

ESSEC

ISUP

Promotion 2003

Mémoire présenté devant

**l'Institut de statistique
de l'Université Pierre et Marie Curie**

et

l'Ecole Supérieure des Sciences Economiques et Commerciales

pour l'obtention du

Diplôme de l'Institut des Actuaires Français

Assurance Finance

Par **M. Pierre VENDÉ**

Sujet : **Les Couvertures Indicielles en Réassurance Cat :
Prise en compte de la dépendance spatiale dans la tarification**

Lieu du stage : Benfield Paris

Responsable du stage : Emmanuel DUBREUIL

Invité(s) :

CONFIDENTIEL

Remerciements

Je tiens à remercier tout particulièrement M. Emmanuel Dubreuil pour son attention, sa patience, et ses conseils avisés. Je remercie aussi M. Philippe Renault de m'avoir donné l'occasion de réaliser ce mémoire au sein de Benfield Paris.

Je souhaite également remercier M. Jean Berthon pour le soutien qu'il m'a apporté.

J'ajoute un vif remerciement à M. Jean Pierre Indjehagopian, qui m'a fait découvrir les études d'actuariat.

Je remercie également la société CatiXL d'avoir accepté de mettre à ma disposition les données nécessaires à la réalisation de ce mémoire.

Enfin, j'adresse mes derniers remerciements à mes collègues de Benfield Paris, M. Yves Ledanvic en particulier, pour leur soutien durant ces mois de travail.

Promotion 2003**M. Pierre VENDÉ**

Lieu du stage : Benfield Paris

**Les Couvertures Indicielles en Réassurance Cat :
Prise en compte de la dépendance spatiale dans la tarification****RESUME**

Ces dernières années ont vu le développement d'un nouveau marché, celui des dérivés climatiques. Un produit dérivé « traditionnel » est un contrat financier (ou un contrat d'assurance) dépendant d'un autre actif (action, taux d'intérêt, matière première, etc.), appelé actif sous-jacent. L'actif sous-jacent d'un produit dérivé climatique est un indice construit à partir de données météorologiques, comme par exemple les précipitations (pluie, neige), le vent ou l'ensoleillement. Toutefois, le marché actuel comporte essentiellement des produits liés à la température, permettant à des entreprises de couvrir une perte de marge ou une perte de chiffre d'affaires consécutives à un hiver trop doux ou un été trop frais.

La couverture indiciaire Cat est une forme de dérivé climatique spécialement adaptée à la réassurance des événements catastrophiques (tempête, vague de froid, inondation, etc.). A partir de données climatiques enregistrées dans un panel de stations météorologiques, la construction de l'indice permet de capturer les variations extrêmes de la météo. La couverture indiciaire offre alors la garantie d'une éventuelle indemnisation, en fonction de la valeur de l'indice à maturité. Comment effectuer la cotation d'une couverture indiciaire ? A travers l'exemple d'une couverture contre le risque de tempête en France basée sur un indice de vitesses de vent, un ensemble d'approches envisageables en vue de tarifier une couverture indiciaire Cat sera présenté.

Plus intéressantes que les tarifications fondées sur les valeurs historiques de l'indice, les tarifications obtenues à partir des variables climatiques elles-mêmes posent le problème de la dépendance entre les stations. Au moyen d'outils de mesure et de modélisation de la dépendance, en particulier les copules, une tarification affinée de la couverture sera proposée.

MOTS-CLES

Réassurance, dérivé climatique, indice, dépendance, valeurs extrêmes, copules

Promotion 2003**M. Pierre VENDÉ**

Lieu du stage : Benfield Paris

**Pricing Cat Reinsurance Index covers :
Dealing with spatial dependency****ABSTRACT**

The last past years have seen the emergence of a new fascinating market : the weather derivatives market. A “traditional” derivative product is a financial contract (or an insurance contract) relying on another asset (share, interest rate, commodity...) called the underlying asset. The underlying asset of a weather derivative is an index based on weather data, such as precipitations (rain, snow), wind speed or sunshine. The market is today dominated by temperature-related products, which allow the purchasing companies to hedge the risk of a loss in margins or turnover due to a warmer than normal winter or to a cooler than normal summer.

The Cat index cover is a kind of weather derivative, which is especially designed to fit reinsurance needs against catastrophic events (windstorm, cold wave, flood...). Based on weather data recorded in a network of meteorological stations, the structure of the index can capture extreme weather events. The index cover may lead to a payoff, depending on the index value at maturity. What is the price of an index cover ? Through the example of a cover against windstorm in France based on a wind speed index, some pricing methods for Cat index cover are developed.

More interesting than index-based pricing methods, a pricing approach relying on weather data fitting introduces the problem of dependencies between stations. With the help of dependency modelling tools, such as copulas, an enhanced quotation of the index cover is studied.

KEY WORDS

Reinsurance, weather derivative, index, dependency, extreme values, copulas

Sommaire

INTRODUCTION	5
1 LA REASSURANCE	7
1.1 DEFINITIONS ET PRINCIPES GENERAUX DE LA REASSURANCE	7
1.1.1 Définition et but de la réassurance	7
1.1.2 Quelques données sur l'assurance et la réassurance.....	9
1.1.3 Les modes de réassurance.....	10
1.2 LES TYPES DE REASSURANCE, LES MECANISMES DES TRAITES DE REASSURANCE.....	12
1.2.1 La Réassurance proportionnelle	12
1.2.2 La Réassurance non proportionnelle	15
1.3 LE BESOIN DE REASSURANCE EN MATIERE DE RISQUES NATURELS	17
1.3.1 Définitions du risque naturel, des catastrophes naturelles	17
1.3.2 Risque de premier plan pour l'assurance	18
1.3.3 Un risque croissant	21
2 DE LA REASSURANCE CLASSIQUE A LA REASSURANCE NON TRADITIONNELLE DES RISQUES NATURELS	29
2.1 LA REASSURANCE TRADITIONNELLE DES RISQUES NATURELS	29
2.1.1 Les formes de réassurance traditionnelle des risques naturels.....	29
a. En France	29
b. A l'étranger	34
2.1.2 Les limites de la réassurance traditionnelle des risques naturels.....	35
a. De façon générale	35
b. Du point de vue de la cédante	38
c. Du point de vue du réassureur	38
2.2 LA REASSURANCE NON TRADITIONNELLE, LE TRANSFERT ALTERNATIF DES RISQUES NATURELS	39
2.2.1 Définition de la Réassurance non traditionnelle.....	39
2.2.2 Utilisation de la Réassurance non traditionnelle pour les risques naturels.....	40
2.3 LA TITRISATION DU RISQUE NATUREL, LES CAT BONDS	41
2.4 AUTRES SOLUTIONS ALTERNATIVES A LA REASSURANCE TRADITIONNELLE ...	45
2.5 LES INSTRUMENTS DERIVES CLIMATIQUES	46

3	PRINCIPES DE LA COUVERTURE INDICIELLE CAT.....	51
3.1	L'ORIGINE DE LA COUVERTURE INDICIELLE CAT.....	51
3.1.1	<i>Une vocation spécifique</i>	51
3.1.2	<i>Un produit de réassurance</i>	52
3.2	FONCTIONNEMENT DE LA COUVERTURE INDICIELLE CAT.....	53
3.2.1	<i>Utilisation de séries de données climatiques</i>	53
3.2.2	<i>L'indice : définition, utilisation</i>	55
a.	Choix du type de données.....	55
b.	Utilisation d'un panier de stations.....	58
c.	Choix des stations, répartition du portefeuille entre les stations.....	59
d.	Définition d'un indice par station.....	66
e.	Mécanisme de l'indice final, indemnisation due au contrat.....	67
3.3	ETUDE D'UN EXEMPLE.....	69
3.4	LIMITES ET DEBAT SUR LA COUVERTURE INDICIELLE CAT.....	81
4	PREMIERES APPROCHES DE TARIFICATION DE LA	
	COUVERTURE.....	83
4.1	HYPOTHESES SUR LA QUALITE DES SERIES DE DONNEES.....	83
4.1.1	<i>Qualité de mesure au cours du temps</i>	84
4.1.2	<i>Le problème des données manquantes</i>	84
4.1.3	<i>Conclusion</i>	85
4.2	TARIFICATIONS SUR LES VALEURS D'INDICE.....	86
4.2.1	<i>Burning Cost sur l'indice $S(T)$</i>	86
4.2.2	<i>Modélisation de l'indice $S(T)$ au premier dam/h</i>	89
4.2.3	<i>Modélisation de l'indice $S(T)$ avec introduction d'un seuil</i>	99
4.2.4	<i>Modélisation des indices $S_i(T)$</i>	112
4.3	TARIFICATIONS SUR LES VALEURS QUOTIDIENNES DE VITESSE DE VENT.....	113
4.3.1	<i>Présentation de la démarche</i>	113
4.3.2	<i>Ajustements au premier m/s</i>	116
4.3.3	<i>Ajustements avec introduction d'un seuil</i>	127
4.3.4	<i>Conclusions</i>	144
4.4	AUTRES APPROCHES ENVISAGEABLES.....	145
4.5	CONCLUSIONS.....	146

5 ANALYSE DE LA DEPENDANCE SPATIALE & TARIFICATION	
AFFINEE DE LA COUVERTURE.....	149
5.1 MISE EN EVIDENCE DU PROBLEME DE LA DEPENDANCE SPATIALE.....	149
5.1.1 Les caractéristiques de l'évènement climatique.....	149
5.1.2 La notion de dépendance.....	150
5.1.3 La dépendance entre les stations d'une couverture indicelle.....	151
5.2 OUTILS DE MESURE ET DE MODELISATION DE LA DEPENDANCE.....	156
5.2.1 Préliminaire.....	156
5.2.2 Le coefficient de corrélation linéaire.....	156
5.2.3 Le coefficient de corrélation de Spearman.....	157
5.2.4 Simulation des corrélations.....	158
5.2.5 Présentation des copules.....	159
5.2.6 Distribution conditionnelle et simulation des copules.....	161
5.2.7 Le coefficient de corrélation de Kendall.....	162
5.3 PRISE EN COMPTE DE LA DEPENDANCE DANS LA TARIFICATION :	
MODELISATION AU PREMIER M/S.....	163
5.3.1 Modélisation des corrélations.....	163
5.3.2 Modélisation par une Copule de Gumbel.....	167
5.3.3 Modélisation par une Copule de Student.....	175
5.3.4 Conclusions.....	180
5.4 PRISE EN COMPTE DE LA DEPENDANCE DANS LA TARIFICATION :	
MODELISATION AVEC SEUILS.....	181
5.4.1 Modélisation de la corrélation des fréquences.....	182
5.4.2 Modélisation de la dépendance Metz-Orly à l'aide de copules.....	184
5.4.3 Conclusions.....	188
5.5 INTRODUCTION AU PROBLEME DE DEPENDANCE TEMPORELLE.....	189
CONCLUSION.....	190
BIBLIOGRAPHIE.....	193
ANNEXES.....	197
ANNEXE 1 : TABLEAUX STATISTIQUES ET ROSE DES VENTS – STATION DE METZ.....	197
ANNEXE 2 : FICHES DESCRIPTIVES DES SERIES UTILISEES – CATIXL.....	20201
ANNEXE 3 : PRESENTATION DU LOGICIEL DE DFA REMETRICA.....	217
ANNEXE 4 : NUAGES DES VITESSES DE VENT ET QQ PLOTS.....	219
ANNEXE 5 : PROCEDURE DE CALCUL DES VALEURS Z_1 SOUS R.....	229

INTRODUCTION

L'objet de ce mémoire porte sur la tarification d'une forme relativement originale de protection de réassurance contre les événements naturels : la couverture indiciaire cat.

Après un rappel des principes généraux de la réassurance, nous présenterons le contexte actuel du marché de la réassurance des risques naturels. Nous constaterons alors l'émergence au cours de la dernière décennie de solutions novatrices de transfert du risque naturel, qui se posent en alternatives ou en compléments de la réassurance traditionnelle.

La couverture indiciaire Cat constitue l'une d'entre elles. Son fonctionnement repose sur le transfert du risque financier lié à la corrélation entre la sinistralité d'un portefeuille d'assurance et les variations extrêmes de la météo. Son mécanisme, comparable à celui des dérivés climatiques de température inventés par les acteurs du marché américain de l'énergie, s'appuie sur la création d'un indice à partir de données climatiques enregistrées dans un panel de stations météorologiques. La couverture indiciaire offre alors la garantie d'une indemnisation, en fonction de la valeur de l'indice à maturité.

Comment établir le prix d'une couverture indiciaire ? A travers l'exemple d'une couverture contre le risque de tempête en France basée sur un indice de vitesses de vent, nous présenterons un ensemble d'approches envisageables en vue de tarifier une couverture indiciaire Cat.

Nous nous apercevons que certaines de ces approches nécessitent de prendre en compte le phénomène de dépendance spatiale entre les stations météorologiques. Afin d'obtenir une tarification affinée, nous mettrons alors en pratique les outils de mesure et de modélisation de la dépendance entre des variables aléatoires, en utilisant notamment la théorie des copules.

1 LA REASSURANCE

Cette première partie vise à présenter succinctement les concepts et les mécanismes fondamentaux associés à la réassurance en général, et à la réassurance traditionnelle des risques naturels en particulier.

1.1 Définitions et principes généraux de la réassurance

1.1.1 Définition et but de la réassurance

« La réassurance est une opération par laquelle une société d'assurance, ou cédante, s'assure elle même auprès d'une autre société dénommée réassureur, ou cessionnaire, pour tout ou partie des risques qu'elle a pris en charge ».

La réassurance peut être décrite comme un **instrument de transfert de risques pour l'assureur**. Ce dernier est effectivement amené, délibérément ou pas, à souscrire des risques trop élevés en valeur ou en nombre pour qu'il puisse seul les assumer dans leur totalité. Afin de prévenir, de réduire le risque de ruine concomitant, les assureurs utilisent des méthodes visant à diviser ou à répartir leurs risques. La réassurance est ainsi, avec la coassurance¹, une méthode très répandue de partage des risques.

Le volume du risque, en valeur, que décide de conserver l'assureur pour son propre compte est appelé le « plein de conservation », ou rétention. Le montant total des engagements que l'assureur s'autorise alors à souscrire, en s'appuyant à la fois sur ses propres fonds et sur le renfort de la réassurance, est appelé le « plein de souscription ».

