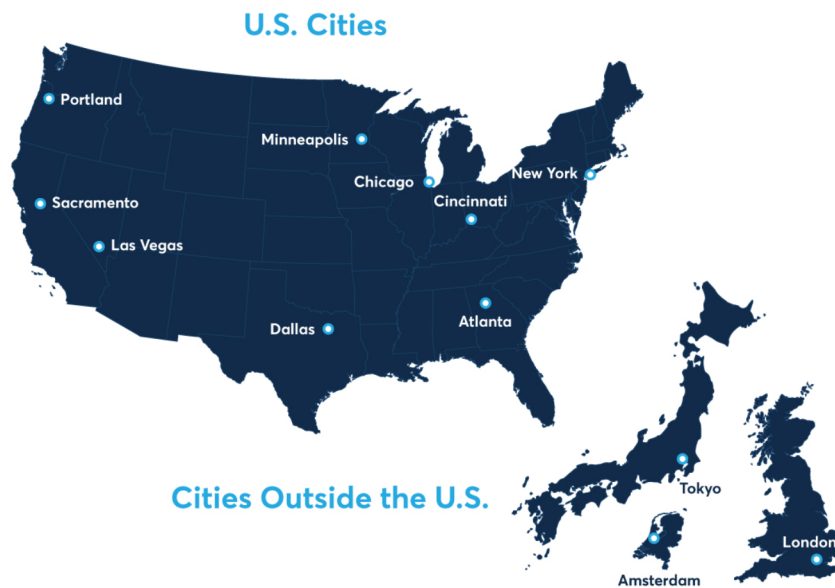


FIGURE 2.2 – CME : Localisation de l'offre des dérivés climatiques

Malgré que le Chicago Mercantile Exchange soit un acteur historique, il semble que la majorité des contrats sont aujourd'hui échangés sur le marché de gré à gré. Cependant, certains acteurs météo-sensibles privilégient toujours le marché organisé car la présence d'une chambre de compensation permet de se prémunir contre le risque de contrepartie.

2.3.3 Prédominance du marché de gré à gré

Il est difficile de recueillir des informations précises sur le volume d'échange de dérivés climatiques, leur composition et leur évolution au cours de ces dernières années. Ces données sont souvent inaccessibles ou très onéreuses. Il a donc fallu s'appuyer sur l'avis de plusieurs experts pour comprendre au mieux l'état actuel du marché.

Selon Bernard Finas, fondateur de Jola et directeur de The Weather Option, le marché est aujourd'hui essentiellement de gré à gré et les principaux intervenants sont Swiss Re, Munich Re et Nephila. Le marché de gré à gré a été en croissance continue et l'utilisation de ce type de couverture contre les intempéries climatiques s'est accrue d'année en année. Aujourd'hui, il estime de la taille du marché en termes de risques échangés c'est-à-dire "la somme sur toutes les transactions des paiements maximum possible de la part de la partie identifié comme preneur de risque à hauteur de 30 milliards de dollars dont 60 % dans la catégorie NAT CAT, 30 % dans la catégorie AGRICULTURE et 10 % dans la catégorie ENERGY".

Selon David WHITEHEAD, co-directeur de weatherXchange, "l'activité du marché est aujourd'hui plus élevée qu'elle ne l'a jamais été au cours de cette dernière décennie". Ce constat est d'ailleurs cohérent avec l'analyse que Bernard FINAS a pu faire du marché en remarquant que "face au défi considérable que représente le changement climatique pour l'humanité les pays du G7 s'étaient engagés lors de la COP 21 à Paris en 2015, à mobiliser 100 milliards de dollars par an à l'horizon 2020 dans le cadre du programme InsuResilience".

Deuxième partie

Théorie

Chapitre 3

Stratégies de couverture par des produits dérivés climatiques

Un acteur météo sensible peut se couvrir contre son risque grâce à un dérivé climatique. Nous allons exposer dans ce chapitre différentes stratégies de couvertures communément échangées sur les marchés financiers et expliquer les avantages qu'elles présentent en terme de couverture.

3.1 Options simples

Commençons tout d'abord par présenter les dérivés climatiques les plus simples : le call et le put qui sont respectivement des options d'achat et de vente. Le sous-jacent étant un indice climatique, il n'est bien entendu pas question d'acheter ou de vendre cet indice, mais plutôt de recevoir un paiement dont le montant est défini par les valeurs prises par cet indice le sous-jacent. Nous appellerons à partir de maintenant payoff le paiement offert par un tel contrat.

Notons :

- T la date d'échéance du contrat
- K le prix d'exercice défini dans le contrat
- S_t la valeur du sous-jacent ou indice climatique à l'instant $t \in [0, T]$
- P_{call} le payoff d'une option d'achat lors de son exercice
- P_{put} le payoff d'une option de vente lors de son exercice

3.1.1 Option d'achat

Au moment de l'exercice, l'option d'achat offre à son détenteur un montant proportionnel au dépassement de la valeur du sous-jacent par rapport au seuil d'exercice préalablement défini dans le contrat. Ainsi, l'expression du payoff d'une option d'achat est la suivante :

$$P_{call} = \max(S_T - K, 0)$$

À ce gain il faut soustraire la prime P qui a été payée pour acquérir l'option. Le gain de l'acheteur de l'option devient donc :

$$G_{call} = \max(S_T - K, 0) - P$$

La détention d'une option d'achat est utile à une entreprise qui s'estime sensible à une variation à la hausse de l'indice climatique sous-jacent au dessus d'un certain seuil. Par exemple, un camping pourrait souscrire un contrat de ce type basé sur un indice de pluie pour compenser un manque de touristes durant un été pluvieux.

3.1.2 Option de vente

L'option de vente a un fonctionnement similaire à celui de l'option d'achat mais le montant qu'elle offre est cette fois proportionnel à la baisse de la valeur du sous-jacent en dessous du seuil d'exercice préalablement défini dans le contrat. Ainsi, l'expression du payoff d'une option de vente est la suivante :

$$P_{put} = \max(K - S_T, 0)$$

En soustrayant la prime P qui a été payée pour acquérir l'option, le gain de l'acheteur de cette option devient :

$$G_{put} = \max(K - S_T, 0) - P$$

La détention d'options de vente est utile à une entreprise qui s'estime sensible à une variation à la baisse de l'indice climatique sous-jacent au dessus d'un certain seuil. Par exemple, un brasseur de bière pourrait souscrire un contrat de ce type basé sur un indice de température pour compenser une faible vente de boissons lors d'un été frais.

Nous remarquons que pour ces deux options, le risque porté par le détenteur est limité par la prime qu'il aura payée pour acheter cette option. Les gains, quant à eux, ne sont en théorie pas limités. Il est cependant improbable que l'indice atteigne des valeurs très importantes sans que la prime, qui est sensée refléter le risque, devienne excessivement élevée. Notons que les options d'achat et de vente sont souvent appelées respectivement call et put.

3.1.3 Différents types d'exercices pour une option

Il existe sur le marché différents types d'options. Les options Européennes ne payent qu'à leur échéance tandis que les options Américaines peuvent être exercées à tout instant entre la souscription et l'échéance. Il existe de nombreux types d'options ayant chacune leurs particularités mais seuls trois d'entre eux, parmi les plus classiques, seront envisagés.

En notant T l'échéance de l'option, le tableau suivant définit l'ensemble J des dates d'exercices possibles pour l'option Européenne, Américaine et Bermudienne :

Type d'option	Possibilité d'exercer
Européenne	$J = T$
Américaine	$J \in [0; T]$
Bermudienne	$J \in \{j_1 < j_2 < \dots < j_n \leq T\}$

Nous pouvons tout de suite remarquer que l'option Américaine offre plus de droits à son détenteur que l'option Européenne car l'exercice peut être fait à n'importe quel instant. Par conséquent, toute chose égale par ailleurs, une option Américaine aura plus de valeur qu'une option Européenne. L'option Bermudienne se situe quant à elle

entre les deux, offrant plus de possibilité d'exercice qu'une option Européenne mais ne dépassant jamais la liberté d'exercice d'une option Américaine. Pour la suite des travaux, lors de la structuration de notre dérivé climatique sécheresse, ce sont ces trois types d'options qui seront envisagées.

3.2 Produits combinés

3.2.1 Écart vertical

Les écarts verticaux ressemblent aux options classiques mais permettent de plafonner le paiement en cas d'écart trop important entre l'indice mesuré à l'exercice et le seuil d'exercice. Ceci permet au porteur de risque de limiter sa perte et à l'acheteur de payer une prime plus faible en ayant des garanties très proches de celles offertes par une option simple.

Écart haussier

Un écart vertical haussier fonctionne comme une option d'achat de seuil d'exercice $K1$, mais son payoff est plafonné dès que le sous-jacent atteint un seuil $K2$, avec $K1 < K2$. Il peut être construit à partir de l'achat d'un call de seuil d'exercice $K1$ combiné à la vente d'un autre call de seuil d'exercice $K2$. L'expression du payoff d'un écart haussier est la suivante :

$$P_{eq.haussier} = \min \left(\max(S_T - K1, 0), K2 - K1 \right)$$

À ce gain il faut soustraire la prime $P1$ qui a été payée pour acquérir le call de seuil $K1$ et additionner $P2$ qui correspond à la vente du call de seuil $K2$. Le gain de l'acheteur de l'option devient donc :

$$G_{eq.haussier} = \min \left(\max(S_T - K1, 0), K2 - K1 \right) - (P1 - P2)$$

Écart baissier

Un écart vertical baissier fonctionne comme une option de vente de seuil d'exercice $K2$, mais son payoff est plafonné dès que le sous-jacent atteint un seuil $K1$, avec $K1 < K2$. Il peut être construit à partir de l'achat d'un put de seuil d'exercice $K2$ combiné à la vente d'un autre put de seuil d'exercice $K1$. L'expression du payoff d'un écart baissier est la suivante :

$$P_{eq.baissier} = \min \left(\max(K2 - S_T, 0), K2 - K1 \right)$$

A ce gain il faut soustraire la prime $P2$ qui a été payée pour acquérir le put de seuil $K2$ et additionner $P1$ qui correspond à la vente du put de seuil $K1$. Le gain de l'acheteur de l'option devient donc :

$$G_{eq.baissier} = \min \left(\max(K2 - S_T, 0), K2 - K1 \right) - (P2 - P1)$$

Les dérivés que nous venons d'analyser permettent également la construction de nombreux autres dérivés plus complexes [GUILLERME, 2009]. Par la suite, nous ne nous intéresseront plus qu'à la tarification d'options simples telles que le call et le put. Nous allons maintenant voir comment le prix de telles options peut être défini.

Chapitre 4

Évaluation d'une option climatique

Un dérivé climatique provoquant une transaction financière à son exercice, il est important pour le vendeur d'estimer au plus juste la prime que l'acheteur devra payer pour obtenir ce produit financier. Une prime trop basse mettrait en danger le vendeur qui serait alors perdant sur les contrats vendus et une prime trop élevée pourrait dissuader l'autre parti d'acheter le dérivé dont il a besoin.

L'une des particularités d'un contrat dérivé climatique est que son sous-jacent est une fonction de variables climatiques qui, par nature, ne peuvent pas être échangées sur le marché. Le modèle de Black-Scholes ne peut donc pas être utilisé pour évaluer le prix d'une option climatique. Pour donner un prix à une option climatique, il est alors nécessaire d'estimer au mieux le paiement attendu par celui-ci à partir de l'historique de données climatiques dont nous disposons. Nous exposons dans ce chapitre différentes méthodes permettant de tarifer une option climatique.

Étant donné que des produits dérivés climatiques sont principalement achetés pour se couvrir contre un risque, ces produits financiers ont un fonctionnement très proche de celui d'un contrat d'assurance. Pour cette raison, nous appellerons "prime" le montant payé par l'acheteur d'un dérivé climatique. Nous utiliserons le terme *payoff* pour définir le montant versé à l'exercice par un produit dérivé.

Rappelons que le type de contrat dérivé climatiques que l'on souhaitera tarifer par la suite est un contrat dont le sous-jacent est un indice climatique qui est lui-même calculé à partir de variables climatiques. L'évaluation d'un tel contrat peut donc se faire à trois niveaux : par l'estimation directe du *payoff* du contrat, par la modélisation de l'indice climatique sous-jacent ou par la modélisation conjointe des variables climatiques servant à calculer l'indice climatique.

4.1 Méthode Burn

Cette première méthode consiste à estimer directement le *payoff* d'un contrat dérivé par sa moyenne empirique. L'application de cette méthode se fait par les étapes suivantes :

- Simuler l'indice sous-jacent dans le passé :

Avant de pouvoir estimer la performance du contrat, il faut calculer les valeurs historiques du sous-jacent. Comme notre estimation repose sur le calcul d'une moyenne empirique, il est important de disposer d'une profondeur d'historique suffisante pour avoir une tarification robuste.

- Simuler le payoff du contrat :

Étant donné que nous travaillons avec des données climatiques saisonnières, nous calculons le payoff du contrat sur chaque année passée, à la date d'exercice du contrat.

- Calcul de la moyenne empirique :

Supposons que l'on dispose d'une profondeur d'historique de n années et notons X_i le payoff simulé il y a i années. Nous pouvons approcher l'espérance du payoff futur par la moyenne empirique des n paiements simulés dans le passé :

$$M = \frac{1}{n} \sum_{i=1}^n X_i$$

- Calcul de la prime :

Pour prendre en compte l'aversion au risque du vendeur, nous pouvons aussi calculer l'écart type des paiements historiques puis définir λ , un réel positif correspondant au poids affecté par le vendeur à l'écart type du payoff du contrat. Nous pouvons alors définir la prime P du contrat :

$$P = M + \lambda * \sigma$$

Avec :

$$\sigma = \frac{1}{n} \sqrt{\sum_{i=1}^n (X_i - E)^2}$$

Limites de validité

Cette méthode permet un calcul rapide de la prime du contrat mais elle comporte certaines limites. Tout d'abord, le fait d'estimer le payoff futur par sa moyenne empirique implique que la tendance du sous-jacent doit être négligeable. Aussi, l'une des principales limites de cette méthode est qu'elle attribue une prime pure nulle à toute option qui n'aurait pas pu être exercée sur la période historique étudiée. Or cela ne signifie en aucun cas que l'option ne sera pas exercée à l'avenir.

4.2 Modélisation actuarielle du sous-jacent

Nous allons maintenant voir une seconde méthode d'évaluation de la valeur d'une option climatique en utilisant une modélisation stochastique de son indice sous-jacent.

Après avoir simulé les valeurs historiques calculées de l'indice, il faut choisir une fonction de répartition qui semble à priori compatible avec la distribution des observations. Il est préférable que la complexité du modèle choisi soit proportionnelle au nombre d'observations dont nous disposons pour éviter une erreur de modèle. Une loi peut également être choisie si la littérature a démontré que l'indice étudié peut être régi par celle-ci. Une fois le modèle choisi, il faut tester si la loi supposée correspond aux valeurs de l'indice observé avec un niveau de confiance suffisamment élevé.

Posons X une variable aléatoire correspondant à la valeur prise par l'indice climatique sous-jacent à la date d'échéance. Soit f la fonction de payoff du contrat. L'objectif est d'estimer son espérance de $\mathbb{E}(f)$.

En supposant que g corresponde à la densité de probabilité de X , nous pourrions estimer l'espérance de f en appliquant la formule de transfert :

Théorème de transfert

Soit $X : (\Omega, \mathcal{F}, P) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire de densité g .

Soit $f : (\mathcal{B}(\mathbb{R})) \rightarrow (\mathcal{B}(\mathbb{R}))$ une fonction mesurable.

Si f est positive sur \mathbb{R} , alors :

$$\mathbb{E}[f] = \mathbb{E}_X[f] = \int_{\mathbb{R}} f dP_X$$

Nous pouvons ainsi estimer l'espérance du payoff à partir de f et de g :

$$E[f] = \int_{\mathbb{R}} f(x)g(x) dx$$

Cas d'une densité normale

La loi normale est souvent utilisée pour modéliser la distribution d'un sous-jacent. Nous allons donc faire quelques rappels qui seront nécessaires à l'application de cette méthode avec une loi de distribution normale.

Rappelons qu'une loi normale de paramètres μ et $\sigma > 0$ est une loi de densité g telle que :

$$\forall x \in \mathbb{R}, \quad g(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Si X possède une telle densité, nous retrouvons facilement sa moyenne et sa variance à partir des paramètres μ et σ :

$$\mathbb{E}[X] = \mu$$

et

$$\text{Var}[X] = \sigma^2$$

Avant de pouvoir supposer que la distribution de l'indice sous-jacent est normale, il est important d'effectuer un test de normalité sur les observations dont nous disposons. Pour cela, nous pouvons utiliser le test de Shapiro-Wilk.

Limite de validité

Cette méthode étant développée à partir de la loi supposée du sous-jacent, elle est sensible à l'erreur de modélisation. Il est donc important de bien justifier le choix de la fonction de répartition. Dans le cas où nous ne parvenons pas à trouver une loi cohérente avec la distribution du sous-jacent, il est possible de construire une distribution de l'indice à partir d'une méthode non paramétrique appelée estimation par le noyau. La description de cette méthode est disponible en Annexe.