L'utilisation de la réassurance apporte des bénéfices multiples. Elle est en premier lieu un moyen **d'appuyer le développement de la société d'assurance**. Une protection de réassurance permet à l'assureur de souscrire des polices à capital

¹ La coassurance est une assurance partagée entre plusieurs assureurs garantissant selon un pourcentage déterminé à l'avance, un même risque. En général, un des co-assureurs, l'apériteur, est chargé de la gestion du risque et des relations avec l'assuré. L'inconvénient de cette méthode pour l'assureur réside d'une part dans la perte d'autonomie et dans la vulnérabilité vis-à-vis de la concurrence qu'elle implique, et d'autre part dans la lourdeur de la gestion qu'elle engendre.

élevé. Elle lui accorde la possibilité d'augmenter son plein de souscription, et, par conséquent, de défendre ou d'accroître ses parts de marché et son volume d'affaires. Par ailleurs, la réassurance peut être utilisée pour financer le développement d'un portefeuille d'assurance en particulier, le réassureur assumant alors une partie des coûts liés à la souscription de nouvelles polices. La cédante peut aussi éventuellement bénéficier de l'expertise du réassureur ; ce dernier, ayant bien souvent une vision globale du marché, est généralement qualifié pour fournir des conseils en matière de création de nouveaux produits ou de méthodes de sélection et d'évaluation des risques.

La cession en réassurance participe également au **respect des impératifs et des objectifs de solvabilité**. En assurant un transfert des déséquilibres par un écrêtement de la sinistralité, c'est à dire en diminuant la probabilité de ruine de la cédante, elle peut indéniablement constituer un moyen de renforcer la solvabilité d'une société d'assurance. Ce rôle est d'ailleurs généralement reconnu par les différentes autorités de contrôles nationales².

Enfin, de façon plus globale, la réassurance peut contribuer à la **stabilité des résultats de la société d'assurance**³. C'est sans doute l'une des fonctions les plus fondamentales de la réassurance contemporaine. Toute société d'assurance, cotée ou non, voit dans l'utilisation de la réassurance un moyen d'atteindre des objectifs de rentabilité que ses dirigeants ou ses éventuels actionnaires lui ont fixés. Le transfert des écarts de sinistralité que peut apporter une protection de réassurance engendre une diminution de la volatilité des résultats de la cédante. A chaque exercice, le risque de ne pas atteindre des objectifs de rentabilité peut ainsi en être limité et la probabilité d'obtenir un certain rendement minimum en être augmentée.

La protection apportée par la réassurance a bien évidemment pour contrepartie le paiement d'une prime (la prime de réassurance) par l'assureur à son ou ses réassureur(s). Ces derniers, poursuivant tout comme l'assureur des objectifs de rentabilité et de stabilité des résultats, mutualisent leurs risques à une échelle géographique supérieure à celle de l'assureur direct. Les sociétés de réassurance sont donc généralement par nature des groupes de taille conséquente à dimension internationale. De plus, elles ont, comme l'assureur direct, recours à la réassurance,

² En France, la Commission de Contrôle des Assurances veille notamment à ce que la richesse propre des sociétés d'assurance soit supérieure à un minimum de marge de solvabilité. Le poids des cessions en réassurance intervient explicitement dans le calcul de ce minimum.

³ Cette contribution est plus ou moins évidente selon les modalités et les caractéristiques de l'accord de réassurance. Cf. la section 1.2 Les types de réassurance, les mécanismes de réassurance.

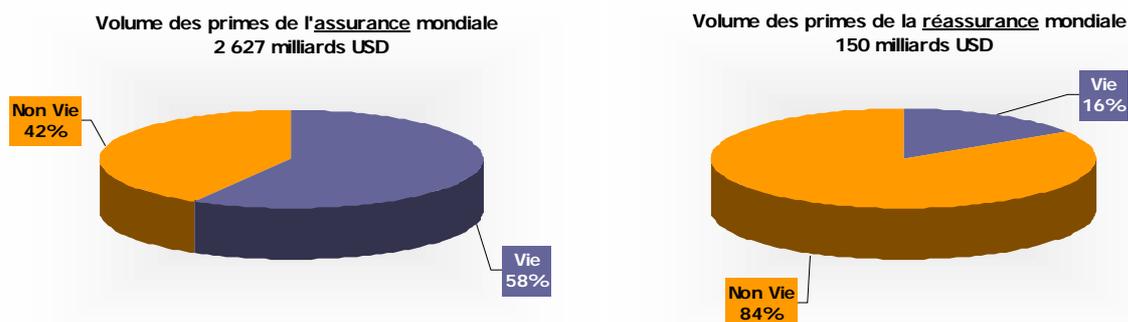
appelée alors rétrocession, c'est-à-dire « l'opération par laquelle un réassureur cède à son tour, une partie des risques qu'il a réassurés à un rétrocessionnaire qui peut être une société de réassurance ou société d'assurance ».

La réassurance contribue à la marge de solvabilité d'une cédante et lui permet de libérer du capital destiné à couvrir des risques. Elle apporte donc à cette dernière la possibilité de croître en limitant l'appel aux fonds propres ou à l'endettement.

En vue de participer à la stabilité des résultats, une protection de réassurance efficace doit également se caractériser par un réel transfert de risque, où la cédante consent à diminuer chaque année l'espérance de son résultat (en transférant une partie de sa marge au réassureur), afin d'atténuer significativement l'impact sur ses comptes d'une sinistralité exceptionnelle lors d'exercices plus difficiles.

1.1.2 Quelques données sur l'assurance et la réassurance

Le poids de l'assurance directe dans l'économie, en termes de chiffre d'affaires, est considérable. En 2002, le volume des primes d'assurance s'élevait à 2 627 milliards USD à l'échelle mondiale (hors organismes publics d'assurance sociale)⁴. Le volume des primes de réassurance était estimé quant à lui à 150 milliards USD.



En 2002, les 495 sociétés d'assurance établies en France ont réalisé sur le territoire national un chiffre d'affaires de 131,7 milliards EUR. Celui des sociétés de réassurance s'élevait quant à lui à 9,4 milliards EUR)⁵. Ainsi, l'activité d'assurance en

⁴ « World Insurance in 2002 ». Swiss Re.

⁵ « L'assurance Française en 2002 ». FFSA.

général (hors système de protection sociale obligatoire), représente **près de 10 % du PIB français**.

Si la majorité du chiffre d'affaires de l'assurance directe est réalisée en Vie, celui de la réassurance à l'inverse s'effectue principalement sur des cessions Non Vie. En Non Vie, le taux moyen de cession en réassurance oscille généralement entre 10 % et 15 % selon les années et les pays, alors qu'il ne dépasse qu'exceptionnellement les 5 % en Vie. Une telle différence illustre l'existence d'une demande de réassurance de la part des cédantes bien plus importante dans la branche Non Vie. Le comportement de la sinistralité de cette dernière est effectivement plus variable, faisant de la réassurance un besoin plus structurel en Non Vie qu'en Vie.

1.1.3 Les modes de réassurance

Il est d'usage de distinguer trois modes de réassurance traditionnelle. Ceux-ci se distinguent par les dispositions du traité de réassurance : nature du risque couvert, part de polices concernées, spécificité des obligations contractuelles.

La réassurance facultative est un contrat conclu entre la cédante et un réassureur portant sur un ou des risques explicitement identifiés (un site de production en particulier, un ensemble de biens clairement délimité, etc.) et qui sont l'objet d'un seul et même contrat. La cession et l'acceptation en réassurance du risque en question résulte d'un libre choix de la part des deux acteurs, d'où la dénomination « facultative ».

Ce mode de réassurance établit un lien direct entre le sort de la cédante et celui du réassureur, ces derniers partageant bien souvent dans la pratique leur appréhension de la qualité intrinsèque du risque concerné, et leur perception de la tarification appropriée.

Si les parties sont pleinement libres de contracter ou non, elles doivent cependant supporter une négociation généralement longue et coûteuse, ainsi qu'une gestion exigeante. La réassurance facultative est donc plutôt adaptée à des risques hors norme, c'est à dire à des polices d'assurance pour lesquelles les valeurs assurées et le montant des cotisations justifient ce déploiement de ressources. Elle est ainsi principalement utilisée pour couvrir les grands risques industriels, et on

l'observe couramment dans les branches construction, individuelle accidents ou annulation d'événements.

La réassurance facultative obligatoire (ou Fac-ob) laisse à la cédante la liberté d'inclure ou non au traité de réassurance les risques de son choix, le réassureur quant à lui s'engageant alors à les accepter dans leur totalité. Les dispositions contractuelles du traité de réassurance facultatif obligatoire sont prévues à l'avance : nature des risques couverts, rétention minimale de la cédante, engagement maximal du réassureur, exclusions éventuelles.

Les avantages de ce mode de réassurance sont indéniables pour la cédante. Celle-ci peut souscrire de façon plus flexible (montant des garanties, rythme du développement), car elle bénéficie d'une plus grande liberté de cession en réassurance. En contrepartie le réassureur court le risque d'avoir un portefeuille déséquilibré, dont la composition peut être exceptionnellement risquée. Enfin, la gestion d'un tel traité demeure assez lourde. De ce fait, le poids de la réassurance facultative obligatoire dans la réassurance en général reste traditionnellement limité.

La réassurance obligatoire représente le mode de réassurance le plus important en volume d'affaires. L'emploi du mot « traité » désigne ainsi implicitement dans la pratique un traité de réassurance obligatoire. Dans le cadre d'un tel traité, l'assureur a l'obligation d'inclure la totalité des polices d'un portefeuille d'une branche définie, le réassureur celle de l'accepter de façon automatique. Ce mode de réassurance donne lieu à une gestion bien plus simple que les modes précédemment décrits. Il se distingue également par un transfert de risques souvent nombreux et homogènes, permettant ainsi l'application de techniques actuarielles d'analyse.

L'ensemble de ces modes de réassurance fonctionne de façon complémentaire. La base d'un **programme de réassurance** est constituée de traités obligatoires, les facultatives étant utilisées parallèlement en vue de couvrir des risques exceptionnels de par leur montant ou difficilement mutualisables de par leur nature.

1.2 Les types de réassurance, les mécanismes des traités de réassurance

Le transfert de risque de la cédante vers son ou ses réassureur(s) peut s'effectuer par le biais de deux types de réassurance : proportionnel ou non proportionnel.

1.2.1 La Réassurance proportionnelle

Un traité de réassurance de type proportionnel se caractérise par l'égalité entre la proportion de primes reçues par le réassureur et la proportion du coût des sinistres transférés au réassureur :

$$\frac{\text{prime de réassurance}}{\text{primes totales reçues par la cédante}} = \frac{\text{montant des sinistres à la charge du réassureur}}{\text{montant brut des sinistres à la charge de la cédante}}$$

Le traité en Quote-Part ou traité en participation constitue une première forme de réassurance proportionnelle. D'un traité en Quote-Part résulte un véritable partage du sort entre la cédante et son réassureur. A la totalité des risques d'un portefeuille donné s'applique un unique taux de cession (des primes et des sinistres). Les dispositions contractuelles d'une Quote-Part définissent : le taux de cession, la période de validité du traité, le champ d'application (la catégorie d'affaires visée assortie d'éventuelles exclusions, l'étendue géographique), mais également le taux de commission de réassurance et les modalités d'une éventuelle participation aux bénéfices.

La commission de réassurance a pour vocation de faire supporter au réassureur une partie des coûts liés aux frais d'acquisition et de gestion des contrats d'assurance engagés par la cédante. Elle est exprimée dans les traités en pourcentage des primes perçues par la cédante.

Si un traité en Quote-Part se traduit par une gestion simple et un principe de répartition équitable, certaines insuffisances peuvent néanmoins être imputées à son mécanisme. D'une part, le montant des primes cédées peut être jugé comme excessif du point de vue de la cédante (notamment lorsque le résultat brut de

souscription est bon). D'autre part, le mécanisme peut se révéler inapproprié, voire inefficace, face à des sinistres de montants individuels élevés. Enfin, pour le réassureur, la participation à l'intégralité du portefeuille de la cédante s'accompagne d'une dépendance vis-à-vis des qualités de souscription et de tarification de cette dernière.

Le traité en excédent de plein, dans une certaine mesure, permet à la cédante de remédier à ces inconvénients. Son mécanisme repose sur la variabilité du taux de cession en relation avec un fractionnement du portefeuille en différents ensembles de polices.

Ainsi les dispositions contractuelles du traité en excédent de plein explicitent les mêmes caractéristiques que celles d'un traité en Quote-Part, et définissent en sus les modalités plus complexes de cession. Pour chacun des sous-ensembles de polices, le traité indique le plein de rétention et le plein de souscription de la cédante, comme l'illustre l'exemple suivant :

Polices A	
Plein de souscription: 200	Plein de rétention : 10
Polices B	
Plein de souscription: 1000	Plein de rétention : 100

Il est alors possible de calculer le taux de cession pour chaque sous-ensemble de polices en fonction des capitaux assurés par police⁶, de la manière suivante :

$$\text{Taux de cession} = \text{Max}\left[0; \frac{\text{Min}(\text{Plein de souscription}; \text{Capitaux assurés}) - \text{Plein de rétention}}{\text{Min}(\text{Plein de souscription}; \text{Capitaux assurés})}\right]$$

On remarque que :

- si les capitaux assurés par police d'un sous-ensemble sont inférieurs au plein de rétention correspondant, le taux de cession est nul et la cédante conserve le risque dans son intégralité.

⁶ Il convient de noter que ceci implique l'existence et la connaissance des sommes assurées par police, ce qui n'est dans la pratique pas toujours le cas. A défaut, on utilise souvent une estimation du SMP par police, ou sinistre maximum probable.

- si ces capitaux assurés sont supérieurs au plein de souscription correspondant, alors le taux de rétention (qui est égal à 1 - taux de cession) est limité au rapport plein de rétention / plein de souscription.

En appliquant la formule à l'exemple précédent, on obtient les différents pourcentages de cession :

Polices A			
Plein de souscription: 200		Plein de rétention : 10	
Capitaux assurés	Rétention	Taux de rétention	Taux de cession
5	5	100,00%	0,00%
10	10	100,00%	0,00%
20	10	50,00%	50,00%
50	10	20,00%	80,00%
70	10	14,29%	85,71%
100	10	10,00%	90,00%
150	10	6,67%	93,33%
200	10	5,00%	95,00%
300	10	5,00%	95,00%

Polices B			
Plein de souscription: 1000		Plein de rétention : 100	
Capitaux assurés	Rétention	Taux de rétention	Taux de cession
50	50	100,00%	0,00%
100	100	100,00%	0,00%
500	100	20,00%	80,00%
800	100	12,50%	87,50%
1000	100	10,00%	90,00%
3000	100	10,00%	90,00%

En conséquence, le traité en excédent de plein permet de faire correspondre ampleur du risque (en termes de capitaux sous risque) et niveau de la cession, source a priori d'une meilleure homogénéisation des conservations nettes. Toutefois sa gestion en est plus complexe que celle d'une Quote-Part.

En pratique, les traités en excédent de plein seront préférés aux Quote-Part lorsque les capitaux sous risques sont très variables d'une police à l'autre, comme par exemple dans les branches Incendie risques industriels ou GAV (Garantie Accidents de la Vie). L'utilisation d'une Quote-Part répondra davantage aux besoins de sociétés jeunes ou de cédantes qui se lancent sur de nouvelles branches, et qui ne maîtrisent pas parfaitement les risques de souscription.

Enfin, les protections de réassurance de type proportionnel peuvent souvent se révéler onéreuses compte tenu de l'efficacité de leur contribution au résultat technique d'une cédante. La cession d'une partie, éventuellement importante, de primes liées à des risques qu'une cédante est en mesure d'assumer seule, constitue généralement un transfert de risque inadéquat.

1.2.2 La Réassurance non proportionnelle

La garantie apportée à une cédante par un traité de réassurance non proportionnelle est liée, indépendamment des sommes assurées, à la seule éventualité de survenance d'un sinistre ou d'une perte.

L'indemnisation due par le réassureur équivaut au montant du sinistre qui excède un seuil préalablement défini dans le traité, **la priorité, ou franchise**, et ne peut généralement pas dépasser une certaine limite, elle aussi définie contractuellement, **la portée**. Le plafond d'un traité de réassurance est alors défini comme la somme de la priorité et de la portée. La part éventuelle du sinistre qui excède ce plafond demeure à la charge de la cédante.

En contrepartie, la cédante verse au réassureur une prime liée au transfert de risque qu'implique la couverture. La réassurance est non proportionnelle car il n'existe plus de relation purement proportionnelle entre le montant de la prime et le montant des sinistres cédés.

Par ailleurs, une couverture non proportionnelle sera fréquemment composée de plusieurs tranches, chacune étant caractérisée par un couple priorité – portée. Considérons par exemple le cas où une cédante bénéficie d'une couverture constituée de N tranches, où les montants P_n et F_n décrivent respectivement la portée et la franchise de la tranche n (n étant un nombre entier de 1 à N). Dans la pratique, toute tranche $n+1$ sera telle que $F_{n+1} = F_n + P_n$, si bien que l'ensemble de la couverture assure à la cédante une protection de $\sum_{n=1}^{n=N} P_n$ en excédent de F_1 . Le plafond de la couverture s'élève alors à $F_N + P_N$. L'ampleur du risque motivant l'emploi d'une couverture non proportionnelle justifie le plus souvent dans la pratique le recours à un tel découpage. Ce dernier a généralement entre autres avantages de

faciliter le placement de la couverture et de diversifier le risque de crédit de la cédante, la multiplicité des tranches engendrant l'intérêt d'une plus grande variété de réassureurs.

Le traité en excédent de sinistre constitue une première forme de réassurance non proportionnelle. On parle d'excédent de sinistre **par risque** lorsque la couverture et ses modalités s'appliquent risque par risque, c'est à dire police par police. On parle d'excédent de sinistre **par événement** lorsque la couverture et ses modalités s'appliquent pour des sinistres qui sont le cumul de risques individuels sinistrés lors d'un même événement. La notion d'événement fait alors l'objet d'une définition particulière au sein du contrat de réassurance. Dans la pratique, il est fréquent d'observer l'utilisation de couvertures en excédent de sinistre par événement pour la réassurance des dommages liés aux périls naturels ou au risque terroriste⁷ notamment.

Le traité en excédent de perte, ou stop loss, est également une forme de réassurance non proportionnelle. L'objet de la couverture ne porte plus sur un montant de sinistres élémentaires survenus à un moment donné, mais sur l'intégralité de la charge sinistre enregistrée au cours d'un exercice ou d'une période préalablement définie. La priorité et la portée d'un stop loss sont généralement exprimées en pourcentage de l'encaissement des primes de la cédante, mais peuvent toutefois être définies par des montants fixes. Enfin, la priorité est traditionnellement fixée à un niveau tel que l'indemnisation de réassurance ne puisse avoir lieu qu'en cas de réelle perte technique de la cédante (où la somme des sinistres et des frais est supérieure aux primes).

Une couverture en excédent de perte est très appréciable pour la cédante, puisqu'elle peut ainsi protéger directement son bilan de toute fluctuation exceptionnelle de sa charge sinistre. Mais le coût du transfert de risque qui en résulte est généralement assez élevé.

La réassurance non proportionnelle est ainsi bien adaptée à la couverture de risques catastrophiques, c'est à dire de risques qui ont une faible probabilité de survenance, mais qui sont susceptibles de mettre en péril de façon significative les comptes d'une cédante. Le montant de la prime est souvent minime par rapport au montant de l'indemnisation en cas de sinistre majeur.

⁷ Risque de conflagration lié à un événement de nature terroriste.

D'une manière générale, un programme de réassurance efficace doit permettre à une cédante de limiter les fluctuations annuelles de la sinistralité qu'elle conserve, et assurer un véritable transfert de la sinistralité catastrophique.

Toutefois, le besoin et l'intérêt d'une cédante vis à vis de la réassurance diffèrent selon les branches d'assurance et la nature même des périls assurés. Il en va consécutivement de même pour les solutions de réassurance (type, forme, modalités) qu'elle est amenée à choisir. La partie suivante va permettre d'approfondir la question en matière de périls naturels.

1.3 Le besoin de réassurance en matière de risques naturels

1.3.1 Définitions du risque naturel, des catastrophes naturelles

Il n'existe pas, au travers des textes réglementaires, de définition unique du risque naturel. Il est toutefois possible de décrire le risque naturel comme la menace, potentielle ou non, que des aléas issus d'un phénomène naturel (tempête, inondation, séisme, cyclone, sécheresse, mouvement de terrain, etc.) font peser sur des enjeux humains (sur des personnes ou sur des intérêts humains). La notion de risque naturel diffère de celle d'évènement naturel par la présence de la vulnérabilité qu'elle suppose pour l'homme et ses activités. Cependant, dans les pages qui suivent, et comme bien souvent dans la pratique, le terme de risque naturel sera utilisé dans un sens très généraliste.

Les risques naturels sont caractérisés par une grande diversité de leur dangerosité, du simple accident naturel à faible conséquence à la catastrophe naturelle dont les répercussions économiques et humaines sont considérables. Au sens large, on parle de catastrophe naturelle, lorsqu'un évènement naturel, localisé dans le temps et dans l'espace, entraîne de lourdes pertes humaines ou matérielles, et compromet gravement le bon accomplissement des fonctions de base de la société.

L'impact des catastrophes naturelles sur la société humaine peut donc être considérable, en témoigne l'historique des dix plus importantes catastrophes naturelles de l'histoire contemporaine en matière de préjudice économique⁸ :

Date	Pays	Evènement	Nombre de victimes	Estimation du préjudice économique (USD milliards)
17/01/1995	Japon	Tremblement de Terre - Kobe	5 502	131,5
24/04/1991	ex URSS	Inondations - région de Volgograd	-	60
24/08/1992	Etats-Unis, Bahamas	Cyclone Andrew - Floride, Bahamas	44	30
17/01/1994	Etats-Unis	Tremblement de Terre - Northridge	60	26
07/12/1988	ex URSS	Tremblement de Terre - Arménie	25 000	20,5
23/11/1980	Italie	Tremblement de Terre - Naples	4 689	20
06/08/1998	Chine	Inondations	3 656	20
sept-1997	Indonésie	Incendies de Forêts - Sumatra et Kalimantan	6	17
juil-1995	Corée du Sud	Inondations	68	15
juil-1996	Chine	Inondations	2 775	12,6

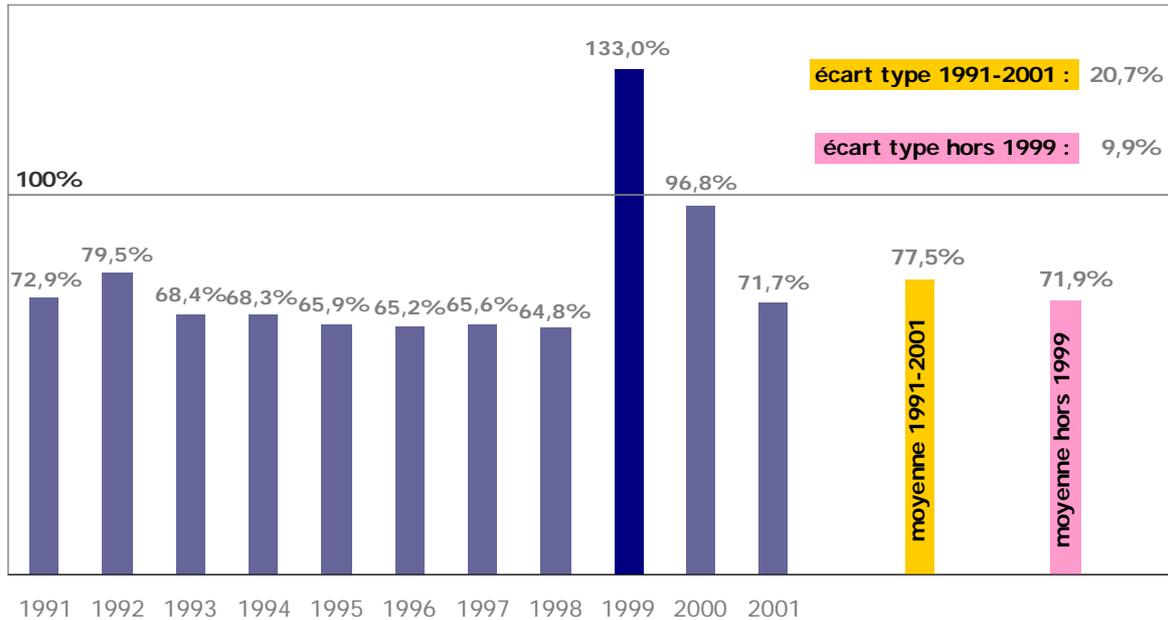
source : CRED, EM-Dat

1.3.2 Risque de premier plan pour l'assurance

Une société d'assurance doit faire face, avec le risque naturel, à l'une des plus sérieuses menaces qui pèsent sur la stabilité de ses résultats. En matière de risque naturel, un seul évènement est en mesure de mettre en péril la pérennité même de l'ensemble des activités d'un assureur. Par exemple, le coût des tempêtes Lothar et Martin a représenté, pour bon nombre de compagnies d'assurances du marché français, près de la totalité des cotisations de l'ensemble de la branche Dommages aux Biens encaissées sur l'exercice 1999. L'évolution sur la période 1991-2001 des rapports sinistres à primes de la branche Dommages aux Biens de l'ensemble du marché français permet de mettre en évidence le caractère exceptionnel de cette catastrophe naturelle :

⁸ Cet historique peut varier sensiblement selon les sources, selon les méthodes de recueil des données et d'estimation des dommages.

Evolution des ratios S/P, branche DAB, France



La sinistralité événements naturels ne représente ordinairement (pour la majorité des exercices) qu'une faible part de la sinistralité annuelle de la branche DAB. Toutefois, la survenance fin décembre 1999 de Lothar et Martin est responsable d'une détérioration brutale des ratios de sinistralité des cédantes du marché français pour l'exercice 1999⁹.

Bien souvent, un portefeuille d'assurances est constitué de polices pour lesquelles les probabilités de survenance d'un sinistre sont indépendantes les unes des autres (incendie d'un bâtiment, dégât des eaux, accident de la route...). La loi des grands nombres tend alors à être applicable, et, par l'effet de la mutualisation des risques, l'assureur est en mesure de prévoir assez correctement la charge de sinistres à laquelle il va être confronté. Mais, dans le contexte d'événements catastrophiques, tels les événements naturels, les sinistralités de chaque risque individuel ne sont manifestement pas indépendantes. Une catastrophe naturelle affecte généralement la totalité des polices d'un portefeuille qui sont situées dans la zone géographique touchée. La mutualisation des risques ne peut ainsi pas jouer son rôle traditionnel lorsqu'il s'agit de catastrophes naturelles. L'accumulation de

⁹ L'effet des tempêtes est pourtant réparti sur les deux exercices comptables 1999 et 2000, leur survenance en toute fin d'année ayant engendré de grosses difficultés d'estimation de la sinistralité.

polices, notamment au sein d'une zone géographique restreinte comme un territoire national, entraîne directement l'augmentation du sinistre maximum probable.

Par ailleurs, l'évaluation du montant probable d'un sinistre catastrophique est souvent problématique. Les événements naturels catastrophiques ont une périodicité très longue, et l'expérience passée est souvent insuffisante pour appréhender le risque réel. A plusieurs reprises au cours des deux dernières décennies, le risque de tempête catastrophique en France a par exemple été sous-estimé.

L'aléa que les catastrophes naturelles font courir sur les activités des sociétés d'assurance est donc gigantesque, comme l'illustre l'historique des dix plus importantes catastrophes naturelles de l'histoire contemporaine en matière de dommages assurés :

Date	Pays	Evènement	Nombre de victimes	Dommages assurés (USD milliards, en valeur 2003)	Coût des sinistres en pourcentage des primes Non Vie
23/08/1992	Etats-Unis, Bahamas	Cyclone Andrew	38	20 900	5,0%
17/01/1994	Etats-Unis	Tremblement de Terre - Northridge	60	17 312	4,0%
27/09/1991	Japon	Typhon Mireille	51	7 598	6,5%
25/01/1990	France, Royaume Uni,	Tempête Daria	95	6 441	–
25/12/1999	France, Suisse,	Tempête Lothar	110	6 382	–
15/09/1989	Porto Rico, Etats-Unis,	Cyclone Hugo	71	6 203	–
15/10/1987	France, Royaume Uni,	Tempête - Inondations	22	4 839	–
25/02/1990	Europe de l'Ouest, Europe Centrale	Tempête Vivian	64	4 476	–
22/09/1999	Japon	Typhon Bart	26	4 445	3,9%
20/09/1998	Etats-Unis, Caraïbes	Cyclone Georges	600	3 969	–

source : Swiss Re, Sigma 1/2004

Aussi une protection en réassurance est-elle, dans son rôle de contribution à la stabilité des résultats et de diminution de la probabilité de ruine de l'assureur, d'autant plus primordiale lorsqu'il s'agit de couverture contre le risque naturel. Le tableau suivant atteste de l'efficacité de l'aide que peut apporter la réassurance : l'impact des tempêtes de décembre 1999 sur les résultats des sociétés d'assurance directes du marché français a été considérablement réduit par la fonction d'écrêtement de la sinistralité remplie par les protections de réassurance.