4.3 Générateur climatique stochastique

La troisième méthode considérée pour la tarification d'un dérivé climatique est la simulation par un générateur climatique stochastique. Nous appelons ici générateur climatique un algorithme qui permet la génération des variables climatiques nécessaires au calcul de notre indice sous-jacent. Déjà, cette approche a notamment été implémentée pour la tarification de dérivés climatiques basées sur les indices de température HDD et CDD [MORENO et ROUSTANT, 2002], [DUTANG et al., 2017] ainsi que pour la "tarification d'une assurance indicielle pour des producteurs de maïs au Mali" [CÔME, 2018]. Leurs résultats montrent que cette approche semble être la plus efficace à ce jour pour tarifier de tels contrats dérivés.

Une fois le générateur construit, il sera possible de calculer la valeur prise par notre indice climatique sous-jacent sur un horizon futur. Ensuite, le calcul de la prime pure d'un contrat basé sur cet indice pourra être tarifé par un calcul similaire à celui fait dans la méthode Burn. La différence essentielle est que nous calculerons le payoff moyen sur un très grand nombre d'années générées et non uniquement sur les observations historiques dont nous disposons dans la base de données.

Un autre avantage de la mise en place d'un générateur climatique stochastique est qu'il permet de prendre en compte les corrélations spatiales des variables utilisées. Cela permet une meilleure gestion du risque. En effet, il est dangereux pour un assureur d'avoir un portefeuille si les dérivés qui le composent sont très corrélés. La prise en compte de la corrélation partielle permet donc à l'assureur de contrôler que le risque issu de son portefeuille est suffisamment diversifié.

Par la suite, nous verrons que l'indice sécheresse que nous retiendrons est basé sur des variables de pluies et de températures. Nous mettrons donc en place un générateur simulant ces deux variables en tenant compte de leur interaction.

Troisième partie
Cadre d'étude

Chapitre 5

Présentation de la base de données

L'objectif est de présenter la base de données sélectionnée, expliquer ses sources, ses propriétés et son extraction sous R. Il s'agit aussi de préciser les variables utiles à notre tarification et leur méthodologie de mesure. Les avantages et les limites de cette base seront aussi énoncés ainsi qu'une ouverture pour aller plus loin dans l'exploitation de la base. Pour la tarification des contrats d'assurance, nous utilisons deux variables qui sont la température maximale et les précipitations journalières.

5.1 Présentation générale

5.1.1 Source

Nous utilisons la base de données issue du projet "Prediction Of Worldwide Energy Resources" (POWER) de la NASA. Il s'agit d'un projet orchestré depuis 2003 par des chercheurs de la NASA ayant pour objectif la construction d'une base de données répertoriant de nombreux indicateurs climatiques à échelle planétaire. La collecte de données se fait à travers un système de satellites. Les données proviennent des modèles d'assimilation GMAO MERRA-2 et GEOS 5.12.4 FP-IT de la NASA [NASA MERRA, 2017], [NASA GEOS, 2016].

D'autres sources de données tel que météo France auraient également pu servir de base à ce mémoire. Cependant, l'accessibilité, l'homogénéité, la pérennité et la gratuité des données de la NASA sont les éléments qui nous ont motivés à travailler sur leur API.

5.1.2 Propriétés

La résolution géographique

Les données brutes sont collectées avec résolution de grille latitude \times longitude de $(\frac{1}{2})^\circ \times (\frac{2}{3})^\circ$. Ensuite, une interpolation bilinéaire est réalisée pour obtenir un maillage géographique de résolution $(\frac{1}{2})^\circ \times (\frac{1}{2})^\circ$ [SHEA, 2014].

Pour donner un ordre de grandeur, en France, cette résolution correspond à un maillage rectangulaire de $40km \times 74km$ interpolé pour créer un maillage de $40km \times 55km$.

Le graphique ci-dessous représente la grille sur laquelle sont effectués les relevés satellitaires pour la base POWER de la NASA :

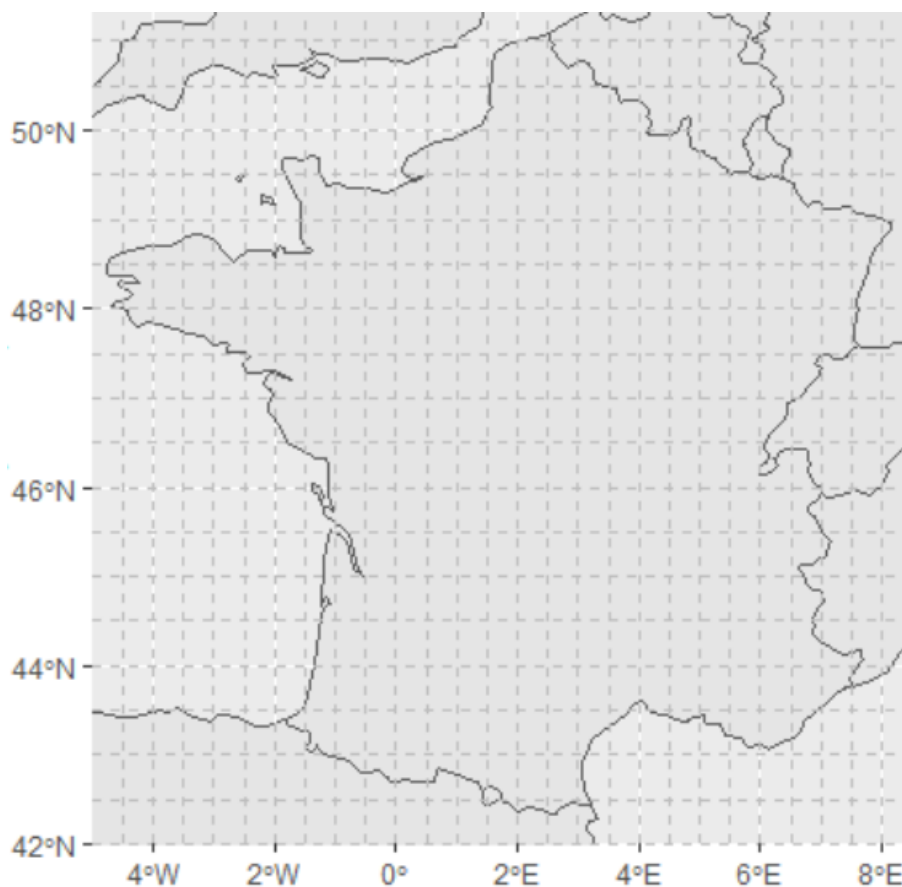


FIGURE 5.1 – Grille GPS de résolution $(\frac{1}{2})^\circ \times (\frac{1}{2})^\circ$

La résolution temporelle

Les données ont été initialement produites sur un pas de temps d'une heure et moyennées pour restituer des valeurs quotidiennes.

Les données météorologiques MERRA-2 disponibles démarrent au 1er janvier 1981.

La publication des données se fait avec latence d'environ 2-3 jours.

5.1.3 Implémentation sous R

Cette base de données s'accompagne de packages R et Python, développés aussi par la NASA, permettant d'exploiter leur base de données. Dans le cadre de ce mémoire, nous utiliserons le package R. Celui-ci permet d'extraire simplement les données souhaitées en entrant les paramètres suivants [SPARKS, 2020] :

- Coordonnées GPS
- Date
- Variables

Ci-dessous nous voyons le tableau créé par fonction "get_power" du package "na-sapower" pour l'import de données journalières de précipitations et de température moyenne à 2 mètres du sol entre 2015 et 2019 :

```
> data
NASA/POWER SRB/FLASHFlux/MERRA2/GEOS 5.12.4 (FP-IT) 0.5 x 0.5 Degree Daily Averaged Data
Dates (month/day/year): 01/01/2015 through 01/01/2019
Location: Latitude -27.48 Longitude 151.81
Elevation from MERRA-2: Average for 1/2x1/2 degree lat/lon region = 434.55 meters Site = na
Climate zone: na (reference Briggs et al: http://www.energycodes.gov)
Value for missing model data cannot be computed or out of model availability range: NA

Parameters:
T2M MERRA2 1/2x1/2 Temperature at 2 Meters (C) ;
PRECTOT MERRA2 1/2x1/2 Precipitation (mm day-1)

# A tibble: 1,462 x 9
  LON  LAT  YEAR  MM  DD  DOY  YYYYMMDD  T2M  PRECTOT
  <dbl> <dbl> <dbl> <int> <int> <int> <date> <dbl> <dbl>
1 152. -27.5 2015 1 1 1 2015-01-01 26.4 1.68
2 152. -27.5 2015 1 2 2 2015-01-02 24.4 4.06
3 152. -27.5 2015 1 3 3 2015-01-03 23.7 5.59
4 152. -27.5 2015 1 4 4 2015-01-04 23.7 1.04
5 152. -27.5 2015 1 5 5 2015-01-05 22.4 3.09
6 152. -27.5 2015 1 6 6 2015-01-06 22.6 2.41
7 152. -27.5 2015 1 7 7 2015-01-07 22.4 0.79
8 152. -27.5 2015 1 8 8 2015-01-08 23.0 0.38
9 152. -27.5 2015 1 9 9 2015-01-09 24.4 0.14
10 152. -27.5 2015 1 10 10 2015-01-10 24.7 0.22
# ... with 1,452 more rows
```

FIGURE 5.2 – Exécution du code R - récupération de données

5.2 Précisions sur les variables utiles

5.2.1 Température maximale

Dans notre base de données, la variable T2M représente la température journalière maximale à deux mètres du sol au point de coordonnées choisi. Ce maximum est calculé à partir de relevés de températures effectué toutes les heures.

La mesure satellitaire est très satisfaisante. En effet, 95 % des variations issues de mesures en surface sont expliquées par les mesures satellitaires [STACKHOUSE, 2019], [WHITE et al., 2008].

5.2.2 Précipitation

Dans notre base de données, la variable PRECTOT représente le volume total des précipitations de la journée au point de coordonnées choisi.

Le graphique suivant nous montre que les relevés satellitaires ont parfois des difficultés à restituer les précipitations extrêmes. Cela ne posera pas problème car l'intensité des pluies extrêmes n'influera pas sur le calcul de notre indice de sécheresse :

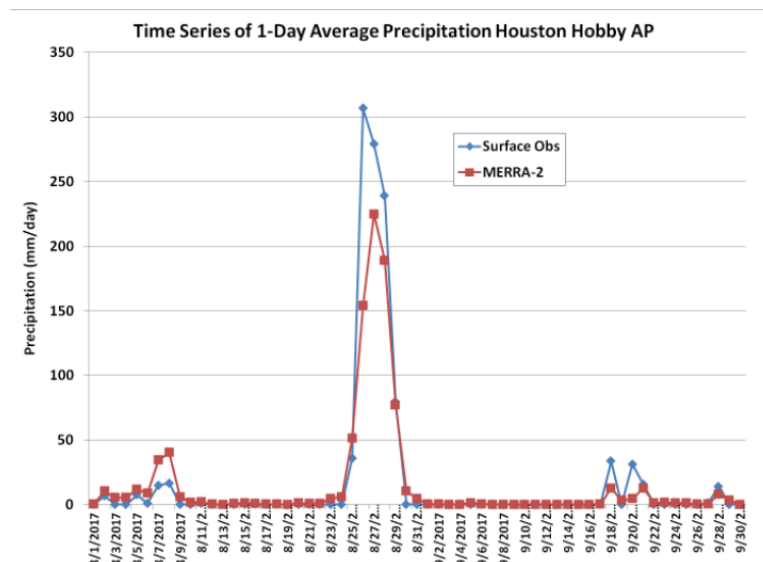


FIGURE 5.3 – Mesure des précipitations lors de l’ouragan Harvey

Des considérations plus physiques sur les méthodes utilisées sont disponibles en Annexe.

5.3 Avantages et limites de cette API

5.3.1 Avantages

- **Fiabilité méthodologique :**

Cette base est administrée par des chercheurs de la NASA. Des méthodes rigoureuses et documentées sont utilisées. En effet, la validation des paramètres météorologiques MERRA-2 (observations par satellite) est basée sur des comparaisons avec les observations prises directement sur le terrain des paramètres correspondants.

- **Accessibilité :**

Cette base de donnée étant liée à un package R, il est très simple et rapide d’obtenir les dernières données disponibles.

- **Dimension :**

Cette base de donnée comprend les données au niveau planétaire.

- **Pérennité :**

Le projet de la NASA est un projet à long terme qui a débuté en 1981. L’organisme est très fiable et son risque de cession d’activité est très faible.

5.3.2 Limites

- **Latence :**

Pour l'accessibilité des données climatologiques, il y a entre 2 et 3 jours de latence. Ce délai est relativement court mais pourrait être un frein au développement de certaines assurances paramétriques qui souhaiteraient indemniser dans les 24 heures suivant la survenance d'un sinistre. Nous pouvons penser en particulier à l'avènement des smart contrats dont la rapidité d'indemnisation est l'un des principaux arguments de vente.

- **Imprécisions de la grille :**

Le maillage $(\frac{1}{2})^\circ \times (\frac{1}{2})^\circ$ reste peu précis et limite la création de produits qui consisteraient à indemniser à un endroit précis et éloigné des points de grille.

5.4 Pour aller plus loin

La base de données du projet POWER ne fournit pas uniquement les deux variables que nous exploitons dans ce projet. Une multitude d'autres variables solaires et météorologiques sont aussi accessibles.

Voici une liste non-exhaustive :

- **Météorologique**

- Température moyenne (°C)
- Température minimale (°C)
- Température maximale (°C)
- Humidité relative (%)
- Point de rosée (°C)
- Précipitations (mm/day)
- Pression atmosphérique (mb)
- Vitesse du vent (m/s)

- **Solaire**

- Température de la terre (°C)
- Ensoleillement

Chapitre 6

Structuration du dérivé sécheresse

6.1 Choix de l'indice sécheresse sous-jacent

Maintenant que nous avons accès à la base de données climatiques, nous souhaitons choisir un indice qui permet de quantifier la sécheresse dans l'espace et le temps. Le choix de l'indice sera motivé par deux contraintes majeures : tout d'abord, il est impératif que l'indice retenu soit exclusivement calculé par les variables climatiques accessibles par l'API du programme POWER de la NASA. En effet, nous souhaitons évaluer des options basées sur cet indice sans stockage interne de données. L'automatisation complète du processus d'extraction des données sera donc une contrainte pour le choix de l'indice. La seconde contrainte est que l'indice devra être calculable sur l'ensemble du sol français. En effet, l'objectif est de pouvoir tarifier un dérivé qui soit à minima souscriptible en France. Bien entendu, si l'indice sécheresse s'avère calculable sur un territoire plus grand, l'intérêt porté envers celui-ci en sera d'autant accru.

La corrélation entre l'indice sécheresse et la mesure sur le terrain de sinistralité causée par la sécheresse ne fait pas partie des contraintes que nous avons fixées pour le choix de l'indice. Proposer une couverture cohérente avec la sinistralité ne fait pas partie des objectifs de ces travaux car ce lien peut être établi indépendamment de la tarification. Par le fruit de travaux futurs, l'indice que nous retiendrons pourrait être utilisé comme paramètre d'un produit d'assurance structuré. Nos travaux pourront alors permettre sa tarification.

En 2016, l'Organisation météorologique mondiale et le Partenariat mondial pour l'eau ont publié une analyse de l'ensemble des indices sécheresse créés à ce jour [OMM et GWP, 2016]. Avant de considérer certains d'entre eux, reprenons la définition d'indicateur de sécheresse et d'indice sécheresse établie dans ce rapport :

- **Indicateurs :**

"Variables ou paramètres qui servent à décrire les conditions de sécheresse. Citons, par exemple, les précipitations, la température, l'écoulement fluvial, le niveau des nappes et des réservoirs, l'humidité du sol et le manteau neigeux."

• **Indices :**

"Représentations numériques de l'intensité des sécheresses, que l'on calcule à partir de valeurs climatiques ou hydrométéorologiques, dont les indicateurs précités. Ils mesurent l'état qualitatif d'une sécheresse à un emplacement donné pour une période donnée."

Parmi l'ensemble des indicateurs exposés dans le manuel publié par l'Organisation météorologique mondiale et le Partenariat mondial pour l'eau, nous remarquons qu'une grande partie d'entre eux sont uniquement basés sur la quantité de précipitations. Cela est notamment le cas pour "l'indice de précipitations normalisé" et le "rapport à la normale des précipitations". Certains indices prennent en considération, en plus de la pluie, l'impact des températures sur la sécheresse. Parmi eux, nous avons considéré l'indice de sécheresse de Keetch-Byram noté KBDI. En 1968, Keetch et Byram publient cet indice permettant d'estimer un risque de feu de forêt aux États-Unis [KEETCH et BYRAM, 1968]. Cet indice reflète la quantité de pluie nécessaire pour saturer le sol et éliminer totalement le stress hydrique, c'est-à-dire saturer le sol en eau. Il est donc suffisamment générique pour être utilisé dans d'autres domaines tels que l'agriculture.