Compte de résultat technique simplifié
Ensemble des sociétés d'assurances du marché français - Branche DAB

<i>chiffres en EUR millions</i>	1998	1999
Brut de Réassurance		
Primes	8 839	8 811
Charges des prestations	5 726	11 716
Net de Réassurance		
Primes	7 254	7 507
Charges des prestations	4 500	6 559
Produits Financiers	490	749
Autres charges techniques	2 532	2 508
Résultat technique brut de réassurance	1 071	-4 664
Résultat technique net de réassurance	712	-811

source : Commission de contrôle des assurances

Un solde de réassurance nettement positif en 1999, en grande partie lié aux indemnisations perçues au titre des événements naturels, a permis d'améliorer le résultat technique de la branche DAB de près de EUR 4 milliards.

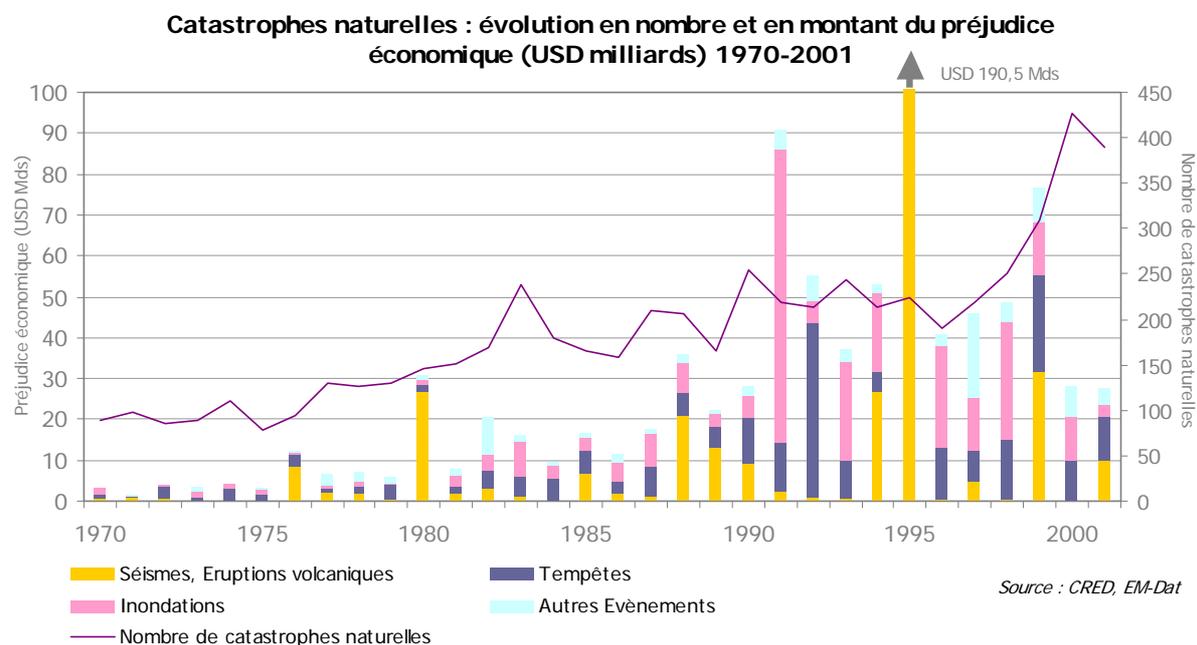
1.3.3 Un risque croissant

Il n'existe aujourd'hui à l'échelle internationale ni institution unique spécifiquement dédiée au recueil et à la vérification des données sur les catastrophes naturelles, ni méthode standard d'évaluation des dommages. Toutefois, Munich Re, Swiss Re et le CRED¹⁰ constituent des organismes de référence en la matière. Si la qualité des données disponibles n'est considérée comme réellement bonne que depuis les années 1980, une observation rapide des statistiques établies à partir de ces trois sources différentes permet de déceler de façon manifeste un accroissement significatif de la fréquence et de la sévérité des catastrophes naturelles au cours des dernières décennies.

Dans le graphique ci-dessous, élaboré à partir des données de EM-Dat, sont considérés comme catastrophes naturelles les événements naturels ayant entraîné un ou plusieurs des maux suivants : le décès de 10 personnes ou plus, un nombre

¹⁰ Le CRED (Centre for Research on the Epidemiology of Disasters, Université Catholique de Louvain – Bruxelles), entretient à des fins humanitaires, en collaboration avec l'OFDA (Office of U.S. foreign disaster assistance), la base de données publique sur les catastrophes mondiales EM-Dat (Emergency Disasters Data Base).

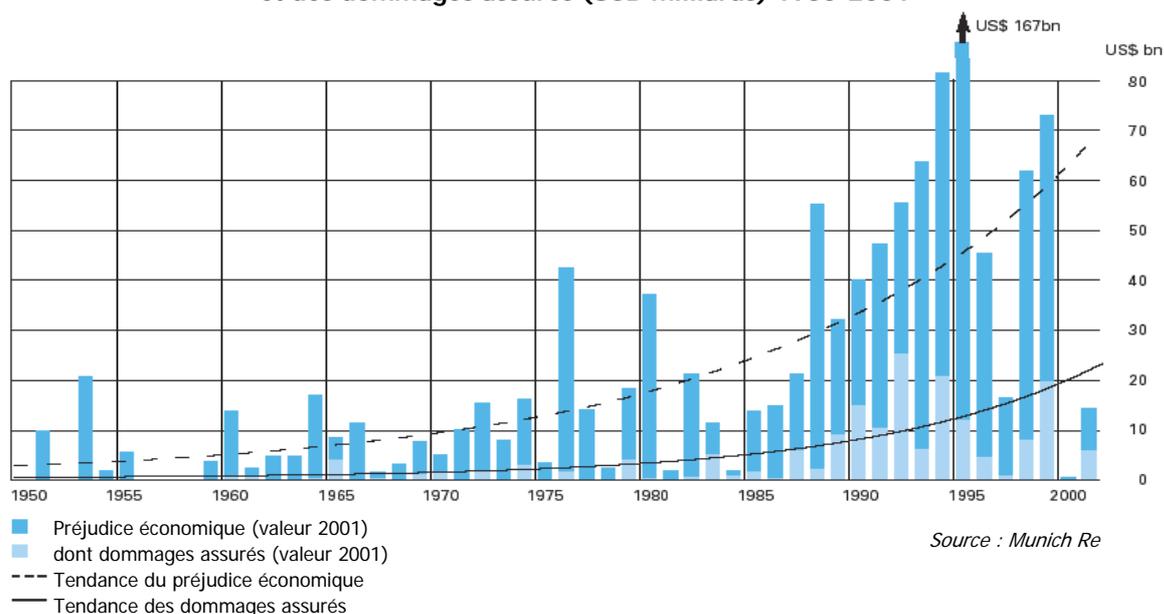
de blessés ou de sans-abri supérieur ou égal à 100, la reconnaissance par des autorités locales de l'état d'urgence, et/ou un appel à l'aide internationale.



Si l'on observe l'évolution des grandes catastrophes naturelles et leurs tendances sur le graphique suivant publié par Munich Re¹¹, on constate à nouveau l'augmentation indéniable des dommages économiques engendrés par les catastrophes naturelles, mais également l'accroissement conjoint des dommages à la charge des sociétés d'assurance. Ces derniers connaissent d'ailleurs une croissance encore plus soutenue, du fait de l'augmentation simultanée des expositions (de plus en plus de biens assurés).

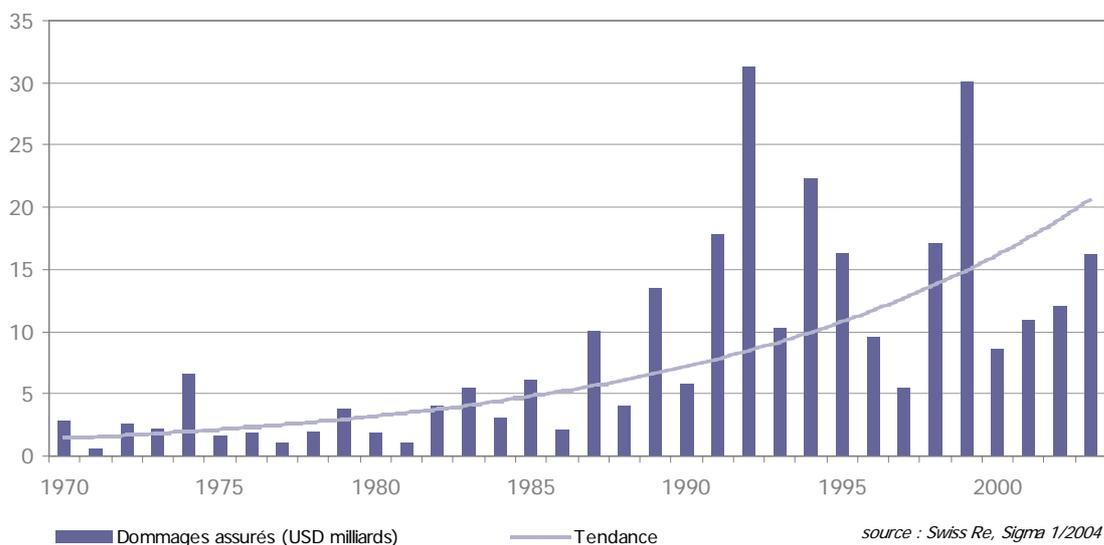
¹¹ « Rétrospective des catastrophes naturelles survenues en 2001 ». Revue Topics, Munich Re Group. Sont définies comme « grandes catastrophes naturelles », les événements naturels qui entraînent pour les régions affectées « l'incapacité de faire face seules à la situation et l'obligation de demander du secours à l'échelon national ou international », ce qui est « généralement le cas lorsque le nombre des victimes atteint le millier, celui des sans-abri la centaine de milliers ou que le pays touché subit un préjudice économique substantiel par rapport à sa situation économique ».

Catastrophes Naturelles : évolution du préjudice économique et des dommages assurés (USD milliards) 1950-2001



Enfin, la tendance est également confirmée par les données publiées par Swiss Re¹², qui sont reprises dans le graphique suivant :

Catastrophes naturelles : évolution des dommages assurés (USD milliards en valeur 2003) 1970-2003



¹² « Catastrophes Naturelles et techniques en 2003 : Un grand nombre de victimes, mais des dommages assurés modérés ». Revue Sigma N°1/2004.

Les différentes sources d'information sur les catastrophes naturelles indiquent d'une part que la fréquence de tels événements aurait quadruplé sur les trente dernières années, et d'autre part que les préjudices économiques et les dommages assurés qui en résultent se seraient accrus d'un facteur compris entre 10 et 20 sur la même période.

Cette croissance peut s'expliquer en grande partie par des facteurs non météorologiques. L'accroissement de la densité démographique et celui du montant des biens assurés, en particulier dans les zones à risques, sont fréquemment cités comme causes principales de cette évolution. La grande fragilité de certains matériaux et technologies modernes est également évoquée¹³.

Toutefois, des analyses scientifiques de plus en plus nombreuses attestent de la vraisemblance d'une recrudescence des catastrophes naturelles en elles-mêmes. L'étude de données météorologiques, tout comme l'observation des évolutions historiques de fréquence et de sévérité des catastrophes naturelles, appuie l'hypothèse d'une croissance du risque naturel.

Le rapport publié en 2001 par le Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC)¹⁴ décrit notamment les principaux changements climatiques détectés à partir de l'étude de plusieurs indicateurs : un réchauffement de couche inférieure de l'atmosphère, une augmentation des précipitations dans certaines régions, et une augmentation du risque de sécheresse dans d'autres. Le tableau suivant, extrait de ce rapport, synthétise les observations les plus significatives.

¹³ « *Catastrophes naturelles et réassurance* ». Publication Swiss Re 2003.

¹⁴ « *Rapport de synthèse, Changements Climatiques 2001* ». IPCC (Intergovernmental Panel on Climate Change).

Indicateur Climatique	Changement observé
Température moyenne mondiale à la surface	Augmentation de 0,6°C (+/- 0,2 °C) au cours du 20 ^{ème} siècle Réchauffement plus important des zones terrestres que des océans (90 à 99 % de probabilité).
Température à la surface dans l'hémisphère Nord	Augmentation au cours du 20 ^{ème} siècle plus importante qu'au cours de tout autre siècle du dernier millénaire ; Années 1990 : décennie la plus chaude du millénaire (66 à 90 % de probabilité).
Fourchette de températures à la surface diurnes	Diminution entre 1950 et 2000 sur les zones terrestres Augmentation deux fois plus rapide des températures minimales nocturnes que des températures maximales diurnes (66 à 90 % de probabilité).
Jours chauds/indice de chaleur	Augmentation (66 à 90 % de probabilité).
Jours froids/de gel	Diminution pour la quasi totalité des zones terrestres au cours du 20 ^{ème} siècle (90 à 99 % de probabilité).
Précipitations continentales	Augmentation de 5 à 10 % au cours du 20 ^{ème} siècle dans l'hémisphère Nord (90 à 99 % de probabilité), mais diminution sur certaines régions (Afrique du Nord et occidentale et certaines parties de la Méditerranée, par exemple).
Fortes précipitations	Augmentation aux latitudes nord moyennes et supérieures (66 à 90 % de probabilité).
Fréquence et intensité de la sécheresse	Absence accrue de précipitations en été et augmentation de la sécheresse associée dans quelques zones (66 à 90 % de probabilité). Dans certaines régions, telles que certaines parties de l'Asie et de l'Afrique, on a observé une augmentation de la fréquence et de l'intensité de la sécheresse au cours des dernières décennies.

Les travaux publiés dans la revue *La Météorologie* - n°38¹⁵, étayent également une partie de ces conclusions à l'échelle de la France :

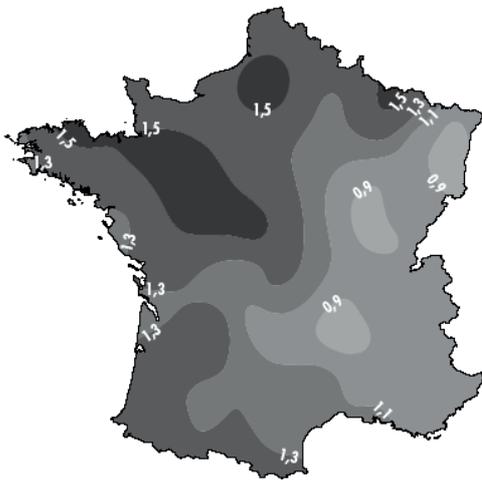
¹⁵ « *Les Changements climatiques en France au XX^e siècle. Etude des longues séries homogénéisées de données de température et de précipitations* ». Jean-Marc Moisselin, Michel Schneider, Claire Canellas et Olivier Mestre. *La Météorologie* – n°38 – août 2002.

La détection d'une tendance à la hausse de la température moyenne sur les 100 dernières années :



*Cartographie des tendances 1901-2000 (en °C/siècle) de la température moyenne (à partir de 70 séries).
Figure extraite de La Météorologie n°38 – août 2002*

La détection d'une tendance à la hausse de la température maximale et d'une tendance à la hausse deux fois plus importante de la température minimale sur les 100 dernières années :



*Cartographie des tendances 1901-2000 (en °C/siècle) de la température minimale (à partir de 70 séries homogénéisées de températures minimales).
Figure extraite de La Météorologie n°38 – août 2002*



*Cartographie des tendances 1901-2000 (en °C/siècle) de la température maximale (à partir de 70 séries homogénéisées de températures maximales).
Figure extraite de La Météorologie n°38 – août 2002*

Ces travaux corroborent également, mais de façon moins évidente, les conclusions concernant la hausse des précipitations (pour la partie Nord de la France) et l'augmentation du risque de sécheresse (pour le Sud de la France).

Pour le GIEC, ces différentes évolutions climatiques, qui pourraient s'accompagner d'une intensification du cycle de l'eau, laissent ainsi présager une augmentation des risques de catastrophes naturelles, en particulier les inondations et la sécheresse, dans de nombreux endroits du globe. Bien que moins consensuelles, les hypothèses d'une recrudescence de l'activité cyclonique tropicale et d'une augmentation de l'intensité des tempêtes aux latitudes moyennes sont également envisageables. Une grande incertitude demeure cependant quant à l'évolution des phénomènes météorologiques extrêmes. Pour les offices météorologiques, des études approfondies sont nécessaires pour mieux appréhender l'évolution des valeurs extrêmes.

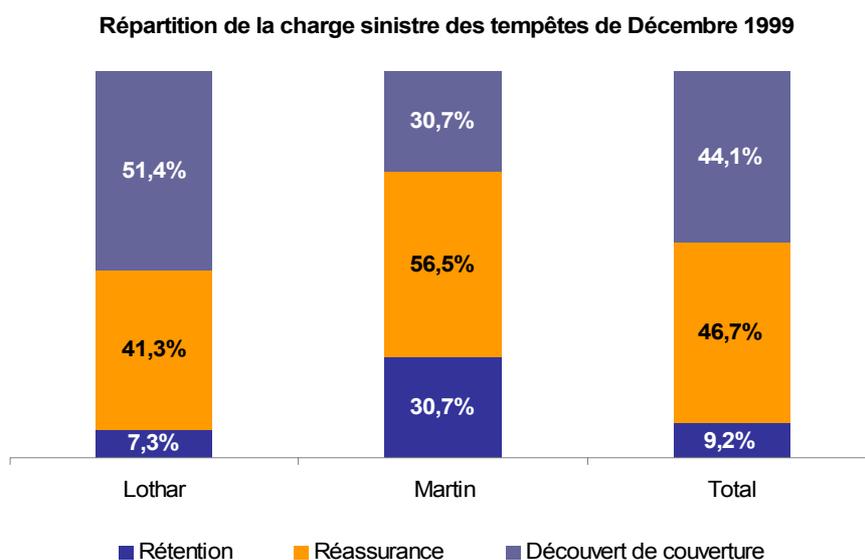
Face à ces incertitudes, les acteurs de l'assurance disposent depuis quelques années d'outils de modélisation de plus en plus perfectionnés des périls naturels¹⁶. Toutefois, ces outils, dont le fonctionnement s'appuie sur une vision simplifiée de la réalité (définition d'un événement naturel, évaluation de son intensité, estimation des dommages concomitants, calcul de la perte économique), souffrent encore d'un manque de crédibilité. Les résultats varient effectivement assez largement d'un modèle à l'autre. De plus, ils reposent quasi exclusivement sur l'expérience passée afin de prévoir les événements futurs (notamment en matière de fréquence des événements), ne prenant donc pas en considération les conséquences de probables changements climatiques. En gardant en mémoire leur sensibilité et leurs limites, ces modèles peuvent tout de même se révéler utiles dans la connaissance de l'exposition d'un portefeuille d'assurances aux périls naturels.

La grande hétérogénéité des risques naturels, l'accroissement en nombre et en intensité des catastrophes naturelles ces dernières années se présentent comme des obstacles de poids pour les acteurs de l'assurance et de la réassurance. Aujourd'hui, la couverture des risques naturels est loin d'être totale¹⁷. La majeure

¹⁶ Ce sont principalement trois compagnies qui commercialisent aujourd'hui ce type de logiciels : AIR, EQECAT et RMS.

¹⁷ Cette carence de protection est bien évidemment flagrante dans les pays en développement, mais elle est également notoire dans les pays riches, lesquels sont l'objet véritable des idées exposées ici, où le secteur assurantiel est très développé.

partie des dommages économiques engendrés par les catastrophes naturelles n'est pas assurée à l'heure actuelle. Selon Munich Re¹⁸, sur les USD 64,6 milliards de dommages économiques provoqués par les catastrophes naturelles survenues en 2003, un peu moins de 25 % (15,8 milliards de USD) étaient assurés. La couverture des risques naturels en réassurance peut également s'avérer insuffisante. Ainsi, comme l'illustre la figure suivante, plus de la moitié du montant des sinistres Lothar et Martin fut à la charge des cédantes¹⁹ :



Le chapitre suivant vise à décrire l'éventail de solutions aujourd'hui utilisées par les sociétés d'assurance en vue de répondre le mieux possible à leur besoin très particulier de réassurance en matière de risques naturels.

¹⁸ « *Rétrospective des catastrophes naturelles survenues en 2003* ». Revue Topics, Munich Re Group.

¹⁹ « *Etude 2003 des Renouvellements des protections tempête du marché français* ». Benfield Paris.

2 DE LA REASSURANCE CLASSIQUE A LA REASSURANCE NON TRADITIONNELLE DES RISQUES NATURELS

2.1 La réassurance traditionnelle des risques naturels

2.1.1 Les formes de réassurance traditionnelle des risques naturels

a. En France

Si le territoire français n'est pas, en comparaison de certaines régions du globe, particulièrement exposé aux risques naturels extrêmes, les inondations constituent de façon récurrente des charges substantielles pour les collectivités et les sociétés d'assurance. La survenance d'une crue centennale à Paris, telle celle de 1910, serait tout de même un évènement d'une ampleur extraordinaire qui pourrait entraîner des dommages économiques estimés à près de 10 milliards d'euros²⁰, dont une partie significative incomberait aux assureurs. Les tempêtes hivernales, dont les conséquences économiques peuvent être colossales (Tempête de 1987, Daria, Vivian, Wiebcke et Herta en 1990, Lothar et Martin en 1999), représentent également un risque naturel de premier plan en France métropolitaine. Par ailleurs, bien que peu concernée par le risque sismique, la France pourrait avoir à faire face à des dégâts avoisinant les 15 milliards d'euros si un important séisme venait à se produire dans le Sud-Est. Enfin, le risque naturel est dans les Dom-Tom particulièrement élevé, du fait de l'exposition de ces derniers aux cyclones et/ou aux risques sismique et volcanique.

L'assurance des évènements naturels catastrophiques repose en France sur un système assez particulier, dans lequel on dénombre quatre régimes différents d'indemnisation des dommages matériels.

La loi du 13 juillet 1982 a instauré le régime d'indemnisation des catastrophes naturelles proprement dit. Son apparition fait suite aux conséquences des grandes inondations qui ont affecté la France à la fin de l'année 1981, et vise à remédier à l'absence de protection contre les dommages provoqués par les évènements naturels, une exclusion jusqu'alors de règle dans les contrats

²⁰ « *Entreprises et risque de crues à Paris et en petite couronne* ». Etude de la CCIP, 1999.

d'assurance²¹. Ainsi, la loi rend obligatoire, au sein de tout contrat d'assurance dommages aux biens, l'inclusion d'une garantie contre les événements naturels catastrophiques. De nature mixte, reposant à la fois sur l'assurance de marché et sur la garantie finale de l'Etat, le système fonctionne selon le principe de mutualisation, son financement est assuré par un taux de cotisation identique pour tous les assurés (12 % depuis 1999 pour les contrats Dommages aux Biens). Le déclenchement de l'indemnisation est subordonné à la reconnaissance de l'état de catastrophe naturelle par les autorités publiques au moyen d'un arrêté interministériel. Le champ d'application des dommages couverts a été tout d'abord défini comme « *les dommages matériels directs ayant eu pour cause déterminante l'intensité anormale d'un agent naturel, lorsque les mesures habituelles à prendre pour prévenir ces dommages n'ont pu empêcher leur survenance ou n'ont pu être prises* »²², puis précisé par la loi du 16 juillet 1992 comme « *les seuls dommages matériels non assurables* ». Bien qu'il n'y ait pas de liste précise de ces événements naturels, les inondations et/ou coulées de boue, les séismes, les mouvements de terrain, la sécheresse²³, les raz de marée, les ruissellements d'eau, de boue ou de lave, et les masses de glace ou de neige en mouvement sont couverts par le système des catastrophes naturelles. Par contre, les dommages dus aux effets du vent n'en font plus partie depuis 1990²⁴. Le tableau suivant indique les principaux événements déclarés catastrophes naturelles depuis l'instauration du régime :

Exercice de survenance	Evènement	Estimation coût marché (en millions d'euros)	Exercice de survenance	Désignation	Estimation coût marché (en millions d'euros)
1982/1983	Tempêtes/Inondations (1)	534	1996	Inondations Décembre (Sud-Ouest)	76
1987	Tempête Octobre (1)	107	1997	Inondations Juin (Normandie)	30-40
1988	Inondations Octobre (Nîmes)	290	1998	Inondations Juin (Nord/Pas-de-Calais)	10-20
1990	Inondations Février	183	1999	Inondations Novembre (Grand Sud)	240-250
de 1989 à 2000	Subsidence (2)	3200	1999	Ouragans José et Lenny (DOM) (3)	50-60
1992	Inondations Septembre (Vaison)	244	1999	Tempêtes Lothar et Martin (3)	220-230
1993	Inondations Septembre/Octobre	305	2000	Inondations Décembre (Bretagne)	60-70
1993/1994	Inondations Décembre/Janvier	259	2001	Inondations Janvier (Bretagne-Normandie)	40-50
1994	Inondations Novembre (Nice)	122	2001	Inondations Avril (Somme)	60-80
1995	Inondations Janvier/Février	396	2002*	Cyclone Dina Janvier (Réunion)	90-100
1995	Inondations Août/Septembre (cyclones Antilles) (3)	110	2002*	Inondations Septembre (Sud)	700
1996	Séisme Juillet (Annecy)	61			

source : Caisse Centrale de Réassurance

(1) L'indemnisation au titre des "Cat-Nat" a été effectuée en complément ou à défaut des garanties contractuelles tempêtes.

(2) Il s'agit des dommages causés aux bâtiments par la sécheresse du sous-sol.

(3) Il ne s'agit que des dommages causés par les eaux, les dommages dus au vent relevant de la garantie contractuelle T.O.C. (Tempête, Ouragan, Cyclone)

* estimations

²¹ Cette absence de couverture était le fruit de plusieurs facteurs : une insuffisance de statistiques fiables en la matière, un risque important et mal connu de cumul des expositions, un mécanisme d'antisélection dû au caractère facultatif de la garantie.

²² Loi n°82-600 du 13 juillet 1982 relative à l'indemnisation des victimes de catastrophes naturelles.

²³ Plus précisément la subsidence, i.e. les dommages causés aux bâtiments par la sécheresse du sous-sol.

²⁴ Loi n°90-509 du 25 juin 1990. Modifiant le code des assurances et portant extension aux départements d'outre-mer du régime d'indemnisation des catastrophes naturelles.

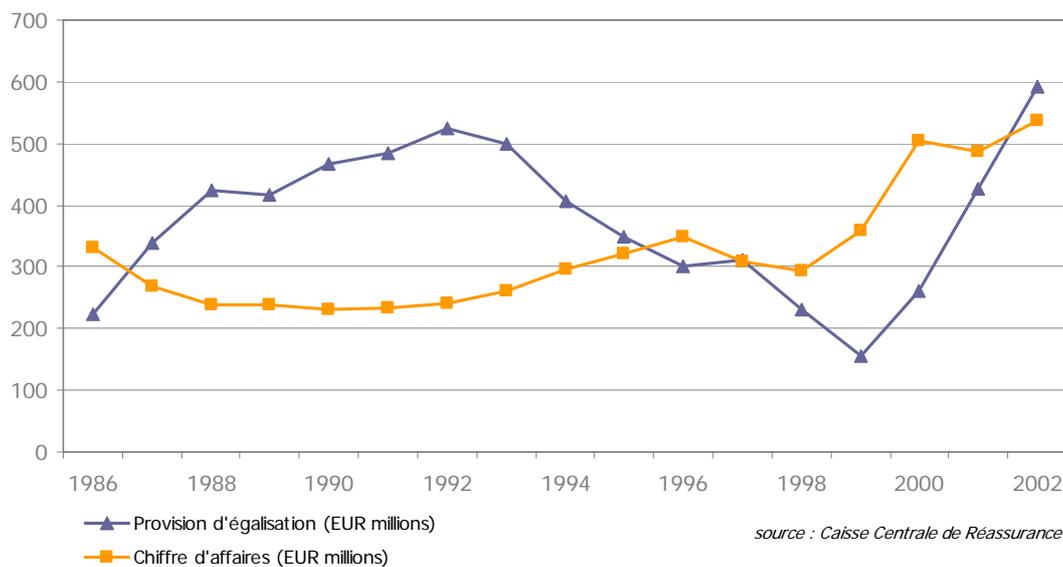
Le régime d'indemnisation des catastrophes naturelles en France tire son originalité et sa mixité de la nature des acteurs engagés financièrement. Les sociétés d'assurance, acteurs de l'économie de marché, effectuent la collecte des primes et le règlement des sinistres catastrophes naturelles. Elles peuvent se réassurer auprès de la Caisse Centrale de Réassurance (CCR), qui est « *habilitée à pratiquer les opérations de réassurance des risques résultant des catastrophes naturelles, avec la garantie de l'Etat* ». Cette garantie financière de l'Etat, visant à pallier une éventuelle insolvabilité de la CCR, assure le principe de solidarité du système.