En plus de respecter les critères de choix d'indice énoncés précédemment, le KBDI présente l'avantage d'être normalisé par le climat historique à partir duquel il est calculé. Cela le rend, à priori, valable aussi bien aux États-Unis, qu'en France ou dans d'autres pays [DOLLING et al., 2009].

6.1.1 Calcul du KBDI

Après avoir corrigé la formule de calcul du KBDI dans laquelle une faute de frappe avait été constatée, une réécriture de la formule originelle a été publiée [ALEXANDER, 1990]. Ainsi, Le KBDI se calcule comme la somme du KBDI précédant, d'un facteur de sécheresse FS et d'un facteur de pluie FP :

$$KBDI_t = KBDI_{t-1} + FS_t - FP_t$$

$$FS_t = \frac{(203 - KBDI_{t-1}) \left(0,968e^{(0,0875 \times T_t + 1,552)} - 8,3 \right) \times 10^{-3}}{1 + 10,88e^{(-0,001736 \times \bar{P})}}$$

$$FP_t = \begin{cases} (P_t - 5,1) & \text{si } P_t \geq 5,1 \\ P_t & \text{si } P_{t-1} \geq 5,1 \\ 0 & \text{si } P_t < 5,1 \end{cases}$$

Avec : $\begin{cases} T_t & \text{Températures maximales du jour } t \text{ en } ^\circ \text{C} \\ P_t & \text{Précipitations totales du jour en } mm \\ \bar{P} & \text{Précipitations annuelles moyennes en } mm \end{cases}$

6.1.2 Simulation sur le sol français

Avant de retenir définitivement le KBDI comme sous-jacent, nous souhaitons voir si son évolution dans le temps est cohérente avec la sinistralité historique mesurée en France. Pour cela nous avons calculé les valeurs prises par le KBDI entre les années 2000 et 2020 pour la ville de Toulouse :

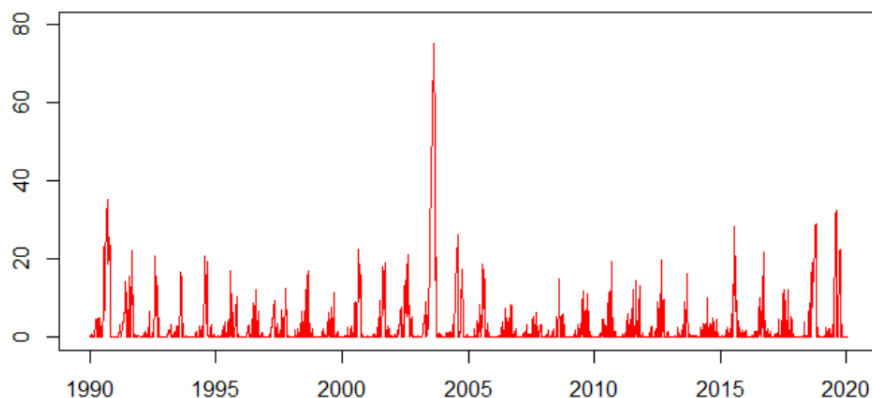


FIGURE 6.1 – KBDI historique pour la ville de Toulouse

Nous pouvons remarquer que la valeur maximale annuelle du KBDI traduit bien les années de grande sinistralité liée à la sécheresse exposée précédemment [Figure1.3]. Nous retrouvons en effet aisément les quatre années les plus sinistrées : 1990, 2003, 2018 et 2019.

Notons également que le fait que le KBDI prenne en considération à la fois la pluie et la température permet à cet indice de détecter la sécheresse, qu'elle soit d'ordre hydraulique ou caniculaire. Pour preuve, selon Météo-France, la sécheresse de 1990 fut essentiellement liée aux "déficits d'humidité du sol et de précipitations". D'après la CCR, la sécheresse de 2003 fut, quant à elle, causée par "une très longue période de fortes chaleurs" comprenant pourtant quelques "puissantes vagues orageuses ponctuelles".

Quatrième partie

Construction d'un générateur
climatique

Afin de tarifier au mieux des options basées sur le KBDI, nous allons, dans cette partie, mettre en place un générateur climatique. L'objectif est de pouvoir, pour un emplacement donné, simuler le KBDI et ainsi définir la prime pure d'une option climatique comme l'espérance de son payoff calculée sur une grande quantité de périodes de sécheresses simulées.

Pour y arriver, nous commencerons par extraire les données de températures maximales et de pluies journalières de la base de données POWER à l'aide de l'API R présentée précédemment. Ensuite, nous modéliserons les jours pluvieux par une chaîne de Markov et l'intensité de la pluie par une loi gamma. Parallèlement à cela, nous modéliserons la température maximale journalière par un processus autorégressif de retour à la moyenne conditionné selon les précipitations du jour précédent. Enfin, nous pourrions générer notre indice sécheresse à partir des pluies et températures simulées.

Afin de simplifier la lecture, nous appellerons désormais "température" la température maximale journalière. Nous n'étudierons pas les températures moyennes.

Sans perte de généralité, nous avons choisi de détailler la calibration et restituer les résultats de notre générateur climatique dans quatre villes françaises : Paris, Lyon, Rennes et Toulouse. Bien entendu, la méthode explorée ici est applicable à tout autre emplacement géographique dont le climat est comparable à celui de la France.

La calibration de nos modèles sera généralement réalisée sur la période du 1^{er} janvier 1990 au 31 décembre 2009. Cette période est dite période de calibration. L'ajustement de notre modèle sera quant à lui généralement étudié sur la période du 1^{er} janvier 2010 au 31 décembre 2020. Cette période est dite période de test du modèle.

Chapitre 7

Modélisation des jours pluvieux

Avant de modéliser la quantité de précipitation journalière, nous allons modéliser les jours pluvieux. Nous allons donc modéliser la succession de jours pluvieux ou secs par un processus de Markov à temps discret. Ce processus comporte deux états notés 0 et 1 qui représenteront respectivement un jour sec et un jour pluvieux [Wilks et Wilby, 1999], [Williams, 1952], [Longley, 1953] et [Gabriel and Neumann, 1962].

7.1 Chaîne de Markov à deux états

7.1.1 Définitions préliminaires

Soit E un ensemble fini d'états et $X = (X_n) \in E^{\mathbb{N}}$ une suite à valeur dans cet espace d'états.

X est une chaîne de Markov si :

$$\begin{aligned} \forall k \in \mathbb{N}, \forall (x_0, \dots, x_{k+1}) \in E^{k+2} \text{ tels que } \mathbb{P}(X_k = x_k, \dots, X_0 = x_0) > 0, \\ \mathbb{P}(X_{k+1} = x_{k+1} \mid X_k = x_k, \dots, X_0 = x_0) = \mathbb{P}(X_{k+1} = x_{k+1} \mid X_k = x_k) \end{aligned}$$

De plus, une chaîne de Markov X est dite homogène si :

$$\begin{aligned} \forall k \in \mathbb{N}, \forall a \in E \text{ et } \forall b \in E, \\ \mathbb{P}(X_{k+1} = b \mid X_k = a) = \mathbb{P}(X_1 = b \mid X_0 = a) \end{aligned}$$

7.1.2 Matrice de transition

Considérons maintenant X comme étant une chaîne de Markov homogène à deux états. La probabilité que X soit dans l'état j à l'instant $t+1$, sachant que X a été dans l'état i à l'instant t , est donc égale à la probabilité que X soit dans l'état j à l'instant $t = 1$ sachant que X a été dans l'état i à l'instant $t = 0$. Nous notons alors p_{ij} cette probabilité appelée probabilité de transition à un pas de l'état i vers l'état j .

Cela se traduit mathématiquement par :

$$p_{ij} = \mathbb{P}(X_{k+1} = j \mid X_k = i) = \mathbb{P}(X_1 = j \mid X_0 = i)$$

Pour un processus comprenant deux états notés 0 et 1, les probabilités de transitions peuvent être regroupées en une matrice \mathcal{P} appelée matrice de transition et définie de la manière suivante :

$$\mathcal{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

Par construction, une telle matrice vérifie :

- $\forall (ij) \in \{0, 1\}^2, 0 \leq p_{ij} \leq 1$
- $\forall i \in \{0, 1\}, \sum_{j \in \{0, 1\}} p_{i,j} = 1$

7.2 Estimation des coefficients

Nous allons maintenant modéliser la succession de jours pluvieux et de jours secs par X un processus de Markov homogène à temps discret comportant deux états. Ainsi, nous aurons pour tout n :

$$X_n = \begin{cases} 1 & \text{si le jour } n \text{ est pluvieux} \\ 0 & \text{si le jour } n \text{ est sec} \end{cases}$$

Afin d'éliminer des pluies parasites, nous considérons qu'un jour est pluvieux dès que la précipitation journalière dépasse 1mm.

7.2.1 Protocole

Pour i et j appartenant aux états 0 et 1, notons p_{ij} la probabilité de transition de l'état i à l'état j .

Nous cherchons maintenant à estimer la matrice de transition associée au processus de Markov X .

Pour prendre en compte la saisonnalité, nous définirons une matrice différente pour chaque mois m de l'année.

Nous allons donc déterminer pour chaque mois m compris entre 1 et 12, la probabilité $p_{00}^{(m)}$ de transition d'un jour sec à un jour sec et la probabilité $p_{11}^{(m)}$ d'un jour pluvieux à un jour pluvieux et définir la matrice de transition mensuelle $\mathcal{P}^{(m)}$ définie ci-dessous :

$$\mathcal{P}^{(m)} = \begin{pmatrix} p_{00}^{(m)} & p_{01}^{(m)} \\ p_{10}^{(m)} & p_{11}^{(m)} \end{pmatrix} = \begin{pmatrix} p_{00}^{(m)} & 1 - p_{00}^{(m)} \\ 1 - p_{11}^{(m)} & p_{11}^{(m)} \end{pmatrix}$$

Pour chaque mois m , nous calculons $p_{00}^{(m)}$ et $p_{11}^{(m)}$ par la démarche suivante :

- Extraction des données de pluie :

Tout d'abord, nous commençons par extraire les données de pluies journalières ainsi que les données de pluies de la veille sur l'ensemble des mois m entre 1990 et 2010. Nous attribuons ensuite à chaque coordonnée de ces vecteurs la valeur 1 si la quantité de pluie dépasse 1mm et 0 si celle est inférieure à ce seuil.

- Comptage des différents scénarios :

$$\text{Nous définissons : } \begin{cases} S_{sec} = \text{nombre de jours secs suivis d'un jour sec} \\ S_{pluie} = \text{nombre de jours pluvieux suivis d'un jour pluvieux} \\ N_{sec} = \text{nombre de jours secs} \\ N_{pluie} = \text{nombre de jours pluvieux} \end{cases}$$

- Estimation des coefficients de la matrice de transition :

$$\text{Nous estimons maintenant : } \begin{cases} p_{00}^{(m)} = \frac{S_{sec}}{N_{sec}} \\ p_{11}^{(m)} = \frac{S_{pluie}}{N_{pluie}} \end{cases}$$

$$\text{Et déduisons : } \begin{cases} p_{01}^{(m)} = 1 - p_{00}^{(m)} \\ p_{10}^{(m)} = 1 - p_{11}^{(m)} \end{cases}$$

7.3 Résultats obtenus

En appliquant le protocole énoncé précédemment à notre base de calibration qui contient les données de pluies de janvier 1990 à décembre 2009, nous avons pu estimer les coefficients mensuels de la matrice de transition. Nous regardons toujours exclusivement les probabilités $p_{00}^{(m)}$ et $p_{11}^{(m)}$ qui définissent entièrement le processus de Markov modélisant les pluies.

Le graphique présenté ci-dessous confirme que la probabilité de transition p_{00} d'un jour sec à un jour sec possède un caractère saisonnier. Nous remarquons qu'il en est de même pour la probabilité de transition p_{11} . L'hypothèse émise sur l'utilité du fractionnement de la matrice de transition est donc validée au vu de ce résultat.

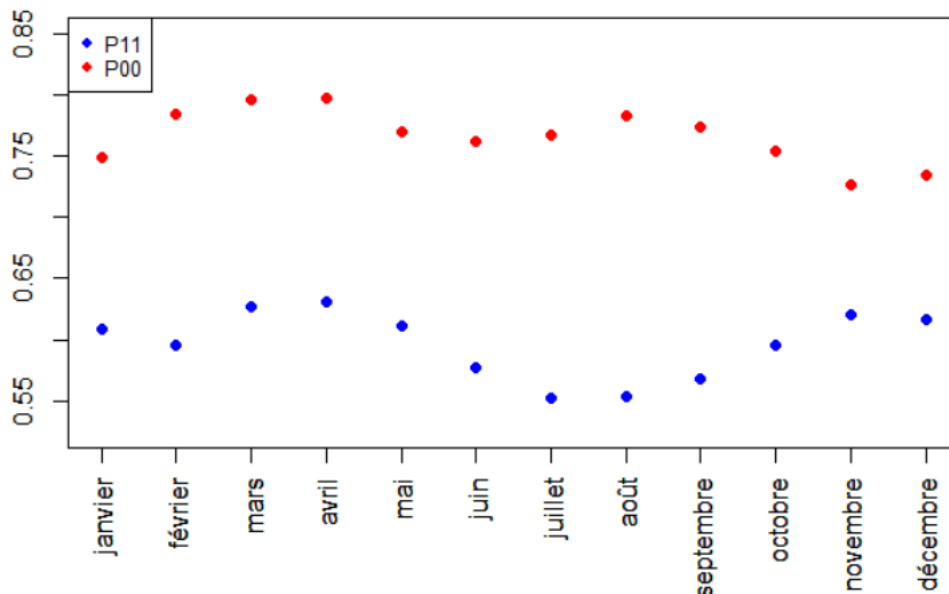


FIGURE 7.1 – Probabilités de transitions estimées pour chaque mois à Paris

La variation saisonnière des probabilités de transitions met en avant des propriétés climatiques propres à chaque mois pour la ville de Paris. Par exemple, nous pouvons voir qu'en juillet et en août la probabilité qu'un jour pluvieux succède à un jour pluvieux est significativement plus faible que pour le reste de l'année. Par conséquent, les phases de pluies ont tendance à durer moins longtemps en été qu'en hiver.

Nous constatons également que durant le mois de novembre, les coefficients $p_{00}^{(11)}$ et $p_{11}^{(11)}$ sont presque identiques. La matrice de transition du mois de novembre s'approche alors d'une matrice symétrique. Ainsi, plus ces deux coefficients sont proches, plus nous pouvons nous attendre à ce que le générateur construise autant de jours secs que de jours pluvieux. En été l'effet inverse se produit, l'écart entre $p_{00}^{(11)}$ et $p_{11}^{(11)}$ augmente et le générateur a tendance à générer davantage de jours secs que pluvieux. La lecture de ce graphique confirme alors certaines intuitions que nous pouvons avoir, comme le fait que le changement de régime entre les jours pluvieux et les jours secs s'effectue plus souvent en hiver et que durant les mois de juillet et août, les pluies ont tendances à être davantage ponctuelles que pendant le reste de l'année.

Étant donné que nous souhaitons vérifier que notre méthode est généralisable sur l'ensemble du sol français, nous avons vérifié que des résultats similaires sont trouvés pour les villes de Lyon, Rennes et Toulouse. Le graphique ci-dessous montre la distribution des probabilités $p_{00}^{(m)}$ et $p_{11}^{(m)}$ pour l'ensemble des quatre villes considérées.

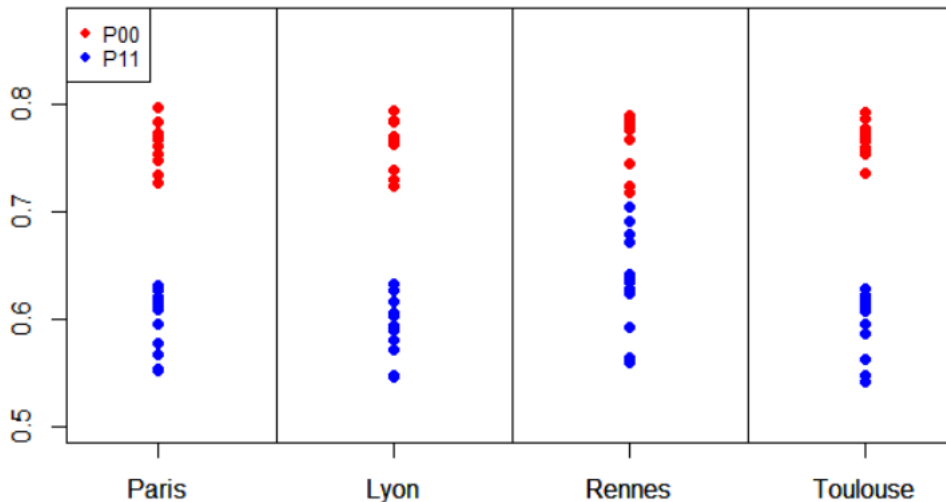


FIGURE 7.2 – Probabilités de transitions estimées pour chaque ville étudiée

La distribution des probabilités de transition est homogène pour les villes de Paris, Lyon et Toulouse. Cependant, la distribution des probabilités $p_{11}^{(m)}$ pour Rennes est plus étalée que pour les trois autres villes. Nous allons donc vérifier la cohérence de la valeur des $p_{11}^{(m)}$ estimés pour la ville de Rennes. Pour cela, nous allons regarder le graphique suivant qui détaille, pour chaque ville, l'évolution mois par mois des coefficients estimés.