Bien que la cédante conserve la liberté de se réassurer ou non, et ce auprès du réassureur de son choix, la CCR, forte de la garantie de l'Etat, domine le marché de la réassurance des catastrophes naturelles. Par l'intermédiaire de protections de réassurance de capacité illimitée, autrement dit en acceptant le transfert intégral du risque de sinistralité extrême, la CCR est en effet en mesure d'offrir aux cédantes une sécurité absolue, fort logiquement appréciée par ces dernières. En pratique, l'architecture des couvertures de réassurance proposées par la CCR est bâtie au moyen de deux formes de réassurance : une cession en Quote-Part au taux de 50%²⁵, et un Stop Loss sur rétention de portée illimitée (et de franchise généralement égale à 200 %), qui complète la Quote-Part.

Depuis sa création, le système a suscité de nombreux débats autour de sa viabilité. Plusieurs ajustements effectués ces dernières années ont notamment permis d'améliorer les réserves de la CCR, comme l'illustre le graphique suivant :

²⁵ L'unicité du taux de cession de 50%, ainsi que l'absence de commission de réassurance, ont été établies en 2000.

**Evolution du chiffre d'affaires et de la provision d'égalisation de la CCR
(EUR Millions) 1986-2002**



Les risques de catastrophes naturelles étant de mieux en mieux connus, des adaptations progressives permettent de maintenir le principe de solidarité du système. Toutefois, plusieurs années de sinistralité élevée seraient susceptibles de mettre à mal la rentabilité et la pérennité même du système.

Enfin, de nombreux acteurs reconnaissent l'exhaustivité du système français d'indemnisation des catastrophes naturelles, qui apporte une garantie généralisée à l'ensemble de la population et permet des indemnisations relativement rapides. On lui reproche cependant l'impact négatif qu'il peut jouer en matière de prévention, ainsi que l'absence de couverture des dommages corporels.

La loi du 25 juin 1990 a d'une part étendu le système des catastrophes naturelles aux Départements d'Outre-Mer, et d'autre part modifié le régime d'indemnisation des dommages résultant des effets du vent. Ces derniers, considérés comme « assurables », ont été exclus en 1990 du champ d'application de la loi de 1982, mais leur couverture est désormais systématiquement incluse dans chaque contrat d'assurance Incendie. Ainsi, la garantie des effets du vent, c'est-à-dire Tempêtes, Ouragans, Cyclones, fonctionne sur le mode contractuel, assureur et assuré s'accordant librement sur les conditions d'assurance.

En conséquence, les sociétés d'assurance, en matière d'évènements naturels hors catastrophes naturelles, ont recours plus librement au marché de la réassurance. Bien souvent, les traités de réassurance qui s'ensuivent offrent des garanties contre les dommages des tempêtes, de la grêle et du poids de la neige. On parle de programmes de réassurance TGN (Tempête, Grêle, Neige).

Bien que la forme proportionnelle soit encore assez répandue, le traité en excédent de sinistre par événement représente la protection traditionnelle la plus adaptée en réassurance TGN. On parle également de traité en excédent de sinistre catastrophe, en anglais « catastrophe excess of loss », et souvent abrégé Cat XL. La garantie accordée par les réassureurs dans le cadre d'une couverture Cat XL est généralement limitée. Afin que la cédante soit couverte en cas de survenance de plusieurs événements au cours de la période de validité du traité (le plus souvent une année), il est fréquent d'inclure des clauses de reconstitution de garantie. En contrepartie d'une prime additionnelle, celles-ci permettent de renouveler automatiquement la garantie (la portée), au fur et à mesure de son absorption. Les cédantes présentes sur le marché français consacrent un budget important aux couvertures TGN. Lors du renouvellement 2003, le montant des primes de réassurance des traités Cat XL s'est élevé à 484 millions d'euros, soit plus de 50 % de l'encaissement TGN perçu par les cédantes²⁶.

Enfin, parallèlement à ces deux principaux régimes d'indemnisation des dommages liés aux évènements naturels catastrophiques, le **Fonds National de Garantie des Calamités Agricoles** (créé en 1964) indemnise les exploitations agricoles victimes de dommages non assurables, et le **Fonds de Prévention des Risques Naturels Majeurs** (1995), dont la gestion est assurée par la CCR, verse des indemnités aux personnes expropriées de leurs biens en raison de la grave menace que ces dernières encourent, de par la présence avérée d'un risque naturel majeur (mouvements de terrain, avalanches ou crues torrentielles notamment).

²⁶ « Etude 2003 des Renouvellements des protections tempête du marché français ». Benfield Paris.

b. A l'étranger

La diversité du risque naturel (nature, fréquence, ampleur) à l'échelle internationale, combinée aux différences culturelles, financières et politiques entre les pays, est à la source de la multiplicité des solutions choisies à l'étranger en matière de couverture des risques de catastrophes naturelles. Les modalités des régimes d'indemnisation se distinguent selon le niveau d'intervention de l'Etat, et selon le caractère obligatoire ou facultatif de la protection.

Plusieurs pays européens ont opté, tout comme la France, pour un système d'assurance obligatoire, avec un degré d'engagement financier de l'Etat plus ou moins élevé. Le tableau suivant synthétise les principales caractéristiques des systèmes en vigueur dans ces pays :

Les régimes de couverture des catastrophes naturelles en Europe, avec intervention de l'Etat

Pays / Organisme	Evénements couverts	Evénements exclus	Biens garantis	Dommages garantis	Financement
Belgique / Fonds des Calamités	événements qui résultent de l'intensité anormale d'un agent naturel	ceux qui sont indemnisés par les clés d'assurance	risques simples	dommages directs seuls	Etat
Belgique (projet en cours)	séismes, inondations, glissements de terrain	tempêtes	risques simples, véhicules terrestres	dommages directs, perte d'exploitation	polices incendie, surprime de 9 %
Danemark / Stormflodsradet	inondations dues à une élévation du niveau de la mer	tempêtes, pluies torrentielles, dégel	tous les biens assurés contre l'incendie	dommages directs seuls	taxe forfaitaire annuelle par police incendie
Espagne / Consorcio de Compensacion	séismes, inondations, éruptions volcaniques, cyclones atypiques	neige, grêle, vent, pluie sauf intensité exceptionnelle	bâtiments à usage privé, industriel, agricole ou commercial	dommages directs seuls	Surprime proportionnelle à la valeur assurée
France / Caisse Centrale de Réassurance	événements qui résultent de l'intensité anormale d'un agent naturel	TOC	tous les biens assurés contre l'incendie	dommages directs, perte d'exploitation	Taxe 12 % sur polices dommages aux biens
Italie (projet en cours)	séismes, inondations, glissements de terrain, éruptions volcaniques, raz de marée		tout immeuble privé assuré contre l'incendie		surprime maxi 50 % de la prime incendie
Norvège / Norsk Naturskadepool	séisme, inondations, tempêtes, avalanches, éruptions volcaniques, raz de marée	gel, sécheresse, foudre	bâtiments à usage privé, industriel, agricole ou commercial	dommages directs seuls	Surprime proportionnelle à la valeur assurée
Suisse (19 cantons avec monopole d'assurance)	tempêtes, poids de la neige, grêle, crues, glissements de terrain, chute de pierres, avalanches	éruption volcanique, rupture de barrage, séismes	bâtiments à usage privé, industriel, agricole ou commercial	dommages directs seuls	Surprime proportionnelle à la valeur assurée

Source : SCOR

Dans de nombreux autres pays, tels l'Allemagne, le Royaume Uni, les Etats-Unis ou le Japon, le système d'assurance est de nature facultative, et repose à des degrés divers sur les principes libéraux. Il peut en résulter dans certains cas un degré de couverture moindre et des niveaux d'indemnisation plus faibles que dans les pays à système d'assurance obligatoire. La couverture des événements naturels y constitue toutefois un véritable marché compétitif où les assureurs s'efforcent de

mieux connaître les risques naturels et de maîtriser leurs expositions. La réassurance y est également très développée.

En fonction notamment de leur propre expérience des risques naturels, chaque pays a vu émerger des situations d'assurance et de réassurance assez particulières en matière de risques naturels. Pour l'ensemble des sociétés d'assurance, les protections de réassurance sous forme traditionnelle demeurent cependant de loin les plus utilisées. Mais, la connaissance et la perception du risque par les différents acteurs (états, assurés, assureurs et réassureurs) évoluent continuellement, et font évoluer avec elles les situations d'assurance et de réassurance, lorsque des dysfonctionnements apparaissent.

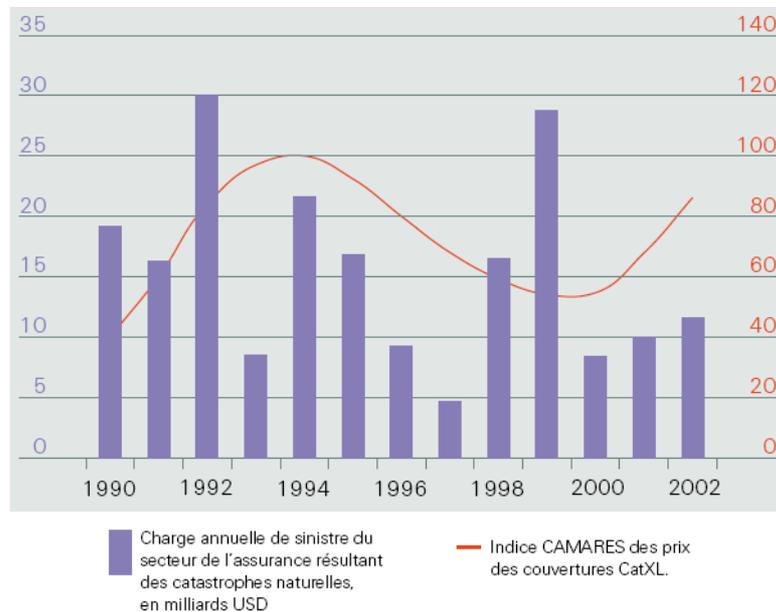
2.1.2 Les limites de la réassurance traditionnelle des risques naturels

a. De façon générale

L'utilisation de la réassurance traditionnelle ne semble pas totalement satisfaisante en matière de risques naturels. La difficulté à quantifier l'ampleur du risque naturel est telle que ce dernier peut dans certains cas être considéré comme non assurable. Jusqu'à présent, la seule réassurance traditionnelle, en tant que produit négociable sur un marché concurrentiel, n'a pas pu permettre aux sociétés d'assurance de proposer aux assurés une couverture optimale contre les périls naturels. Pour bon nombre de périls naturels, l'obstacle majeur réside dans la méconnaissance du risque et la difficulté de proposer des couvertures à des prix jugés acceptables. Les progrès réalisés dans la connaissance des risques, et le développement des méthodes de modélisation, semblent toutefois participer à un meilleur fonctionnement du marché de la réassurance traditionnelle des événements naturels. En France, la loi de 1990 portant sur les dommages résultant des effets du vent illustre cette évolution.

Mais le marché traditionnel de la réassurance des événements naturels reste sujet à des cycles tarifaires pluriannuels très marqués. La volatilité des prix des couvertures Cat XL est effectivement très liée à l'évolution de la charge sinistre, et dans une moindre mesure, à l'état des marchés de capitaux. Le graphique suivant,

issu d'une publication de Swiss Re²⁷, permet de mettre en évidence l'existence de cycles tarifaires.



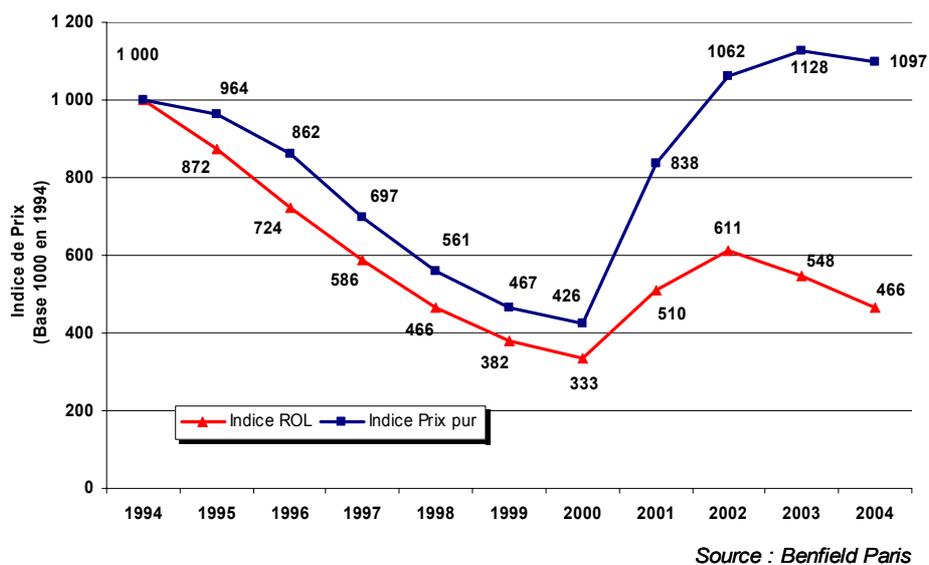
Source : Swiss Re 2003 - « Catastrophes naturelles et Réassurance »

La survenance d'un ou plusieurs sinistres majeurs au cours d'un exercice (par exemple les tempêtes hivernales en Europe en 1990 puis l'ouragan Andrew en 1992) entraîne lors des exercices suivants à la fois une croissance de la demande et une diminution de l'offre de réassurance. L'effet induit est alors une évolution à la hausse des prix des couvertures pendant quelques années. Par la suite, l'amélioration de la rentabilité du secteur, consécutive à l'absence de catastrophe, renouvelle l'offre de réassurance et accroît la concurrence sur le marché. Il s'en suit une inversion du cycle tarifaire (exemple en 1994) et le début d'une phase de diminution des prix. Le cycle s'inverse à nouveau quand surviennent de nouvelles catastrophes majeures (comme en 1999), et lorsque les marchés financiers s'effondrent (en 2001).

Sans l'ériger en théorie, l'existence de cycles tarifaires marqués est bien palpable dans l'expérience des quinze dernières années. L'évolution de l'indice des prix des protections tempête en excédent de sinistre sur le marché français depuis 1994²⁸ abonde également dans ce sens.

²⁷ « Catastrophes naturelles et réassurance ». Risk Perception - Swiss Re, 2003

²⁸ « Etude 2004 des Renouvellements des protections tempête du marché français ». Benfield Paris.



La présence de cycles tarifaires est bien évidemment un inconvénient de taille pour assureurs et réassureurs qui cherchent une stabilité de leurs résultats sur le long terme. Mais d'autres aspects de la réassurance traditionnelle sont source de difficultés en matière d'événements naturels.

Une couverture Cat XL est une protection par événement. La notion d'événement est donc fondamentale et fait l'objet d'une définition précise dans le traité. Elle est souvent définie au moyen d'une clause de limitation géographique et d'une clause de limitation dans le temps. Si la majorité des traités comportent aujourd'hui des clauses relativement consensuelles concernant la notion d'événement, le cadre de la définition n'est pas toujours approprié à la réalité des événements naturels. Pour le risque tempête en France, la clause de limitation dans le temps est généralement de 72 heures. Ainsi, Lothar et Martin, deux tempêtes distinctes d'un point de vue météorologique, étaient survenues entre le 25 et le 28 décembre 1999, et ont constitué deux événements pour les couvertures de réassurance alors en vigueur. Pour certains périls naturels, tels le gel ou la sécheresse, la notion d'événement est plus problématique, et les couvertures Cat XL traditionnelles sont assez peu adaptées.

Enfin, le placement des couvertures Cat XL demande de plus en plus d'investissement de la part des assureurs et des réassureurs. Une bonne utilisation des outils de modélisation des périls naturels requiert un important travail de

préparation des données liées au portefeuille et d'analyse des résultats, consommant donc des ressources non négligeables.

b. Du point de vue de la cédante

L'utilisation de couvertures traditionnelles contre les périls naturels présente des inconvénients particuliers pour les cédantes. Tout d'abord, il leur est difficile d'évaluer de façon optimale les caractéristiques de leur besoin de couverture. Une fois fixées la priorité et la portée de leur protection, une grande incertitude demeure quant à la distribution probable de la charge de sinistre qu'elles vont conserver sur l'exercice à venir. Mais il leur est surtout difficile d'arbitrer entre le coût de la réassurance qu'on leur propose et les bénéfices qu'elles peuvent en espérer.

Par ailleurs, compte tenu des sommes engagées, le marché de la réassurance des événements naturels présente un risque important de contrepartie. C'est bien évidemment en cas de survenance d'événements de grande ampleur qu'une cédante a le plus besoin de sa couverture de réassurance, mais qu'elle a simultanément le plus de chances de voir l'un de ses réassureurs lui faire défaut.

c. Du point de vue du réassureur

Si la cédante peut rencontrer des difficultés dans l'évaluation de son besoin de couverture, le réassureur est quant à lui confronté à la difficulté d'estimer la distribution de la sinistralité sur laquelle il s'engage. Par conséquent, la tarification de couvertures Cat XL contre les périls naturels s'avère délicate pour le réassureur, et force est de constater qu'il est souvent lui aussi dépendant de l'état du marché et de la situation du cycle tarifaire. Il lui est parfois problématique de renoncer à souscrire lorsqu'une situation de sous-tarification n'est pas reconnue. A ceci s'ajoute bien souvent une distorsion de l'équilibre concurrentiel, lorsque, à la suite d'une catastrophe, de nouvelles sociétés de réassurance sont créées, vierges de tout passif.

Par ailleurs, bien que de plus en plus réduite, une asymétrie d'information existe entre la cédante et le réassureur. Ce dernier ne connaît pas le portefeuille qu'il protège dans son intégralité, et doit faire confiance aux capacités de souscription et

de gestion de la cédante. En cas de sinistre majeur, il est également tributaire des évaluations réalisées par la cédante.

Sur un marché très concurrentiel, le réassureur doit faire preuve d'expertise et connaître avec précision le montant de capital que nécessitent ses souscriptions. C'est dans ce cadre que peut lui poser problème la couverture traditionnelle des risques naturels.

La réassurance traditionnelle des risques naturels comporte ainsi plusieurs inconvénients pour chacun des acteurs concernés, ce qui peut en partie expliquer qu'une part encore importante des dommages causés par les événements naturels ne soit aujourd'hui pas assurée. Afin de remédier à ces inconvénients, les acteurs de la réassurance ont développé de nouvelles formes de réassurance, de nature plus souvent complémentaire qu'alternative.

2.2 La réassurance non traditionnelle, le transfert alternatif des risques naturels

2.2.1 Définition de la Réassurance non traditionnelle

Il est difficile de donner une définition précise de la réassurance non traditionnelle, si ce n'est par opposition à la réassurance traditionnelle, dont les mécanismes ont été décrits dans le chapitre précédent. Le marché de la réassurance traditionnelle permet en effet aux cédantes de trouver des couvertures standards. La réassurance non traditionnelle englobe une très large gamme de produits qui répondent à des besoins très spécifiques des cédantes en matière de réassurance. Le marché de la réassurance non traditionnelle est inclus de façon plus générale dans le marché du transfert alternatif de risque, « alternative risk transfer » en anglais, ou ART.

Le terme ART permet de décrire à la fois le transfert de risque par le biais de structures alternatives, et le transfert de risque au moyen de produits alternatifs. Par structures alternatives, on entend principalement les captives ou les pools. La variété de produits alternatifs est quant à elle très large. Les produits de réassurance

financière, tels les traités à transfert de risque limité (« finite ») ou les solutions de transfert de portefeuille sinistre (« run-off »), en font notamment partie.

Les solutions ART se sont développées avant tout pour pallier les inconvénients des instruments traditionnels de transfert de risque, notamment lors des périodes où le coût de ces derniers s'avèrent trop onéreux. Bien que difficile à mesurer, le poids du marché ART est significatif. En 2000, le volume mondial des primes liées à des supports ART était estimé à 88 milliards de USD, alors que le volume des primes directes des produits traditionnels, pour le secteur commercial, s'élevait à 370 milliards USD²⁹.

2.2.2 Utilisation de la Réassurance non traditionnelle pour les risques naturels

Le transfert de risque lié aux périls naturels fait l'objet depuis quelques années d'un vif développement de produits de transfert alternatif de risque. La cause principale de cette évolution est le développement conjoint de l'offre et de la demande de nouveaux produits de transfert du risque naturel.

En effet, la survenance de nombreuses catastrophes naturelles dans les années 1990, notamment celle du cyclone Andrew³⁰ en 1992, est à l'origine d'un accroissement notable de la demande de couverture en assurance et réassurance contre les périls naturels. L'hiver 1997/1998, exceptionnellement doux aux Etats-Unis du fait d'El Niño, conduisit également des sociétés américaines du secteur de l'énergie à chercher à se protéger contre le risque climatique en général. La situation des marchés traditionnels de l'assurance et de la réassurance n'ayant pas toujours permis aux acteurs concernés de trouver des solutions acceptables, ils ont alors cherché à se protéger à l'aide de nouveaux types de couvertures.

Parallèlement, le développement des applications de la théorie financière moderne a amené de plus en plus d'investisseurs à s'intéresser à de nouvelles sources de risques, tels le risque d'assurance et le risque climatique. En effet, les développements technologiques, notamment en matière d'outils de tarification, ont

²⁹ « Transfert alternatif des risques (ART) : état des lieux ». Revue Sigma N°1/2003, Swiss Re

³⁰ Le cyclone Andrew dévasta une partie de la Floride en 1992, et on a recensé entre USD 20 et 25 Milliards de dommages assurés. La vitesse de vent a dépassé les 235 Km/h

favorisé cette évolution, rendant disponible des capitaux des marchés financiers pour la couverture de ces risques. Ces nouveaux placements répondent alors de l'application de la théorie de Markovitz sur la diversification des portefeuilles, puisque le risque d'assurance (notamment catastrophe) et le risque climatique sont en général indépendants des risques du marché. Par ailleurs, assureurs et réassureurs, s'appuyant également sur l'amélioration des outils de tarification, ont continué de développer leur métier de couverture des risques, et ont pris part au marché naissant de la protection contre le risque climatique en général.

La dernière décennie a donc vu le développement d'un marché du risque climatique et l'apparition d'une grande variété de protections. Parmi elles, les instruments dérivés climatiques, qui s'adressent à un très large éventail d'acteurs, dépassant le simple cadre de l'assurance, suscitent un vif intérêt. L'émergence d'un marché de titrisation du risque des catastrophes naturelles s'inscrit dans la même mouvance, mais constitue plus spécifiquement une solution alternative de réassurance.

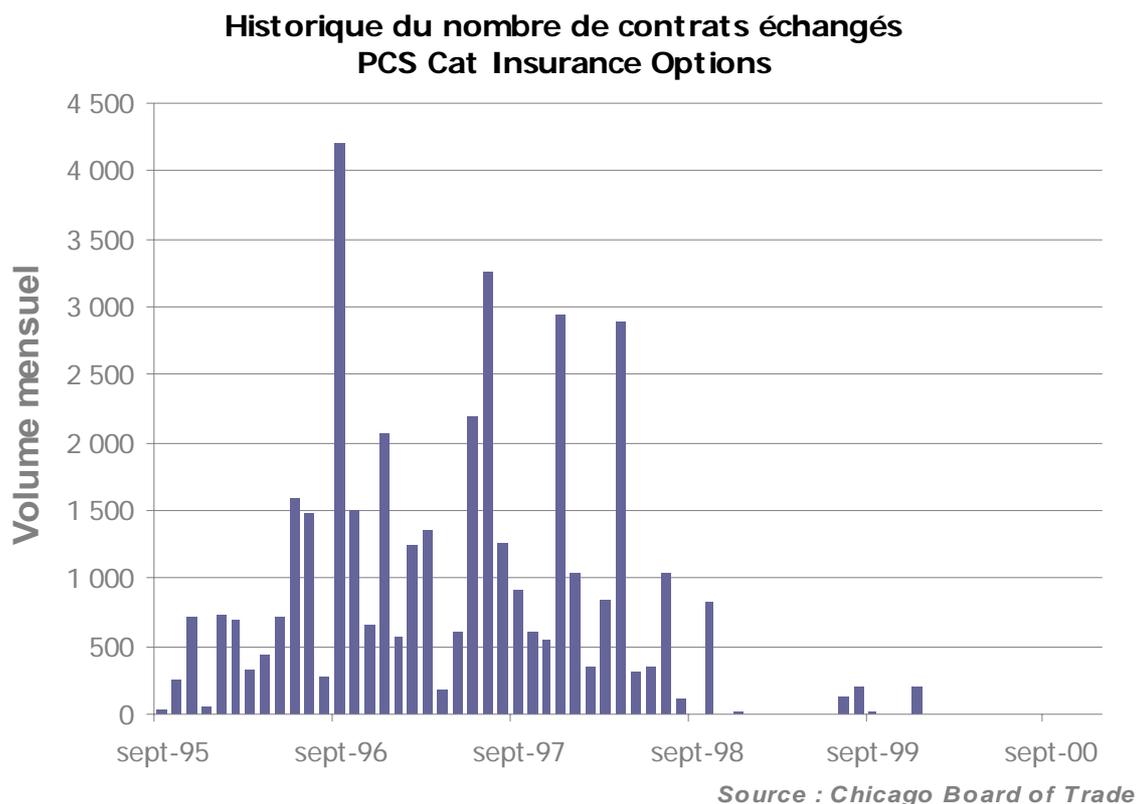
2.3 La titrisation du risque naturel, les Cat Bonds

A la suite du début difficile des années 1990 pour les acteurs de l'assurance et de la réassurance, le recours aux marchés des capitaux pour le transfert du risque d'assurance est devenu une alternative intéressante à la réassurance traditionnelle. Assureurs et réassureurs y ont vu la possibilité de compenser le tarissement de l'offre de réassurance par l'utilisation de la masse bien plus grande et bien plus liquide des capitaux provenant des marchés financiers.

La première initiative fut lancée par le Chicago Board of Trade (CBOT) en 1992, avec la création d'un marché organisé de produits dérivés de (ré)assurance, ayant pour sous-jacent des indices de sinistralité. En l'occurrence, la valeur de l'indice reposait sur la sinistralité moyenne des catastrophes naturelles pour un panel de sociétés d'assurance américaines³¹. Plusieurs problèmes techniques (notamment la faible transparence liée au caractère trimestriel de l'indice) furent responsables du

³¹ Il s'agissait du panel ISO (Insurance Service Office), constitué des 30 premières sociétés d'assurances américaines. L'ISO calculait quatre indices trimestriels différenciés selon les régions (East, West, Middle West et National), à partir des sinistres survenus au cours d'un trimestre donné et déclarés dans un délai de 6 mois, rapportés au quart des cotisations annuelles estimées.

très faible volume d'échanges de ces options, et conduisirent le CBOT à lancer en 1995 une nouvelle gamme de produits, les PCS³² Cat Insurance Option. Plus représentatifs de la sinistralité réelle du marché (neuf indices régionaux couvrant 70 à 85 % du marché de l'assurance américain) et publiés sur une base quotidienne, ces nouveaux produits n'ont pourtant pas connu le succès escompté. Après un démarrage très modéré, le négoce de ces contrats a été stoppé à la fin de l'année 2000, comme l'illustre le graphique suivant.



Si le recours aux marchés financiers au moyen de titres d'assurance (en anglais insurance-linked securities, ILS) a été utilisé dans plusieurs domaines (comme la réassurance Vie par exemple), la majorité des innovations et des échanges porte sur le transfert du risque naturel, catastrophique principalement. Près de la moitié des transactions de titrisation du risque d'assurance concerne les obligations catastrophe, également appelées Cat Bonds.