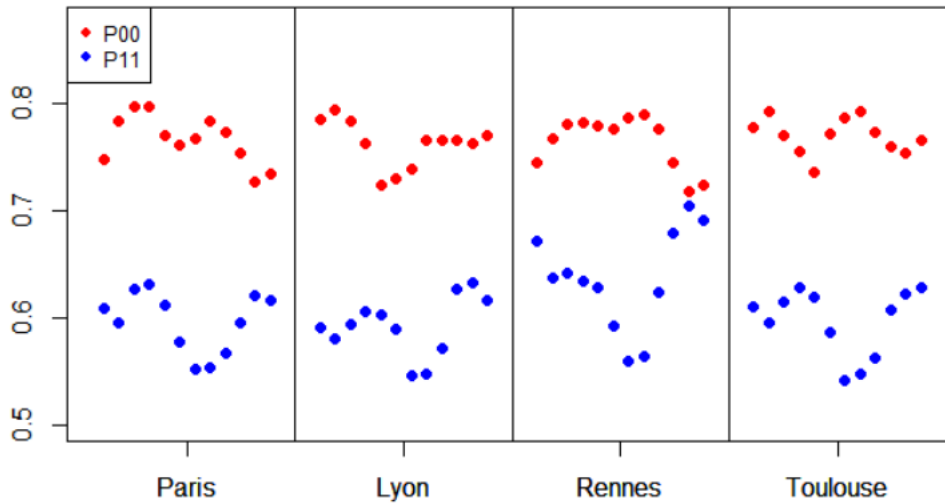


FIGURE 7.3 – Probabilités de transitions détaillées pour chaque ville et chaque mois

Nous pouvons conclure que l'évolution mensuelle des probabilités de transition est cohérente sur l'ensemble des quatre villes étudiées. Bien que la distribution des probabilités $p_{11}^{(m)}$ est davantage étalée pour la ville de Rennes, ce détail mois par mois montre que l'évolution des coefficients $p_{11}^{(m)}$ est cohérente et peut se traduire par le fait qu'en hiver, les périodes de pluies durent plus longtemps à Rennes qu'à Paris, Lyon ou Toulouse.

Enfin, nous avons voulu tester si notre modélisation restitue correctement la longueur des périodes sèches. Pour cela nous avons effectué un test d'adéquation du khi-deux sur l'hypothèse H_0 selon laquelle les durées des périodes sèches de notre historique suivent la même loi les durées des périodes sèches simulées. En effectuant ce test nous obtenons les résultats suivants :

Ville	statistique du test	p.value du test	Présomption contre H_0
Paris	80	0,2424	pas de présomption
Lyon	120	0,2372	pas de présomption
Rennes	72	0,1951	pas de présomption
Toulouse	154	0,09244	pas de présomption

Ayant choisi un niveau de confiance de 95%, les résultats du test d'adéquation du khi-deux ne remettent donc pas en cause l'hypothèse H_0 .

Au vu de ces résultats, nous validons la modélisation des jours pluvieux ainsi que le fractionnement mensuel de la matrice de transition de la chaîne de Markov.

Chapitre 8

Modélisation de l'intensité des pluies

Maintenant que nous savons générer des jours secs et des jours pluvieux, il reste à modéliser le volume des précipitations associé à ces jours de pluie. La modélisation du volume des précipitations est un sujet qui a déjà été abordé par le passé [Thom, 1958], [Katz, 1977], [FALLOT et SAUTEBIN, 2018]. Des résultats ont alors montré qu'il est possible de modéliser la quantité de précipitation journalière par une loi gamma.

8.1 Rappels sur la loi gamma

8.1.1 Définition de la loi gamma

La loi gamma provient de la fonction gamma notée Γ . Définie pour tout nombre complexe dont la partie réelle est strictement positive. Cette fonction s'écrit :

$$\Gamma : z \mapsto \int_0^{+\infty} t^{z-1} e^{-t} dt$$

Nous retrouvons cette fonction dans la densité de probabilité de la fonction gamma. En effet, soit X une variable aléatoire suivant une loi gamma de paramètres α et β . Sa densité de probabilité f est définie pour tout réel strictement positif et s'écrit :

$$f : x \mapsto \frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-x/\beta}$$

$$\text{Où : } \begin{cases} \alpha \text{ est un paramètre de forme} \\ \beta \text{ est un paramètre d'échelle} \end{cases}$$

Le graphique suivant illustre l'impact du paramètre de forme et d'échelle sur la distribution de probabilité de la loi gamma.

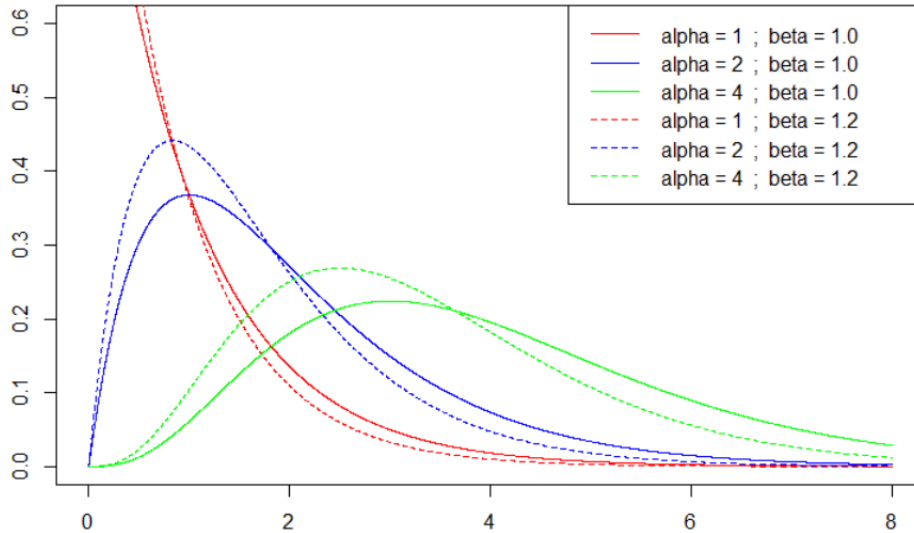


FIGURE 8.1 – Influence des paramètres α et β sur la distribution de la loi gamma

Nous pouvons remarquer que selon la valeur du paramètre α l'aspect de la densité de probabilité change très rapidement. Cela est particulièrement visible lorsque ce paramètre devient strictement inférieur à 1 provoquant une divergence de la densité en 0. Nous pouvons voir que lorsque l'on augmente la valeur de β , la densité se décale vers les valeurs plus faibles.

8.2 Estimation des paramètres de la loi gamma

Nous allons maintenant modéliser la pluie grâce à une loi gamma. Comme la présence d'une saisonnalité a été démontrée lors de la modélisation des jours pluvieux, nous allons estimer les paramètres α et β indépendamment pour chaque mois de l'année. La méthode des moments, des moindres carrés ou du maximum de vraisemblance peuvent être envisagées pour l'estimation de ces paramètres. Parmi elles, la méthode des moments à été retenue car elle est appropriée à la modélisation des précipitations quotidiennes [FALLOT et SAUTEBIN, 2018].

8.3 Méthode des moments

Si X suit la loi gamma de paramètres α et β , son espérance et sa variance valent :

$$\mathbb{E}[X] = \frac{\alpha}{\beta} \quad \text{et} \quad \text{Var}[X] = \frac{\alpha}{\beta^2}$$

Grâce à ces expressions, il est possible d'exprimer α et β en fonction de $\mathbb{E}[X]$ et de $\text{Var}[X]$.

$$\alpha = \frac{\mathbb{E}[X]^2}{\text{Var}[X]} \quad \text{et} \quad \beta = \frac{\mathbb{E}[X]}{\text{Var}[X]}$$

Ainsi, α et β peuvent respectivement être estimés par $\hat{\alpha}$ et $\hat{\beta}$ définis par :

$$\hat{\alpha} = \frac{\bar{X}^2}{\bar{V}} \quad \text{et} \quad \hat{\beta} = \frac{\bar{X}}{\bar{V}}$$

avec :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad \bar{V} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Nous pouvons noter que cette méthode présente l'avantage d'estimer les paramètres par une formule fermée. Cela permettra donc un calcul rapide et ainsi, notre générateur pourra calibrer simultanément des lois gamma sur un grand nombre de sites géographiques de manière efficace.

8.3.1 Résultats obtenus

Pour la calibration de ce modèle, nous avons tout d'abord extrait les jours pluvieux de notre base de test. Nous avons ensuite extrait les données pour chaque mois et estimé les paramètres de forme et d'échelle par la formule énoncée précédemment.

Le graphique ci-dessous confirme que les paramètres α et β possèdent un caractère saisonnier. L'oscillation de α pourrait s'apparenter à une sinusoïde. Au vu de ce résultat, l'hypothèse émise sur l'utilité du conditionnement mensuel de la loi gamma calibrée est donc validée.

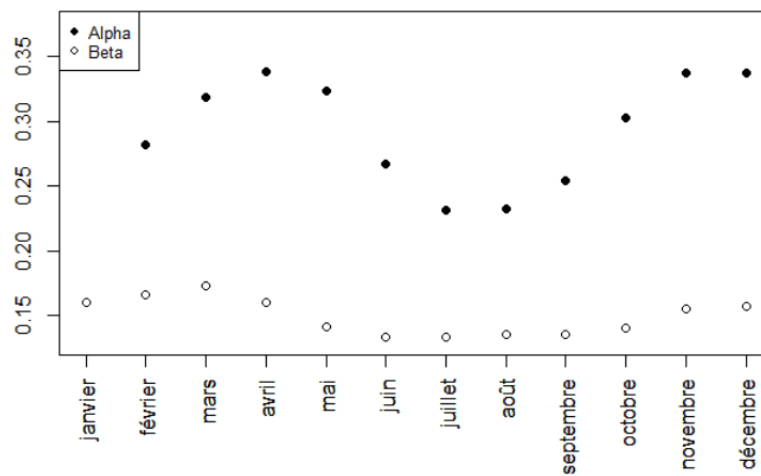


FIGURE 8.2 – Estimation de α et β en fonction du mois pour la ville de Paris

La distribution des paramètres estimés semble dépendre du climat. En effet, nous remarquons sur le graphique ci-dessous que les climats pourraient être séparés en deux groupes : les villes du nord et celles du sud. Malgré ces quelques nuances, les distributions sont bien cohérentes pour l'ensemble des quatre villes étudiées.

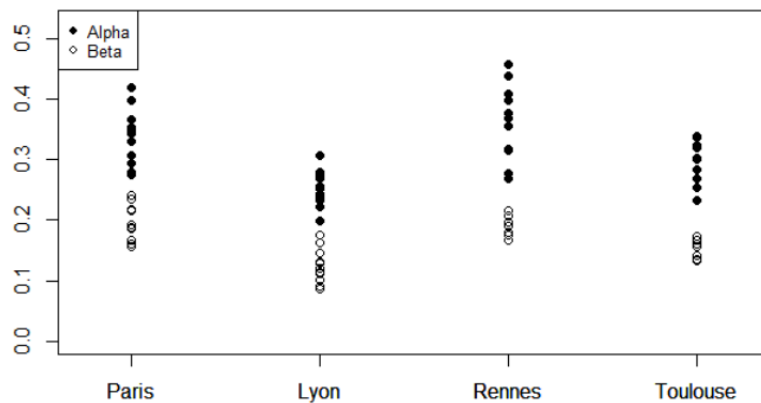


FIGURE 8.3 – Estimation de α et β pour chaque ville étudiée

8.4 Ajustement de la loi gamma sur nos données de précipitations

Nous allons maintenant voir si la loi gamma, dont les paramètres ont été calculés sur notre base de calibration, est cohérente avec le volume de précipitations présent dans notre base de données. Nous regarderons d'abord l'ajustement du modèle sur les données de la base de calibration, puis ferons de même sur la base de test afin de repérer d'éventuelles divergences du modèle dans le temps.

Pour simplifier la lecture, nous allons montrer des résultats trimestriels. Le trimestre 1 correspondra ainsi au mois de février, trimestre 2 au mois de mai, le trimestre 3 à août et enfin le mois de novembre représentera le trimestre 4.

Nous pouvons voir ci-dessous le diagramme quantile-quantile entre notre modèle gamma et la précipitation issue de la base de calibration pour chaque trimestre :

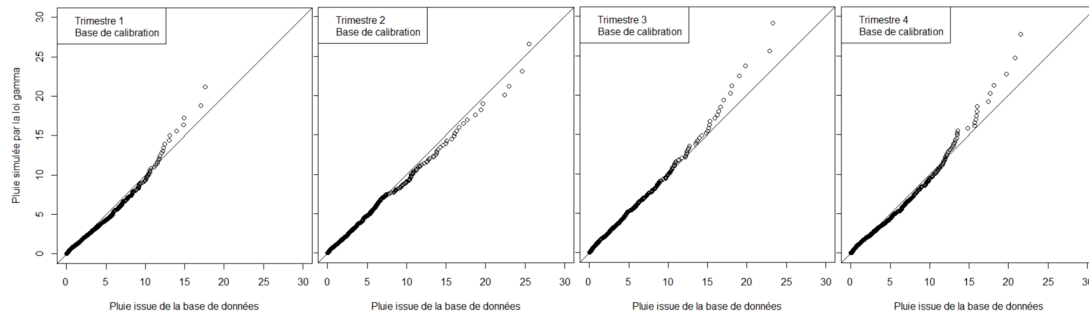


FIGURE 8.4 – Quantile de la loi gamma en fonction du quantile base de calibration pour la ville de Paris

Nous pouvons constater que la modélisation est très bien ajustée pour les pluies inférieures à 15 millimètres et que la loi gamma a tendance à diverger légèrement pour précipitations extrêmes.

Malgré cette divergence dans la queue de la distribution, nous avons choisi de valider l'ajustement de la loi gamma sur notre base de calibration. Ce choix a été motivé par deux raisons principales. Tout d'abord, nous avons analysés dans notre base de données que 99 % des précipitations sont inférieures à 14 millimètres. Le deuxième argument en faveur de cette validation est le fait qu'une surestimation éventuelle d'une pluie extrême sera sans conséquence pour le calcul de notre indice sécheresse car il sera systématiquement réinitialisé à 0 en cas de pluie extrême. À l'inverse, si la modélisation de la pluie était réalisée pour l'étude d'inondations, un tel écart n'aurait pas été acceptable.

Analysons maintenant le diagramme quantile-quantile entre notre modèle gamma et la précipitation issue de la base de test pour chaque trimestre :

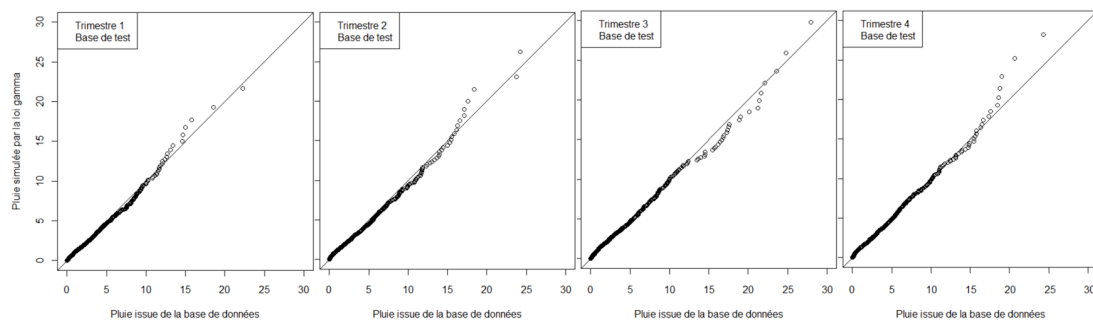


FIGURE 8.5 – Quantile de la loi gamma en fonction du quantile base de test

La loi gamma reste toujours aussi bien ajustée sur la base de test. Il est donc envisageable d'utiliser ce modèle pour effectuer des projections de pluie dans le temps. Nous avons ainsi justifié le choix de la loi gamma pour simuler la précipitation journalière de manière cohérente avec notre problématique de construction d'un indice sécheresse.

Chapitre 9

Modélisation de la température

La température journalière possède un caractère saisonnier évident et c'est pour cette raison que modéliser cette grandeur par une série temporelle est à priori légitime. Dans le cadre de la tarification de dérivés climatiques basés sur un indice de température il a été démontré que les températures journalières peuvent bien être modélisées par une série temporelle couplée à un processus autorégressif de retour à la moyenne [MORENO et ROUSTANT, 2002], [DUTANG et al., 2017]. Avant d'implémenter cette modélisation, nous rappellerons certaines notions théoriques essentielles sur les séries temporelles.

Afin de rendre plus claire la restitution des résultats, nous détaillerons dans ce chapitre la construction de notre modèle de température pour la ville de Paris. Pour les villes de Lyon, Rennes et Toulouse, la démarche a été identique et a abouti à des conclusions similaires.

9.1 Introduction aux séries temporelles

Dans le cadre général, une série temporelle modélise la valeur X_t d'une variable climatique à l'instant t par la somme de trois composantes : la moyenne M_t , la saisonnalité Θ_t et un processus R_t qui représente le caractère aléatoire de l'évolution de la variable climatique dans le temps. Ainsi, à partir de n observations T_1, \dots, T_n , nous souhaitons établir une relation de la forme :

$$\forall t \in \mathbb{N}, \quad X_t = M_t + \Theta_t + R_t$$

Nous allons voir que la principale difficulté de cette modélisation réside dans la bonne appréhension du caractère aléatoire de la composante R . Bien modéliser cette composante probabiliste est d'ailleurs la clé d'une projection efficace de notre série dans le temps. Une attention particulière sera donc portée à la modélisation de la composante R car notre objectif est justement de pouvoir projeter la série des températures sur un grand nombre d'années.

Nous allons aborder ici les notions et définitions essentielles à l'étude d'une série temporelle. Pour cela, notons $X = (X_t)_{t \in \mathbb{N}} \in L^2(\Omega, \mathcal{F}, P)$ une série temporelle indexée par le temps.

Stationnarité d'un processus

Le processus X est stationnaire au sens large si :

- $\mathbb{E}(X) \perp t$
- $\mathbb{V}(X) \perp t$
- $\forall t \in \mathbb{Z}$ et $\forall s \in \mathbb{Z}$, $\text{cov}(X_t, X_s) = \text{cov}(X_{t-s}, X_0)$

Vous trouverez en Annexe, la définition de stationnarité au sens strict.

La notion de stationnarité illustre le fait que les propriétés stochastiques d'un processus ne varient pas au cours du temps. En pratique, il est rare qu'un processus soit rigoureusement stationnaire. Cependant, chacune des trois conditions de stationnarité énoncées ci-dessus doit être dans une certaine mesure respectée. En effet, si les propriétés stochastiques varient dans le temps, cela remet en cause la viabilité d'une projection de la série dans le temps.

Autocorrélation d'un processus

La fonction γ d'autocovariance du processus stationnaire X mesure la covariance entre X et cette même variable décalée de h pas de temps. Cette fonction est définie par :

$$h \mapsto \gamma(h) = \text{cov}(X_t, X_{t-h}) = \mathbb{E}(X_t X_{t-h}) - \mathbb{E}(X_t) \mathbb{E}(X_{t-h})$$

Une approximation empirique de γ à partir de n observations peut être :

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X})$$

où \bar{X} est la moyenne empirique de nos n observations.

La fonction ρ d'autocorrélation du processus stationnaire X mesure la corrélation entre X et cette même variable décalée de h pas de temps. Cette fonction est définie par :

$$h \mapsto \rho(h) = \text{cor}(X_t, X_{t-h}) = \frac{\text{cov}(X_t, X_{t-h})}{\sqrt{\text{Var}(X_t)} \sqrt{\text{Var}(X_{t-h})}}$$

Nous pouvons en déduire l'expression de ρ en fonction de l'autocovariance γ :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

Cette formule permet de trouver directement une approximation empirique de ρ à partir de n observations :

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

L'étude de l'autocorrélation est importante car lors de la projection de la série temporelle, il ne faut pas que l'on simule la composante aléatoire chaque jour de manière indépendante si celle-ci dépend de la valeur qu'elle a prise dans le passé.

Bruit blanc :

$(\varepsilon_t)_t \subset L^2(\Omega, \mathcal{F}, P)$ est un bruit blanc de variance σ^2 si et seulement si :

$$\forall s, \forall t, \quad \mathbb{E}(\varepsilon_t) = 0, \text{ et } \text{cov}(\varepsilon_t, \varepsilon_s) = \begin{cases} 0 & \text{si } s \neq t \\ 1 & \text{si } s = t \end{cases}$$

Si $(\varepsilon_t)_t$ est également indépendant et identiquement distribué, on dit que c'est un bruit blanc iid.

$(\varepsilon_t)_t \subset L^2(\Omega, \mathcal{F}, P)$ est un bruit blanc gaussien de variance σ^2 si et seulement si $(\varepsilon_t)_t$ est un bruit blanc iid tel que :

$$\forall t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

Nous verrons à la fin de l'étape de modélisation que la composante aléatoire de notre série de températures pourra se résumer à un simple bruit blanc gaussien.

9.2 Estimation de la tendance et de la saisonnalité

Nous cherchons dans un premier temps à extraire la tendance M et la saisonnalité Θ du processus T des températures journalières.

Tendance

Pour l'estimation de la tendance, nous avons estimé le coefficient directeur de la composante linéaire T avec la méthode des moindres carrés. Le tableau suivant dresse le bilan des coefficients estimés pour les villes de Paris, Lyon, Rennes et Toulouse.

Ville	Hausse journalière	Réchauffement depuis 1986 (° C)
Paris	$6,48 \cdot 10^{-5}$	0.8
Lyon	$6,35 \cdot 10^{-5}$	0.8
Rennes	$2,56 \cdot 10^{-5}$	0.3
Toulouse	$3,72 \cdot 10^{-5}$	0.5

Le fait d'estimer le réchauffement climatique par une composante linéaire est un moyen efficace de projeter les températures sur un horizon de quelques années. Cependant cette modélisation comporte ses limites si l'on veut simuler des températures sur plusieurs dizaines d'années. En effet, le Groupe d'experts Intergouvernemental sur l'Evolution du Climat (GIEC), organisme scientifique de référence sur le réchauffement climatique, a annoncé qu'après avoir connu une tendance linéaire par le passé, la température pourrait être en hausse exponentielle jusqu'à la fin du siècle. Le modèle implémenté ici ne permet donc pas de rendre compte automatiquement d'une hausse des températures pouvant être anticipée à l'avenir. Cela présente un risque car il est possible que le tarif d'un dérivé ayant une date d'exercice trop lointain ne reflète pas

correctement un risque de sécheresse qui serait accru dans le futur et la prime demandée serait alors trop faible pour couvrir l'espérance du payoff futur du dérivé.

En 2014, le GIEC a publié un éventail de scénarios de hausse de températures pouvant être attendus à la fin du XXI^e siècle. Nous conseillons donc de prendre ces scénarios pour référence dans le cadre de projections de températures sur un horizon dépassant une dizaine d'années.

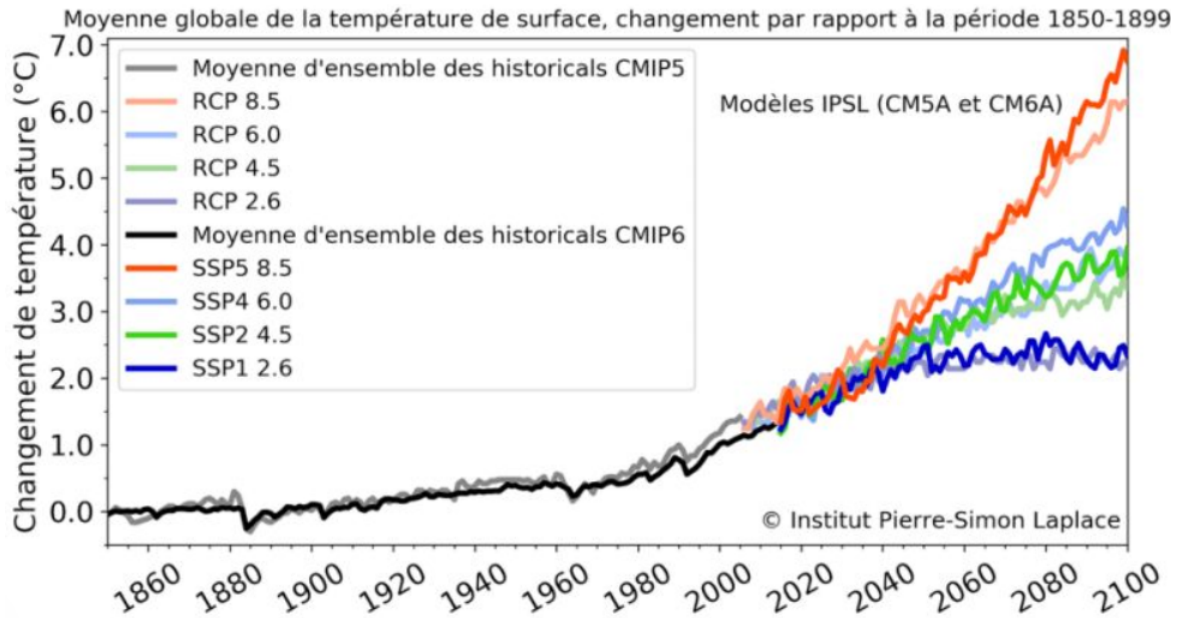


FIGURE 9.1 – Projection des scénarios du GIEC jusqu'à 2100

Saisonnalité

Nous allons maintenant estimer la composante saisonnière Θ grâce à une régression linéaire entre la température T et un ensemble de fonctions sinusoïdales 365.25 périodiques. Nous avons fait ce choix car nous savons que le cycle saisonnier des températures dure un an, ce qui correspond en moyenne à 365.25 jours. Nous cherchons donc à estimer les coefficients $A1, B1, A2, B2, A3$ et $B3$ dans la formule suivante :

$$\begin{aligned} \Theta_t = & A1 \sin\left(\frac{2\pi}{365,25}t\right) + B1 \cos\left(\frac{2\pi}{365,25}t\right) \\ & + A2 \sin\left(\frac{4\pi}{365,25}t\right) + B2 \cos\left(\frac{4\pi}{365,25}t\right) \\ & + A3 \sin\left(\frac{8\pi}{365,25}t\right) + B3 \cos\left(\frac{8\pi}{365,25}t\right) \end{aligned}$$

Le tableau suivant dresse le bilan des coefficients estimés sous R par la régression linéaire exposée ci-dessous :

Coefficient	Estimation	Significativité (de 0 à 5)
A1	-9.03	5
B1	-2.63	5
A2	-0.81	5
B2	1.11	5
A3	-0.21	4
B3	0.24	4

Nous constatons que chacune des composantes a un niveau de significativité important. Nous avons également essayé d'ajouter à ce modèle d'autres composantes sinusoïdales de période 365,25 mais aucune d'entre elles n'est significative.

Nous allons donc retenir cette modélisation pour la composante Θ . Posons alors :

$$\Theta_t = -9.03 \sin\left(\frac{2\pi}{365,25}t\right) - 2.63 \cos\left(\frac{2\pi}{365,25}t\right) - 0.81 \sin\left(\frac{4\pi}{365,25}t\right) + 1.11 \cos\left(\frac{4\pi}{365,25}t\right) - 0.21 \sin\left(\frac{8\pi}{365,25}t\right) + 0.24 \cos\left(\frac{8\pi}{365,25}t\right)$$

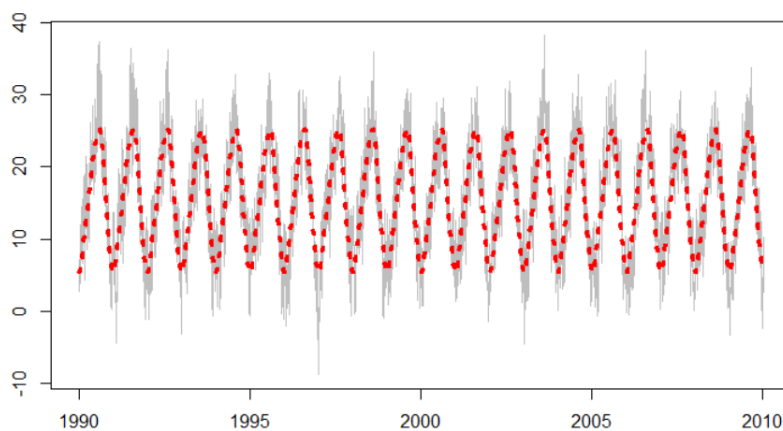


FIGURE 9.2 – Modélisation de la saisonnalité des températures

Maintenant que nous avons extrait la tendance et la saisonnalité de la série T , nous allons devoir modéliser le résidu R .

Analyse du résidu R

Après avoir extrait la saisonnalité de la série temporelle des températures, nous allons étudier les propriétés du résidu R . Nous rappelons que c'est cette composante qui contient le caractère aléatoire de l'évolution des températures dans le temps. Il est donc important de bien comprendre sa structure pour pouvoir la modéliser de manière fine et ainsi être capable de projeter efficacement les températures dans le temps.

Nous remarquons tout d'abord que la distribution de R est très proche d'une gaussienne :

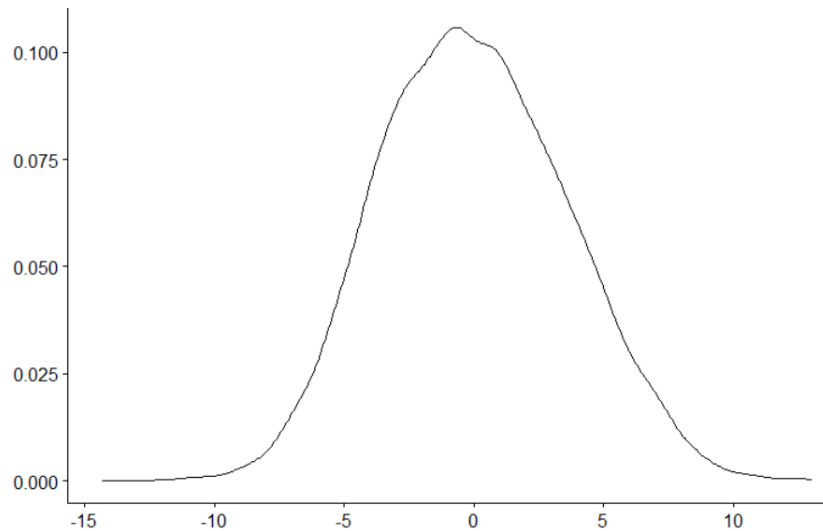


FIGURE 9.3 – Distribution du résidu R

Au vu de cette courbe, nous pourrions penser que les résidus peuvent être directement modélisés par une loi normale. Cependant, il est primordial de prendre en compte le fait que la valeur de notre résidu peut être conditionnée par ses valeurs passées.

Le graphique suivant, représentant l'autocorrélation de R pour un retard h variant de un jour à un mois, suffit à invalider l'hypothèse selon laquelle nous pouvons modéliser ce résidu par un bruit blanc indépendant :

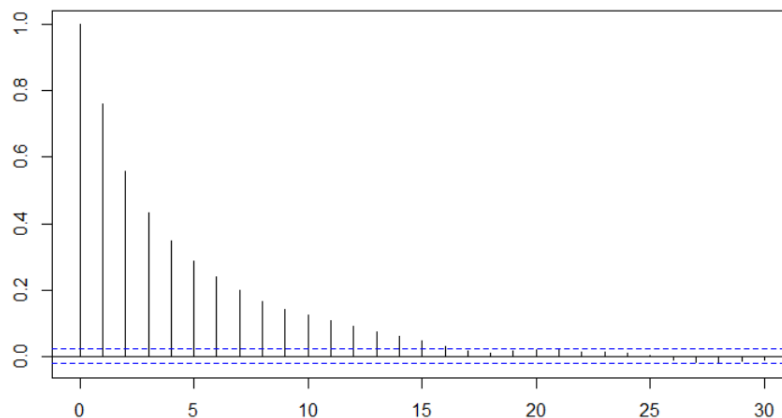


FIGURE 9.4 – Autocorrélation du résidu R

Nous pouvons en effet constater que $\rho(1) = \text{cor}(R_t, R_{t-1}) = 0,76$. Cela signifie que l'écart avec la température saisonnière est fortement corrélé à cet écart mesuré la veille. Ce constat est logique avec la réalité physique des températures. Par exemple, il y a effectivement plus de chances que l'on soit au dessus des moyennes de saison s'il a fait très chaud toute la semaine précédente, comparativement à un scénario où les températures auraient été récemment en dessous de la moyenne saisonnière.

Différenciation du résidu R

Une méthode simple pour éliminer cette autocorrélation existe. Elle consiste à différencier le résidu et modéliser l'incrément du résidu avec un pas de 1 jour.

Notons $\varepsilon_t = R_t - R_{t-1}$ le résidu différencié de notre série. Le graphique suivant représente l'autocorrélation de la série ε pour un retard h variant de un jour à un mois :

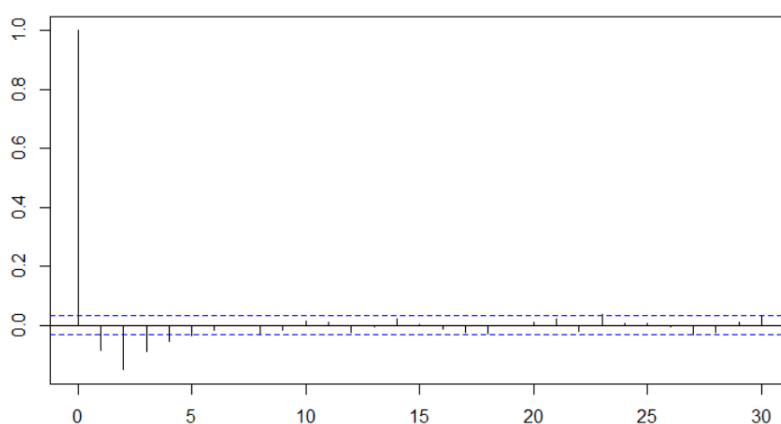


FIGURE 9.5 – Autocorrélation du résidu différencié ε

Nous pouvons constater que la différenciation permet d'éliminer en grande partie ce problème d'autocorrélation. Nous remarquons tout de même que les autocorrélations d'ordre 1 à 5 sont toujours significatives. La négativité de leur signe peut se traduire physiquement par le fait que lorsque la température varie dans un sens, elle a tendance à varier dans le sens opposé les jours suivants. Cette réalité physique fait donc immédiatement penser que la température, après s'en être éloignée, revient progressivement vers la tendance saisonnière.

Nous venons ainsi de mettre en évidence le phénomène de retour à la moyenne présent dans notre série de température. Nous allons donc maintenant modéliser le résidu différencié par un processus de retour à la moyenne.

Afin d'illustrer la nécessité d'une modélisation plus fine, nous avons essayé de simuler les températures à Paris de 2020 à 2030 en modélisant directement le résidu différencié par un mouvement brownien :

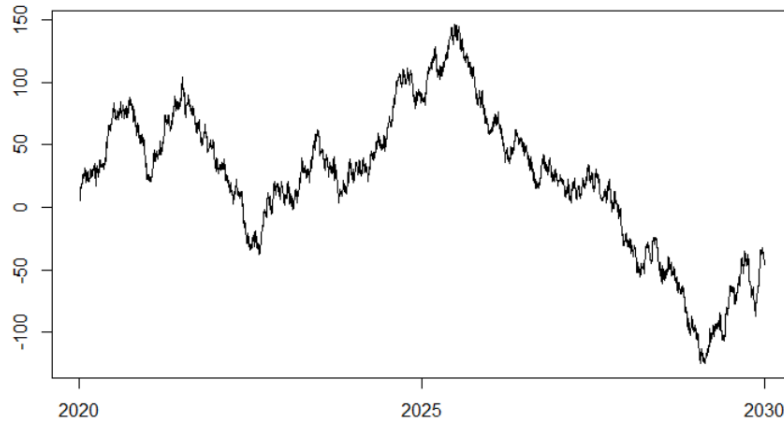


FIGURE 9.6 – Projection des températures sans retour à la moyenne

Il est bien évident qu'un tel modèle n'est pas réaliste car il pourrait prédire des températures de plus de 150°C d'ici 2025. Ceci est bien entendu impossible. De plus, un tel modèle ne permettrait pas de faire des prévisions car l'incertitude portant sur la valeur de la composante ε augmenterait fortement dans le temps et nous pourrions aussi bien prévoir des températures extrêmement chaudes qu'extrêmement froides pour un horizon donné.

9.3 Processus de retour à la moyenne

9.3.1 Formulation

Rappelons que nous avons commencé par estimer la tendance M ainsi que la composante saisonnière Θ du modèle de température suivant :

$$T_t = M_t + \Theta_t + R_t$$

Nous avons ensuite différencié notre série pour enlever l'autocorrélation. Le modèle est alors devenu :

$$T_t - T_{t-1} = M_t - M_{t-1} + \Theta_t - \Theta_{t-1} + \varepsilon_t$$

$$\text{Avec } \varepsilon_t = R_t - R_{t-1}$$

Nous pouvons alors écrire :

$$R_t = R_{t-1} + \varepsilon_t$$

Le modèle de retour à la moyenne consiste à modéliser le résidu ε par la relation suivante :

$$\varepsilon_t = -\alpha R_{t-1} + \omega_t$$

Nous voyons que ce modèle générera un résidu qui se rapprochera de la tendance saisonnière. Le coefficient α est à ce titre appelé force de rappel car c'est ce terme qui détermine la vitesse à laquelle le processus rejoint sa composante déterministe après s'en être éloigné.

Souhaitant construire un générateur climatique qui permettra de simuler conjointement des températures et des précipitations, nous proposons d'améliorer le modèle de retour à la moyenne en ajoutant un terme d'interaction avec la quantité de pluie P_{t-1} tombée la veille. En effet, il semble probable que la température baisse plus rapidement s'il a plu durant la journée précédente. Nous souhaitons établir la relation suivante :

$$\varepsilon_t = -\alpha R_{t-1} + \beta P_{t-1} + \omega_t$$

Compte tenu du caractère saisonnier important de notre série temporelle, nous avons décidé que le paramètre α sera estimé pour chaque mois, grâce à une approche similaire à celle mise en oeuvre lors de la modélisation des précipitations. Nous noterons $\alpha^{(m)}$ ces coefficients, m désignant le mois sur lequel la force de rappel a été estimée. Un fractionnement similaire sera également appliqué pour la composante de pluie.

9.3.2 Estimation de la force de rappel

Pour déterminer la force de rappel mensuelle, commençons par remarquer que, pour un mois m donné, le processus que nous cherchons à modéliser peut s'écrire :

$$R_t = (1 - \alpha^{(m)}) R_{t-1} + \beta^{(m)} P_{t-1} + \omega_t$$

Nous pouvons ainsi estimer la valeur des coefficients $\alpha^{(m)}$ et $\beta^{(m)}$ à l'aide d'une régression linéaire :

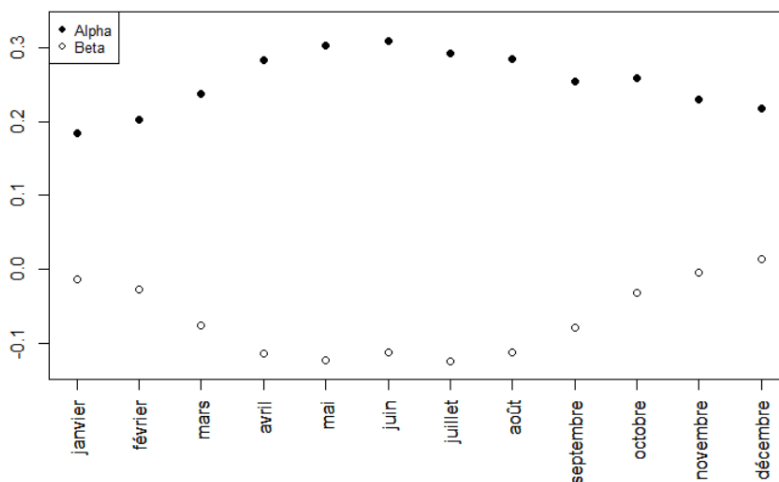


FIGURE 9.7 – Estimation de α et β en fonction du mois

Le caractère saisonnier de la force de rappel et du poids attribué à la pluie tombée la veille est de nouveau bien visible ici. La négativité des coefficient $\beta^{(m)}$ illustre bien le fait qu'une précipitation a tendance à faire baisser la température du lendemain.

9.3.3 Modélisation du résidu ω

Il ne reste maintenant qu'à pouvoir modéliser correctement notre dernier résidu ω pour pouvoir valider notre modèle de température.

Contrairement aux résidus précédemment analysés, nous pouvons constater que l'aurocorrélacion de ω peut être considérée comme nulle pour tout retard allant jusqu'à 1 mois. Nous avons donc réussi à isoler une composante aléatoire qui est indépendante de sa valeur historique. Cela nous permettra de projeter cette composante en générant une valeur qui dépendra uniquement de la loi régissant ω .

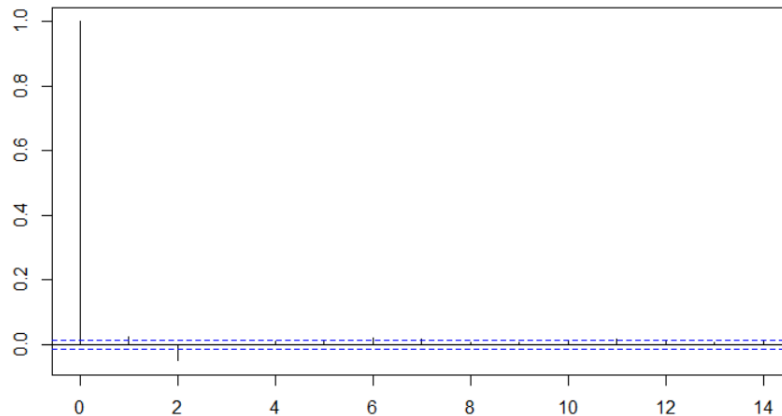


FIGURE 9.8 – Autocorrélation du résidu final ω

Nous allons maintenant voir si nous pouvons modéliser ω par un bruit blanc gaussien. Le graphique ci-dessous représente les quantiles de la modélisation par une loi normale de ω en fonction des quantiles de ω :

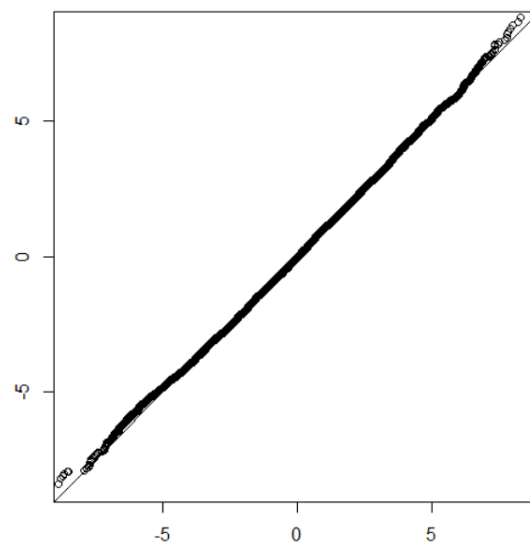


FIGURE 9.9 – Quantile de la loi normale en fonction du quantile de ω

Au vu de ce graphique, nous pouvons considérer cohérente la modélisation ω par une loi normale.

Un autre élément important à contrôler est l'homoscédasticité des résidus. En d'autres termes, il faut s'assurer que les propriétés stochastiques du résidu n'évoluent pas de manière significative dans le temps.

Bien que la variance de ω puisse être considérée comme constante d'une année à l'autre, nous remarquons que cette dernière est saisonnière :

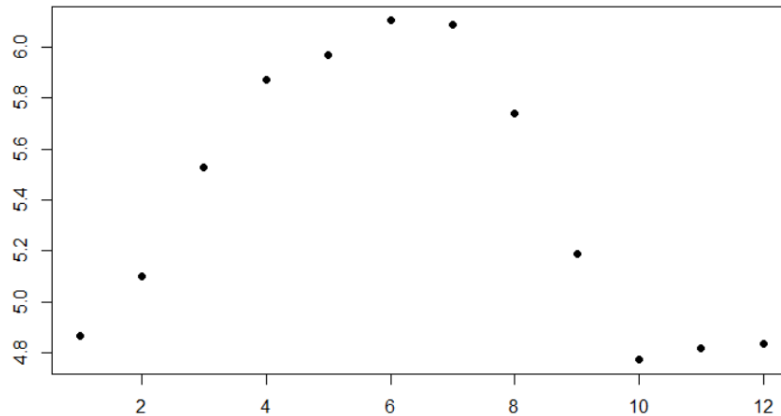


FIGURE 9.10 – Évolution de la variance du résidu en fonction du mois

La simple présence d'une saisonnalité dans la variance des résidus n'est pas quelque chose de grave car elle n'évolue pas sur le long terme. Face à ce scénario, il existe un moyen simple de rendre le résidu homoscédastique. Pour cela, il suffit de modéliser ω par le produit d'un mouvement brownien standard W avec la variance de périodicité annuelle $\sigma^{(m)}$.

Le modèle final deviendra donc le suivant :

$$R_t = (1 - \alpha^{(m)}) R_{t-1} + \beta^{(m)} P_{t-1} + \sigma^{(m)} W_t$$

C'est-à-dire :

$$T_t = M_t + \Theta_t + (1 - \alpha^{(m)}) (T_{t-1} - M_{t-1} - \Theta_{t-1}) + \beta^{(m)} P_{t-1} + \sigma^{(m)} W_t$$

avec :

- M la tendance des températures
- Θ la composante saisonnière des températures
- P la série du volume des précipitations
- W bruit blanc gaussien

Chapitre 10

Construction du KBDI

10.1 Simulations

Maintenant que nous avons validé notre générateur de pluie et de températures, nous pouvons simuler chacune de ces deux variables dans le temps et ainsi projeter l'indice KBDI par la méthode de calcul précédemment exposée.

La simulation suivante équivaut à vingt simulations de l'année 2021. Cependant, pour plus de lisibilité, nous affichons ces vingt années de simulation comme projection jusqu'à 2040 :

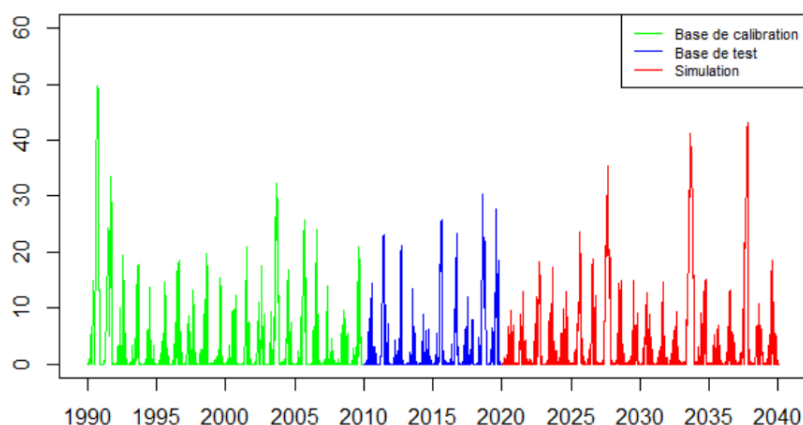


FIGURE 10.1 – KBDI simulé pour la ville de Paris

Nous pouvons voir sur ce graphique que l'amplitude du KBDI simulé est cohérente avec l'amplitude historique. Notons que la cohérence entre l'indice projeté et celui calculé sur la base de calibration n'était pas particulièrement prévisible. Le KBDI n'a pas directement été paramétré sur les données de la base de calibration. Seuls les paramètres du modèle de pluie et température contiennent l'information présente dans la base de calibration. Ci-dessous, nous pouvons faire un constat similaire pour les villes de Lyon, Rennes et Toulouse :

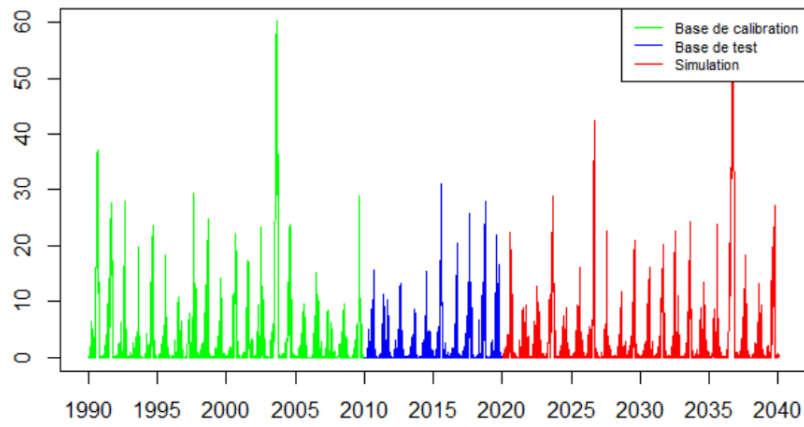


FIGURE 10.2 – KBDI simulé pour la ville de Lyon

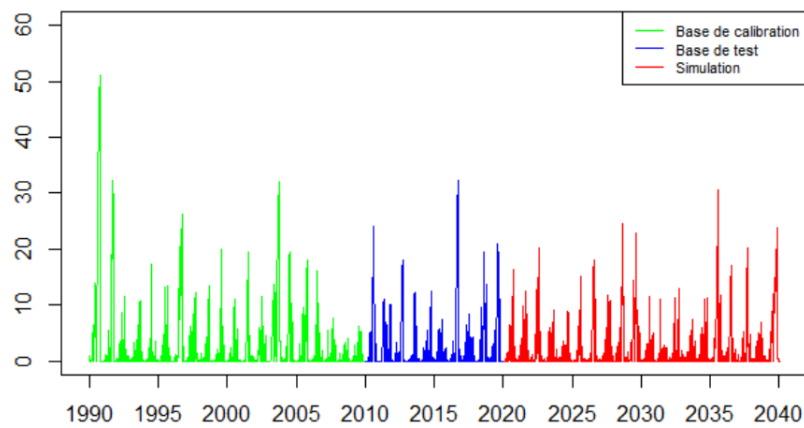


FIGURE 10.3 – KBDI simulé pour la ville de Rennes

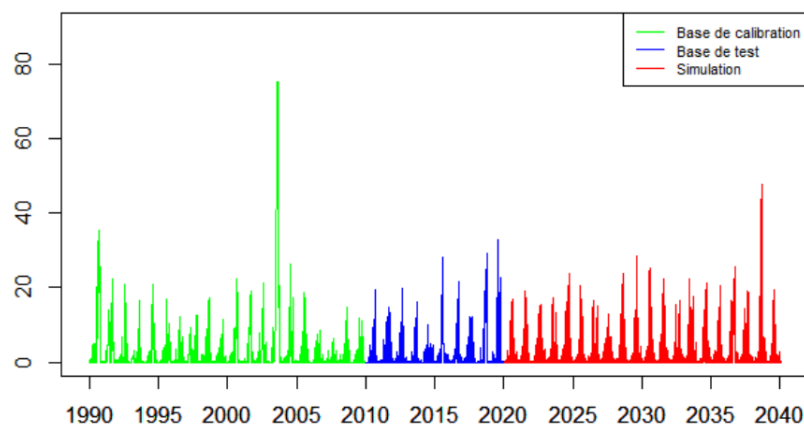


FIGURE 10.4 – KBDI simulé pour la ville de Toulouse

10.2 Cohérence du modèle

Nous allons maintenant étudier la cohérence de notre simulation avec les données historiques et par la suite nous intéresser exclusivement au pic annuel de sécheresse dont l'amplitude varie chaque été. Pour cela, nous allons voir si la distribution des KBDI annuels maximaux simulés sur mille années 2021 est cohérente avec la distribution des KBDI annuels maximaux mesurés de 1990 à 2020.

Le graphique suivant illustre la bonne cohérence entre la simulation et les valeurs historiques :

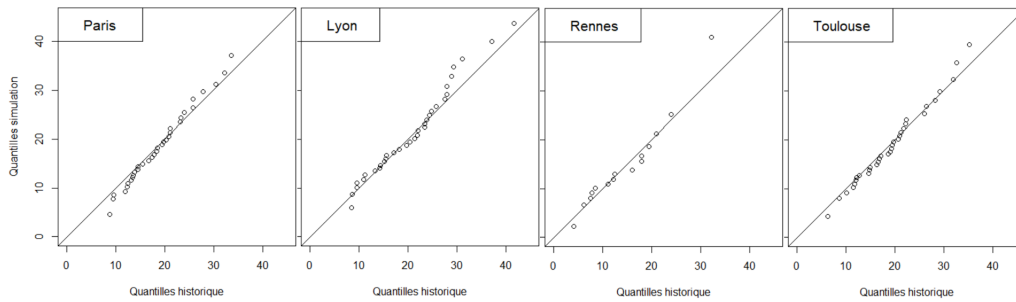


FIGURE 10.5 – Quantiles des KBDI annuels maximaux simulés en fonction des quantiles des KBDI annuels maximaux historiques pour Paris, Lyon, Rennes et Toulouse

Nous allons maintenant formuler l'hypothèse H_0 selon laquelle les KBDI annuels maximaux simulés suivent la même loi que les KBDI annuels maximaux historiques et allons tester cette hypothèse avec un test de Kolmogorov-Smirnov et un niveau de confiance de 95%. Les résultats sont synthétisés dans le tableau suivant :

Ville	p.value du test	Présomption contre H_0
Paris	0,8331	pas de présomption
Lyon	0,9535	pas de présomption
Rennes	0,2061	pas de présomption
Toulouse	0,9431	pas de présomption

À la vue de ces résultats, notre simulation par générateur climatique semble bien cohérente avec les données historiques.

Cinquième partie

Application

Chapitre 11

Calcul de la prime pure de contrats basés sur le KBDI

11.1 Couverture considérée

Définition du sous-jacent

Nous avons vu que l'évolution du KBDI est périodique. L'indice atteint généralement son maximum chaque été avant de tomber proche de zéro en hiver. Nous avons donc décidé de construire des options ayant pour sous-jacent la valeur maximale atteinte par le KBDI dans l'année en cours.

Étant donné que le risque est à priori corrélé positivement avec le niveau de sécheresse, nous considérerons ici qu'options d'achats et écarts haussiers.

Possibilité d'exercice

Rappelons que les options européennes peuvent uniquement être exercées à leur date d'échéance. Ce type d'exercice n'est pas optimal car si la date d'échéance est le premier janvier et que le KBDI atteint son maximum en août, alors le détenteur de l'option devra attendre 4 mois avant d'encaisser le payoff du contrat. Cela pourrait poser problème si le détenteur a acheté ce contrat pour compenser une perte financière causée par la sécheresse et qu'il souhaite une indemnisation rapide.

Pour cette raison nous choisissons de tarifier des options américaines car elles peuvent être exercées à n'importe quelle date, jusqu'à l'échéance. Ainsi, si son détenteur a besoin de récupérer rapidement le payoff du contrat, suite à une sécheresse, il pourra décider de l'exercer directement. Notons qu'en réalité la mise à jour des données de pluie et de température dans la base de données POWER est faite avec 2 à 3 jours de latence. Ce délai reste cependant très court comparé au délai d'indemnisation pouvant dépasser plusieurs mois avec l'assurance traditionnelle. Pour simplifier la compréhension, chaque contrat considéré par la suite aura donc une date d'effet au 1^{er} janvier 2021 et une date d'échéance au 31 décembre 2021.

Seuils déclenchement considérés

Maintenant que nous avons fixé la structure du dérivé, il reste à définir les seuils de déclenchement que nous souhaitons étudier ici. Pour cela, nous avons défini trois seuils :

- Un premier seuil fixé à 20. Il correspond à un niveau de sécheresse faible observé à Paris une fois tous les trois ans.
- Un deuxième seuil fixé à 25. Il correspond à un niveau de sécheresse intermédiaire observé à Paris une fois tous les cinq ans.
- Un dernier seuil fixé à 30. Il correspond à un niveau de sécheresse extrême observé à Paris une fois tous les dix ans.

Pour chacune des méthodes précédemment présentées ; Le Burning Cost, la modélisation du sous-jacent et le générateur climatique, nous avons calculé la prime pure d'une option d'achat américaine de seuils 20, 25 et 30 pour la ville de Paris.

Afin d'étudier l'influence qu'a le choix de la base de calibration sur le tarif, nous avons décidé de tarifier ces produits sur 3 bases de calibration différentes :

- Base 1 : de 1990 à 2010
- Base 2 : de 1995 à 2015
- Base 3 : de 2000 à 2020

Nous définirons la prime moyenne calculée comme la moyenne des primes sur ces trois bases. Nous ferons de même pour la variance qui sera exprimée relativement à la prime moyenne. Ceci facilitera leur comparaison.

11.2 Méthode Burning Cost

Regardons tout d'abord les résultats obtenus par la méthode Burning Cost :

Seuil	Prime pure moyenne	Variance relative
20	2,51	40 %
25	1,10	67 %
30	0,50	88 %

Nous constatons que la variance de l'estimation est relativement élevée. Cela confirme que cette méthode est très sensible à l'historique choisi pour calculer la prime pure. Il est également normal que la variance relative augmente avec le seuil de déclenchement car plus ce seuil est élevé, plus le nombre d'années ayant connu un KBDI supérieur à ce seuil est faible. La très forte variance s'explique par le fait que sur la première base de 1990 à 2010, nous répertorions les deux grandes sécheresses de 1990 et de 2003. Celles-ci ont donc été prises en compte dans le calcul de la prime pure. En effet, sur cette base, la prime calculée pour le seuil 20 atteignait 3,35.

Une autre limite de cette méthode d'évaluation est de pouvoir obtenir une prime nulle si le seuil est plus élevé que le KBDI maximal observé sur l'ensemble de la base servant à calculer la prime pure.

11.3 Modélisation du sous-jacent

Nous souhaitons maintenant tarifier ces options avec la méthode de la modélisation du sous-jacent. Pour cela, nous devons trouver une loi qui représente de manière cohérente la distribution du KBDI annuel maximal. Nous émettons alors l'hypothèse H_0 selon laquelle le KBDI annuel maximal suit une loi normale. Ci-dessous nous pouvons voir le résultat du test de l'hypothèse H_0 pour chacune de nos quatre villes :

Ville	Valeur-p du test	Présomption contre H_0
Paris	0,72	pas de présomption
Lyon	0,71	pas de présomption
Rennes	0,43	pas de présomption
Toulouse	0.07	faible présomption

À la vue du résultat de ce test, il n'est pas déraisonnable de modéliser le KBDI maximal annuel par une loi normale.

Après simulation d'un million d'année de KBDI maximal, nous obtenons les résultats suivants :

Seuil	Prime pure moyenne	Variance relative
20	2,65	64 %
25	1,02	60 %
30	0,33	39 %

Nous remarquons que les résultats sont très proches de ceux obtenus par la méthode du Burning Cost. Cela était prévisible car la distribution de la loi normale est, par construction, très proche de la distribution empirique du KBDI maximal. La variance dans la tarification reste toujours élevée avec cette méthode. Cependant, nous remarquons que la modélisation du sous-jacent est à privilégier pour tarifier les options ayant un seuil de déclenchement élevé.

11.4 Générateur climatique

Enfin, la troisième méthode consiste à calculer la prime pure d'une option climatique par le générateur climatique que nous venons de construire.

Après génération de dix mille années 2021 de KBDI, nous obtenons les résultats suivants :

Seuil	Prime pure moyenne	Variance relative
20	2,63	18 %
25	0,99	10 %
30	0,39	5 %

Nous pouvons tout de suite constater que la variance relative est ici beaucoup plus faible qu'avec les deux méthodes précédentes. Ceci est vrai pour l'ensemble des trois seuils considérés. Ce résultat prouve que notre générateur est beaucoup moins sensible

au choix de la base de calibration que le sont les autres méthodes. En effet, pour l'ensemble des trois seuils, la variance de cette méthode est environ cinq fois plus faible que celle de la méthode de la modélisation du sous-jacent.

11.5 Tarification de l'écart haussier

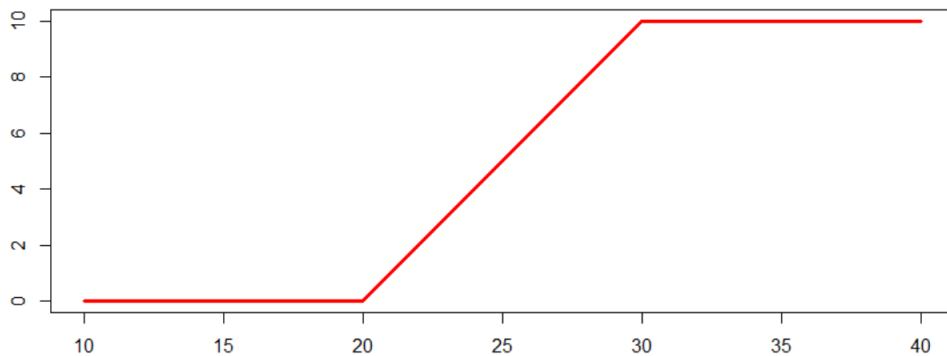
En pratique, l'écart haussier sera privilégié à l'option d'achat aussi bien par l'assureur que par l'assuré. L'assureur préférera que le paiement d'un contrat soit borné afin de ne pas être exposé aux sécheresses les plus extrêmes. L'assuré, particulier ou professionnel, privilégiera également ce type de couverture car au delà d'un seuil de sécheresse extrême, le sinistre aura atteint 100% et n'évoluera plus.

Ainsi, nous considérons par la suite des contrats standardisés sous forme d'écart haussiers de seuil d'exercice 20 et de seuil plafond 30, dont le payoff s'exprime :

$$P = \min \left(\max(KBDIMAX - 20, 0), 10 \right)$$

Où $KBDIMAX$ est le KBDI maximal atteint dans l'année en cours.

La représentation graphique du payoff en fonction de $KBDIMAX$ est la suivante :



En prenant pour base de calibration l'ensemble de notre historique et après génération de dix mille années 2021 de KBDI, nous obtenons les primes suivantes :

Ville	Prime
Paris	2,31
Lyon	3,38
Rennes	0,81
Toulouse	2,34

Par la suite nous noterons respectivement ces primes P_{Paris} , P_{Lyon} , P_{Rennes} et $P_{Toulouse}$

Chapitre 12

Agrégation du risque

12.1 Méthodologie

Si un assureur commercialise des contrats d'assurance paramétrique basés sur le KBDI en France, son portefeuille ne sera pas bien diversifié compte tenu de la corrélation spatiale de la sécheresse.

Le tableau suivant illustre la forte corrélation entre les KBDI annuels maximaux pour l'ensemble des villes étudiées :

Corrélation	Paris	Lyon	Rennes	Toulouse
Paris	1	0.34	0.42	0.49
Lyon	0.34	1	0.14	0.59
Rennes	0.42	0.14	1	0.42
Toulouse	0.49	0.59	0.42	1

Il n'est donc sûrement pas intéressant pour un assureur de commercialiser un portefeuille de dérivés portant sur le KBDI sans qu'il transfère tout ou partie de ce risque. Pour cela, nous proposons de construire un indice agrégé qui prendra en compte l'exposition de l'ensemble du portefeuille et permettra de transférer ce risque à un autre parti.

Nous supposons que l'assureur souhaite transférer l'intégralité de son exposition issue de la commercialisation de contrats d'assurance paramétrique basés sur l'indice KBDI à Paris, Lyon, Rennes et Toulouse.

Nous conservons également l'hypothèse selon laquelle les contrats sont standardisés sous forme d'écart de seuil d'exercice 20 et de seuil plafond 30. Nous rappelons que le premier seuil correspond à un niveau de sécheresse observé à Paris une fois tous les trois ans et le seuil plafond correspond à un niveau de sécheresse extrême observé à Paris une fois tous les dix ans.

L'exposition de l'assureur sur chacune de ces villes sera quand à elle supposée proportionnelle à sa surface :

Ville	Surface (km^2)	Exposition retenue
Paris	105	10/32
Lyon	48	5/32
Rennes	50	5/32
Toulouse	118	12/32

12.2 Construction de l'indice agrégé et calcul de la prime pure

Posons I_{AG} l'indice agrégé du portefeuille. Son calcul est réalisé de la manière suivante :

$$I_{AG} = \frac{10 * I_{Paris} + 5 * I_{Lyon} + 5 * I_{Rennes} + 12 * I_{Toulouse}}{32}$$

Avec :

$$I_{Paris} = \min \left(\max(KBDIMAX_{Paris} - 20, 0), 10 \right)$$

$$I_{Lyon} = \min \left(\max(KBDIMAX_{Lyon} - 20, 0), 10 \right)$$

$$I_{Rennes} = \min \left(\max(KBDIMAX_{Rennes} - 20, 0), 10 \right)$$

$$I_{Toulouse} = \min \left(\max(KBDIMAX_{Toulouse} - 20, 0), 10 \right)$$

Où $KBDIMAX$ est le KBDI maximal atteint dans l'année en cours.

Calculons maintenant la prime pure P_{AG} du produit ayant pour payoff l'indice agrégé I_{AG} . La linéarité de l'espérance permet de calculer cette prime comme fonction linéaire des primes pures P_{Paris} , P_{Lyon} , P_{Rennes} et $P_{Toulouse}$:

$$P_{AG} = \frac{10 * P_{Paris} + 5 * P_{Lyon} + 5 * P_{Rennes} + 12 * P_{Toulouse}}{32}$$

Donc :

$$P_{AG} = 2,25$$

12.3 Étude de la mutualisation du risque

Après avoir calculé la prime pure de chacun des contrats considérés, c'est à dire l'espérance de son indice sous-jacent, nous allons maintenant étudier l'impact de cette agrégation sur la variance de l'indice.

Calculons dans un premier temps la variance des indices I_{Paris} , I_{Lyon} , I_{Rennes} et $I_{Toulouse}$ en prenant pour base les valeurs historiques de l'indice sécheresse :

Ville	Variance
Paris	12,42
Lyon	15,06
Rennes	13,52
Toulouse	14,04

Pour calculer la variance de l'indice agrégé , il faut prendre en compte l'interaction spatiale de l'indice des quatre villes considérées. Pour cela nous calculons les covariances de payoff en prenant également pour base les valeurs historiques de l'indice sécheresse. Les résultats sont synthétisés dans le tableau de covariance suivant :

Covariance	Paris	Lyon	Rennes	Toulouse
Paris	12,42	5,97	6,32	8,41
Lyon	5,97	15,06	5,38	9,08
Rennes	6,32	5,38	13,52	6,28
Toulouse	8,41	9,08	6,28	14,04

Nous pouvons alors calculer la variance de l'indice agrégé grâce à la formule suivante :

$$V_{AG} = \frac{1}{32^2} * V(10 * I_{Paris} + 5 * I_{Lyon} + 5 * I_{Rennes} + 12 * I_{Toulouse})$$

C'est-à-dire :

$$V_{AG} = \frac{10^2}{32^2} V(I_{Paris}) + \frac{5^2}{32^2} V(I_{Lyon}) + \frac{5^2}{32^2} V(I_{Rennes}) + \frac{12^2}{32^2} V(I_{Toulouse}) +$$

$$\frac{10 * 5}{32^2} Cov(I_{Paris}, I_{Lyon}) + \frac{10 * 5}{32^2} Cov(I_{Paris}, I_{Rennes}) + \frac{10 * 12}{32^2} Cov(I_{Paris}, I_{Toulouse}) +$$

$$\frac{5 * 5}{32^2} Cov(I_{Lyon}, I_{Rennes}) + \frac{5 * 12}{32^2} Cov(I_{Lyon}, I_{Toulouse}) + \frac{5 * 12}{32^2} Cov(I_{Rennes}, I_{Toulouse})$$

Ainsi, la variance de l'indice agrégé vaut :

$$V_{AG} = 9,12$$

Le tableau suivant compare la variance et l'écart type réels de I_{AG} à la variance et l'écart type de I_{AG} sous l'hypothèse d'indépendance des indices I_{Paris} , I_{Lyon} , I_{Rennes} et $I_{Toulouse}$, c'est à dire d'indépendance du risque de sécheresse dans ces quatre villes.

	Variance	Écart type
Calcul avec interaction entre les indices	9,12	3,02
Calcul sous hypothèse d'indépendance	3,89	1,97
Écart relatif	+135%	+53%

Nous retrouvons bien ici la concentration du risque lié à la corrélation positive entre les indices de sécheresse calculés à travers la France. L'écart type réellement calculé est ainsi 53% plus important que si les risques à Paris, Lyon, Rennes et Toulouse avaient été indépendants.

Avant d'envisager la commercialisation de contrats d'assurance paramétrique basés sur le KBDI, l'assureur devra donc veiller à ce que les primes commerciales des contrats paramétriques vendus soient suffisamment élevées pour compenser la prime de risque demandée par l'entité auquel le risque agrégé sera transmis et qui s'exposera à une variance importante.

Conclusion

Tout d'abord, nous avons décrit les enjeux du risque sécheresse et ses nombreux impacts financiers et sociétaux. Nous avons montré que ce péril engendre une sinistralité conséquente en France, notamment par le phénomène de sécheresse-réhydratation des sols argileux qui a coûté plusieurs milliards d'euros aux assureurs en 2003. Conscients du besoin de transférer ce risque, nous avons étudié le marché des dérivés climatiques en nous appuyant sur l'expertise de Bernard FINAS, fondateur de Jola et directeur de The Weather Option ainsi que de David WHITEHEAD, co-directeur de weatherX-change. L'échange que nous avons eu nous a conforté dans l'idée que le marché des dérivés climatiques est aujourd'hui en pleine expansion.

Nous avons alors sélectionné la base de données POWER de la NASA pour disposer d'informations nécessaires à l'évaluation d'options climatiques. Ce choix a tout autant été motivé par la pérennité de la base que par son homogénéité et sa simplicité d'exploitation. De plus, son interface étant une API libre d'accès, les travaux réalisés dans ce mémoire pourront être repris et développés librement dans le cadre de travaux futurs. Cependant la base de données présente certaines limites que nous avons pris soin d'énoncer. La mise à jour des données de pluie et de température dans la base de données POWER est faite avec 2 à 3 jours de latence et ce délai pourrait être jugé trop long pour un assuré attendant une indemnisation instantanée de la part de son contrat d'assurance paramétrique. Cette base de données présente également l'inconvénient de fournir exclusivement des données selon une grille GPS de résolution $(\frac{1}{2})^\circ \times (\frac{1}{2})^\circ$

Le cœur technique de ce mémoire est la construction d'un générateur climatique stochastique. Celui-ci a pour objectif d'évaluer des options basées sur l'indice sécheresse de Keetch-Byram, calculable à partir de données de températures et de précipitations journalières présentes dans notre base de données. Nous avons d'abord modélisé la pluie à l'aide d'une chaîne de Markov à deux états correspondant aux jours pluvieux et aux jours secs. Le volume des précipitations a été modélisé par une loi gamma dont les paramètres sont conditionnés selon le mois. Cette granularité a mis en évidence le caractère saisonnier des précipitations en France. Parallèlement à cela, nous avons modélisé la température maximale journalière par une série temporelle combinée à un processus autorégressif de retour à la moyenne. Afin de tenir compte de la corrélation entre la pluie et la température, nous avons proposé l'ajout d'un terme d'interaction entre ces deux grandeurs au modèle préexistant.

Une fois le générateur construit, nous avons pu tarifier des options sécheresse à partir de la projection de l'indice sur dix mille années. Nous avons vu que le principal avantage de la tarification par un générateur climatique est que la prime pure estimée est très peu sensible au choix de la base d'entraînement, alors que les méthodes plus classiques du Burning Cost et de la modélisation du sous-jacent sont grandement influencées par la présence des sécheresses de 1990 et 2003.

Quelques axes d'amélioration peuvent cependant être envisagés pour améliorer la performance de notre générateur. En effet, celui-ci semble avoir du mal à générer les années de sécheresses les plus extrêmes comme celle mesurée en 2003 à Toulouse. Il faut donc prendre garde à ne pas tarifier des options à un seuil trop élevé par cette méthode et privilégier la tarification d'écarts verticaux haussiers à celle de calls. À ce titre, il serait également intéressant de voir si l'ajout d'un nouvel état dans le générateur de jours pluvieux permet de capter de manière plus fine ces périodes de sécheresses les plus extrêmes.

Enfin, si le lecteur souhaite utiliser le fruit de nos travaux pour concevoir un produit d'assurance paramétrique, nous l'invitions à consulter le mémoire de Yann DISERBEAU qui propose une série de méthodes permettant la calibration rapide d'un indice climatique corrélé au risque assuré [DISERBEAU, 2019]. Les accélérateurs qu'il propose dans son étude sont en parfaite complémentarité avec notre démarche de tarification de dérivés climatiques.

Bibliographie

[AFP et FFA, 2019]

<https://www.linfodurable.fr/la-secheresse-risque-dont-le-cout-va-croitre-le-plus-vite-ces-prochaines-annees-federation-de-12968>

[ALEXANDER, 1990]

Martin E. Alexander (1990) Computer Calculation of the Keetch-Byram Drought Index-Programmers Beware!

[Bustinza, 2010]

Bustinza, R., O. Tairou, F., Gosselin, P., Bélanger, D. (2010). Proposition d'indicateurs aux fins de vigie et de surveillance des troubles de la santé liés à la foudre et aux incendies de forêt.

[CCME, 2003]

Le Conseil canadien des ministres de l'environnement. (2003). Le climat, la nature, les gens : Indicateurs d'évolution du climat au Canada.

[CHABANOL, 2013]

CHABANOL, J-J.R-M-L. (2013). Chaînes de Markov

[CIEAU]

<https://www.cieau.com/connaitre-leau/secheresse-comment-agir>

[CÔME, 2018]

CÔME, T. (2018). Tarification d'une assurance indicielle pour les producteurs de maïs au Mali.

[CONSIDINE]

Geoffrey Considine, Ph.D., Weather Derivatives Group, Aquila Energy. Introduction to Weather Derivative.

[DAI, 2010]

DAI, A. (2010). Drought under global warming : a review. Wiley Interdisciplinary Reviews : Climate Change, 2(1), p 45-65.

[DISERBEAU, 2019]

DISERBEAU, Y. (2019). Une proposition d'accélérateurs pour la mise en place d'une couverture indicielle des risques météo-sensibles.

- [DOLLING et al., 2009]
Dolling, K., Chu, P.-S., Fujioka, F. (2009). Natural variability of the Keetch - Byram Drought Index in the Hawaiian Islands. *International Journal of Wildland Fire*, 18(4), p 459.
- [DUTANG et al., 2017]
Christophe, D, Mathieu, B, Michaël, S, et Romain, B. (2017, février). Les dérivés climatiques et énergétiques.
- [FALLOT et SAUTEBIN, 2018]
FALLOT, J. M., SAUTEBIN, T. (2018). Evaluation des différentes lois statistiques pour l'estimation des précipitations journalières extrêmes en Suisse.
- [FEBRIANI et al., 2010]
Febriani, Y., Levallois, P., Gingras, S., Gosselin, P., Majowicz, S. E., Fleury, M. D. (2010). The association between farming activities, precipitation, and the risk of acute gastrointestinal illness in rural municipalities of Quebec, Canada : a cross-sectional study.
- [Figaro, 2019]
<https://www.lefigaro.fr/conjoncture/la-secheresse-coute-entre-700-et-900-millions-aux-assureurs-chaque-annee-20190725>
- [Gabriel and Neumann, 1962]
Gabriel, K.R. and Neumann, J. 1962 : A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of the Royal Meteorological Society* 88, p 90–95.
- [GUILLERME, 2009]
GUILLERME, P. Y. (2009). Dérivés Climatiques Evaluation d'options.
- [HINZ, 2007]
Hinz, J. (2007). Weather Derivative Valuation : The Meteorological, Statistical, Financial and Mathematical Foundations. *Journal of the American Statistical Association*, 102(477), p 1.
- [INSPQ]
Institut national de santé publique du Québec
<http://www.monclimatmasante.qc.ca>
- [Katz, 1977]
Katz, R.W. 1977 : Precipitation as a chaindependent process. *Journal of Applied Meteorology* 16, p 671–676.
- [KEETCH et BYRAM, 1968]
J. Keetch, J et M. Byram, G. (1968). A Drought Index for Forest Fire Control.
- [LLOYD et al., 2007]
LLOYD et al. (2007). Global diarrhoea morbidity, weather and climate. *Climate Research*, 34, p 119-127.

[Longley, 1953]
Longley, R.W. (1953) : The length of dry and wet periods. Quarterly Journal of the Royal Meteorological Society 79, p 520–527.

[MORENO et ROUSTANT, 2002]
MORENO, M., Roustant, O. (Éds.). (2002). Modélisation de la température Application aux dérivés climatiques.

[NASA GEOS, 2016]
<https://gmao.gsfc.nasa.gov/news/geos>

[NASA MERRA, 2017]
<https://gmao.gsfc.nasa.gov/reanalysis/MERRA>

[OMM et GWP, 2016]
Organisation météorologique mondiale (OMM) et Partenariat mondial pour l'eau (GWP), 2016 : Manuel des indicateurs et indices de sécheresse (M. Svoboda et B.A. Fuchs), Programme de gestion intégrée des sécheresses, Integrated Drought Management Programme Tools and Guidelines Series 2, Genève.

[SCHULTE, 2016]
SCHULTE, J. F. (2016). Modélisation du risque subsidence en France métropolitaine.

[Sénat]
<https://www.senat.fr/rap/r09-039/r09-0392.html>

[SHEA, 2014]
SHEA, D. (2014). Regridding overview
<https://climatedataguide.ucar.edu/climate-data-tools-and-analysis/regridding-overview>.

[SPARKS, 2020]
SPARKS, A. H. : R Package 'nasapower'. 2020.
<https://cran.r-project.org/web/packages/nasapower/nasapower.pdf>

[STACKHOUSE, 2019]
STACKHOUSE, P. (2019). Heating/Cooling Degree Days.
<https://power.larc.nasa.gov/docs/methodology/meteorology/degree-days/>

[STACKHOUSE, 2020]
STACKHOUSE, P. (2020). Précipitation.
<https://power.larc.nasa.gov/docs/methodology/meteorology/precipitation>

[STANKE et al., 2013]
STANKE et al.(2013). Health Effects of Drought : a Systematic Review of the Evidence.

[Thom, 1958]
Thom, H.C.S. (1958). A note on the gamma distribution. Monthly Weather Review 86, p 117–122.

[WHITE et al., 2008]

A, J. W. W. (2008). Evaluation of NASA satellite- and assimilation model-derived long-term daily temperature data over the continental US. *Annalen der Physik*. 2008, 148, p 1574-1587.

[Wilks et Wilby, 1999]

Wilks, D.S. et Wilby, R.L. (1999). The weather generation game : a review of stochastic weather models

[Williams, 1952]

Williams, C.B. (1952). Sequences of wet and of dry days considered in relation to the logarithmic series. *Quarterly Journal of the Royal Meteorological Society*, 78, p 91-96.

[WxRiskGlobal]

Weather Derivative Solutions

<https://www.wxriskglobal.com>

Annexe

Formations argileuses en France

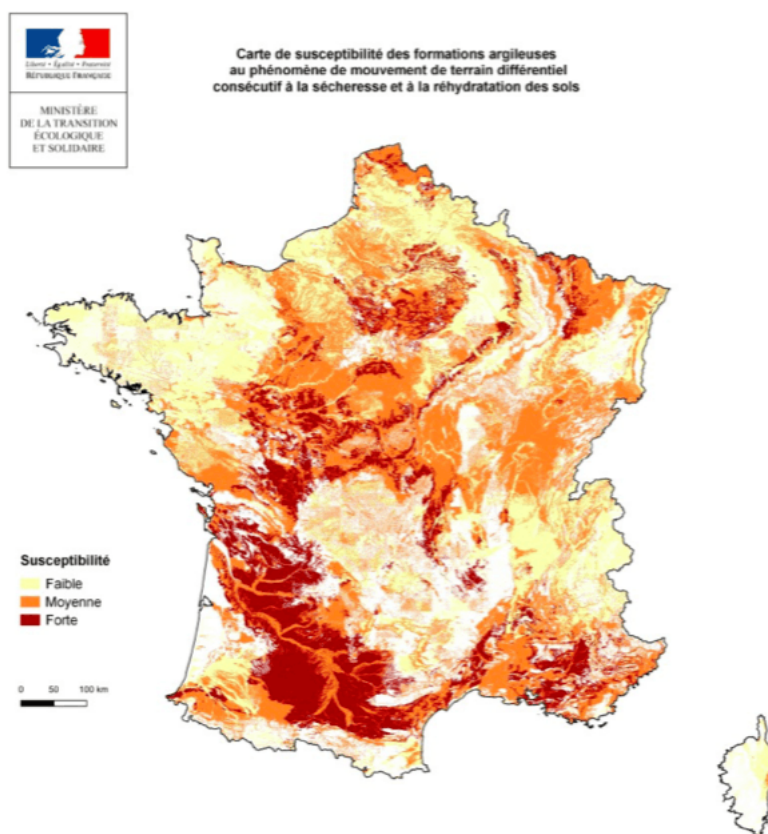


FIGURE 12.1 – Carte de susceptibilité des formations argileuses en France

Méthode d'estimation par noyau

La densité de X , dont les réalisations observées sont x_1, x_2, \dots, x_n , peut être estimée par la fonction g définie par :

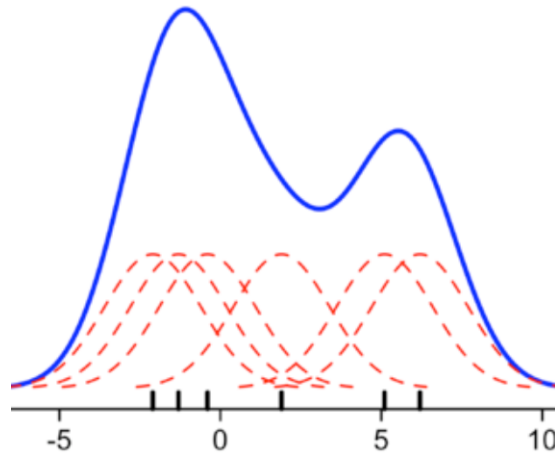
$$g = \frac{1}{nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right)$$

Le noyau K peut par exemple être choisi gaussien :

$$\forall x \in \mathbb{R}, \quad K(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

h est quand à lui un paramètre de lissage qui doit être soigneusement choisi en fonction du nombre d'observations. Pour une valeur trop grande, le lissage serait trop important tandis qu'une valeur trop faible impliquerai un niveau de détail de la densité non justifié.

le graphique suivant illustre cette méthode pour six observations et pour un noyau gaussien :



Nous retrouvons ainsi en noir les observations, en rouge l'application du noyau gaussien centré sur chacune de ces observations et en bleu la densité estimée.

Processus stationnaire au sens strict :

Le processus X est stationnaire au sens strict si :

$\forall n \in \mathbb{N}$, $\forall h \in \mathbb{Z}$ et $\forall t_1 < t_2 < \dots < t_n$ le vecteur $(X_{t_1+h}, \dots, X_{t_n+h})$ est égal en loi au vecteur $(X_{t_1}, \dots, X_{t_n})$

Compléments sur les mesures satellitaires :

Voici un extrait des explications techniques des relevés météorologique [STACKHOUSE, 2020] :

"Le modèle cherche à assimiler et à optimiser les données d'observation et les estimations du modèle des variables atmosphériques. Les types d'observations utilisées dans l'analyse comprennent les observations de la surface terrestre de la pression de surface ; les observations à la surface de l'océan de la pression au niveau de la mer et des vents ; les vents au niveau de la mer déduits des retours de rétrodiffusion des radars spatiaux ; les données conventionnelles en altitude provenant de sources brutes (par exemple, hauteur, température, vent et humidité) ; les sources supplémentaires de données en altitude comprennent les sondes de largage, les ballons pilotes et les vents des avions ; et des informations détectées à distance à partir de satellites (par exemple, profils de hauteur et d'humidité, eau totale perceptible et vents de vecteur de mouvement de nuage à un seul niveau obtenus à partir d'images de satellites géostationnaires). Les estimations mondiales horaires de la distribution verticale d'une gamme de paramètres atmosphériques émergent de MERRA-2."