Les premières émissions de cat bonds sont apparues aux Etats Unis, suite à la prise de conscience des insuffisances du marché de l'assurance et de la

³² PCS : Property Claims Service

réassurance après la survenance du cyclone Andrew. Face à la probabilité réelle d'un événement encore plus catastrophique et à l'absence sur le marché traditionnel des capitaux nécessaires pour l'assumer, assureurs et réassureurs ont trouvé dans les cat bonds un moyen complémentaire de couverture du risque catastrophe³³.

Depuis, l'utilisation des cat bonds s'est répandue à l'échelle mondiale, pour couvrir principalement les risques de tremblement de terre en Californie ou au Japon, d'ouragan en Floride, ou de tempête en Europe. Il n'existe pas de marché secondaire organisé pour les cats bonds, ces derniers demeurant donc des produits négociés de gré à gré (OTC).

Bien que chaque cat bond soit caractérisé par certaines spécificités, le principe de base est l'émission par une cédante d'une dette obligataire dont les modalités de remboursement et de rémunération dépendent de la réalisation d'un risque prédéfini ou de la survenance d'un sinistre, généralement suite à un événement catastrophique. L'élément qui détermine les caractéristiques de l'éventuelle indemnisation de la cédante peut ainsi être défini de plusieurs façons, comme le montant des pertes subies par la cédante lors d'un certain type d'événement naturel, ou bien à l'aide d'un indice assurantiel ou d'indice météorologique³⁴.

Si la cédante peut directement émettre ses titres auprès des investisseurs, la structure d'un cat bond comporte bien souvent un intermédiaire, le SPV (special purpose vehicle)³⁵, notamment en raisons d'avantages fiscaux. Le rôle de cet intermédiaire est comparable à celui d'une captive ou d'un pool. Il reçoit les dépôts des investisseurs correspondant au nominal de l'obligation. Ces sommes sont alors bloquées et placées sur des supports sûrs à rendement certain³⁶. Parallèlement, la cédante verse une prime au SPV au titre d'un contrat de réassurance qui les lie. En cas de non survenance de l'événement catastrophique déclencheur, les investisseurs auront, à la maturité de l'obligation (généralement entre 1 et 5 ans), récupéré leur capital initial et les intérêts du placement, mais également perçu la prime de risque versée par la cédante. Ceci leur assure un rendement plus élevé,

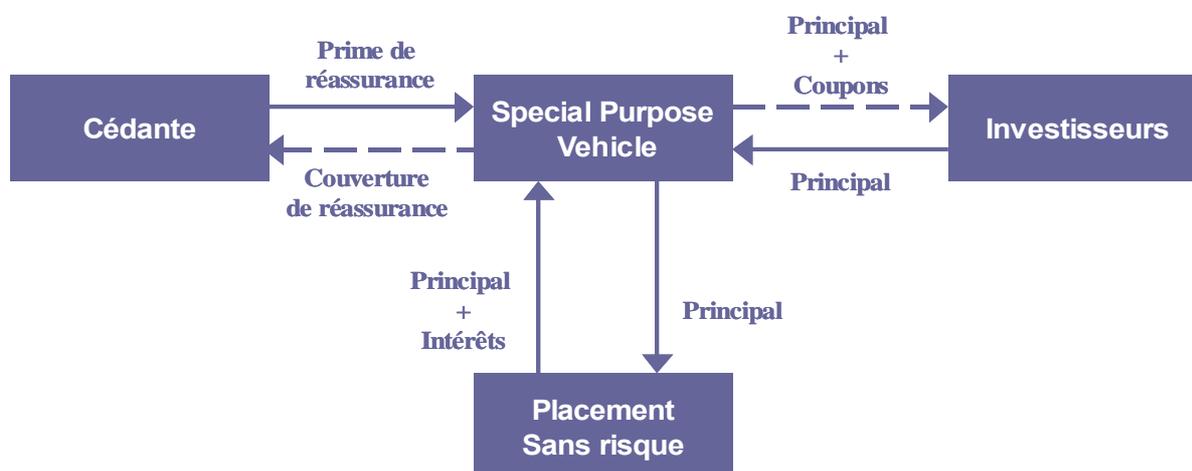
³³ Les Cats Bonds ne sont pas réservés aux seuls assureurs et réassureurs. Ils peuvent également assurer le transfert du risque catastrophe de grandes compagnies.

³⁴ Lorsqu'il s'agit d'un indice, on parle de Cat Bond paramétrique.

³⁵ L'intermédiaire peut également être un SPR (special purpose reinsurer).

³⁶ Obligations du trésor ou obligations de corporate notées AAA et suffisamment liquides.

conforme au risque pris en cas de survenance de l'événement catastrophique. Dans ce cas, le SPV utilise tout ou partie des sommes immobilisées (nominal et/ou coupons) afin d'indemniser la cédante. Le schéma suivant synthétise le principe général d'un cat bond.



La légitimité des cat bonds réside donc pour les assureurs dans l'accès à une nouvelle source de financement du risque de catastrophe naturelle, les marchés des capitaux, dont la capacité potentielle est gigantesque. Grâce aux caractéristiques de leur structuration, les cat bonds ont également pour avantage principal d'éliminer le risque de contrepartie, qui peut être très préjudiciable avec les couvertures traditionnelles de réassurance. Par ailleurs, ils offrent aux cédantes la possibilité d'établir une véritable protection sur mesure, pour une durée supérieure à l'année si nécessaire.

Toutefois, plusieurs inconvénients expliquent certainement que l'utilisation des cats bonds ne soit pas aussi répandue que celle à laquelle on aurait pu s'attendre. Tout d'abord, dans la majorité des structurations de cat bonds, la cédante conserve un risque de non corrélation (le « risque de base »), entre la valeur réelle de son sinistre et le montant de l'indemnisation qu'elle perçoit en cas d'événement déclencheur³⁷. Les protections de réassurance traditionnelles ne présentent pas ce désagrément, dans la mesure où elles reposent sur le principe indemnitaire, l'indemnisation de la cédante étant directement liée au montant de son sinistre. De plus, les coûts de structuration (conseil, intermédiation...) liés à l'émission des cat

³⁷ C'est le cas des cat bonds paramétriques.

bonds impliquent un coût total qui est souvent très élevé en comparaison de la réassurance traditionnelle. Enfin, les délais de mise en place sont relativement longs.

2.4 Autres solutions alternatives à la réassurance traditionnelle

En parallèle des cat bonds, plusieurs protections novatrices se sont présentées comme alternatives potentielles à la réassurance traditionnelle, et plus particulièrement celle relative aux événements naturels. Toutefois, le volume des transactions est pour l'instant assez limité, leur utilisation restant exceptionnelle.

Parmi elles, les swap catastrophe permettent à la cédante de participer à un échange de paiements fixes prédéfinis contre des paiements variables dont le montant est fonction de la survenance d'une catastrophe. Plus souples dans leur structuration et plutôt moins onéreux que les cat bonds, ils présentent cependant un risque de contrepartie.

Les industry loss warranties (ILW) sont des protections structurées de façon similaire à un contrat de réassurance traditionnel. La cédante se voit apporter la garantie d'une indemnisation conditionnée par un mécanisme de double déclenchement : les sinistralités enregistrées lors d'un événement, d'une part par l'ensemble des assureurs d'un marché donné, et d'autre part par la cédante elle-même, doivent toutes deux dépasser un certain seuil défini au préalable. Ces protections sont toutefois conçues de telle sorte que l'élément déterminant est bien souvent le niveau du seuil concernant les dommages subis par le marché dans son ensemble.

Enfin, le recours à des contrats de capital contingent constitue une alternative assez différente dans son principe, puisqu'il ne s'agit plus pour la cédante d'acquérir la garantie d'une indemnité. Un instrument de capital contingent correspond en fait à une option, achetée par la cédante, qui lui permet de trouver une solution de financement en cas de survenance d'un événement déterminé. Selon les cas, il peut s'agir d'acquérir le droit d'émettre de la dette ou des actions, à un niveau de prix fixé à l'avance. Si ce type d'instrument donne l'accès en cas de besoin à des solutions de financement plutôt bon marché, il ne permet pas un véritable transfert du risque.

Le potentiel des instruments de transfert du risque d'assurance aux marchés financiers, au premier plan desquels les produits de titrisation du risque naturel et les cat bonds, est souvent perçu comme substantiel. Les innovations se multiplient et les volumes progressent, quand bien même légèrement. Toutefois ces solutions ne se posent pas jusqu'à présent en véritable alternative aux couvertures de réassurance traditionnelles. En ce qui concerne les événements naturels, la plus grande partie des volumes réalisés concerne des produits utilisés avant tout par les cédantes comme moyen complémentaire de couverture contre les très gros risques de catastrophe.

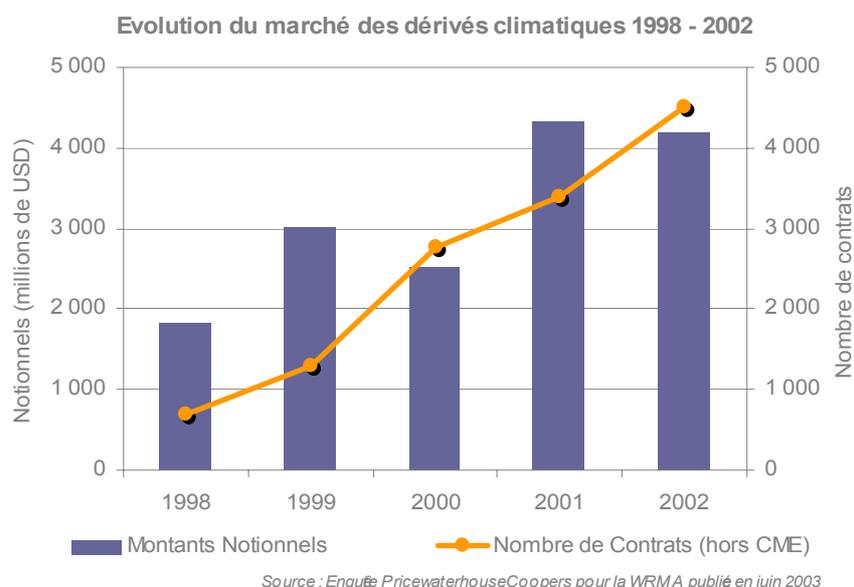
2.5 Les instruments dérivés climatiques

Ces dernières années ont vu le développement d'un nouveau marché, celui des produits dérivés climatiques. Un dérivé climatique est un nouvel actif de type Arrow Debreu, qui se caractérise par des paiements prédéfinis en cas de survenance d'événements climatiques prédéfinis. L'actif sous jacent d'un dérivé climatique est un indice météorologique, ce qui distingue notablement ce dernier des produits dérivés plus traditionnels dont les actifs sous jacents ont une dimension pécuniaire (action, indice boursier, taux d'intérêt, taux de change ou matière première...).

Les premières transactions de produits dérivés climatiques ont été conclues de gré à gré aux Etats-Unis en 1997 par des sociétés du secteur énergétique (Aquila, Enron et Koch). Pour ces sociétés, une saison hivernale trop douce ou une saison estivale trop fraîche sont synonymes de chute du chiffre d'affaires, puisque la demande de leurs clients est amoindrie. Dans le contexte de la déréglementation du marché américain de l'énergie, elles purent composer plus librement la gestion du risque climatique qui pèse sur leurs revenus. Dans un premier temps, des futures et des forwards furent utilisées afin de protéger leurs revenus d'activités contre les fluctuations des prix. Les premiers dérivés climatiques leur permirent par la suite de couvrir les fluctuations du volume liées à des températures, selon la saison, ou trop faibles ou trop élevées.

Depuis 1997, le marché des dérivés climatiques a connu une croissance non négligeable. Le phénomène El Niño, qui rendit l'hiver 1997-1998 exceptionnellement doux aux Etats Unis, participa sans nul doute au vif intérêt adressé au dérivés

climatiques de température. Selon l'enquête réalisée par PricewaterhouseCoopers et publiée en juin 2003 pour la Weather Risk Management Association (WRMA)³⁸, le volume annuel des transactions de produits dérivés climatiques s'élevait à plus de 4 milliards de USD en 2001 et 2002. Parallèlement, le nombre annuel de contrats progressait quant à lui de 695 à 4 517 entre 1998 et 2002³⁹.



Bien que les dérivés climatiques échangés jusqu'à présent aient concerné plusieurs types de variable climatique, les produits basés sur la température représentent depuis l'origine plus de 90 % des transactions, tandis que ceux basés sur les précipitations (pluie, neige), sur le vent ou autres restent marginaux. Par ailleurs, la plus grande partie des transactions demeure sur le marché américain, même si ces deux dernières années ont vu un développement notable en Europe et en Asie.

La diversité des acheteurs de dérivés climatiques tend à s'accroître, mais la part des acteurs du secteur de l'énergie reste prépondérante. Plusieurs analyses s'accordent à montrer le rôle éminent joué par le risque climatique sur les revenus de

³⁸ « *The Weather Risk Management Industry - Survey Findings for April 2002 to March 2003* ». Troisième publication de l'enquête réalisée par PricewaterhouseCoopers pour le compte de la Weather Risk Management Association (WRMA), basée sur les réponses à un questionnaire de 19 sociétés internationales. (20 sociétés pour l'étude précédente).

³⁹ Ces chiffres n'incluent pas les dérivés climatiques échangés sur le Chicago Mercantile Exchange (CME).

la plupart des activités économiques⁴⁰, et les dérivés climatiques s'adressent potentiellement à nombre de secteurs aussi variés que l'agriculture, l'agroalimentaire et les boissons, le tourisme et les loisirs, la construction, les transports ou l'habillement. Banques et assureurs sont quant à eux les principaux preneurs de risque sur ce nouveau marché, qui nécessite à la fois compétences financières et expérience de gestion des risques.

Les dérivés climatiques prennent la forme des instruments dérivés plus traditionnels. Cap, floor, collar ou swap sont utilisés pour transférer le risque météo par le biais d'un produit dérivé. Les transactions OTC, à l'origine du développement de ces produits, représentent encore la plus grosse part du volume. Le premier marché organisé fut créé par le Chicago Mercantile Exchange en février 1999. Le CME a lancé le 22 septembre 1999 une gamme de futures et d'options portant sur les indices de température HDD (Heating Degree Day Index) et CDD (Cooling Degree Day Index) mesurés dans dix villes (Chicago, Cincinnati, New York, Atlanta, Philadelphie, Dallas, Des Moines, Las Vegas, Tucson et Portland)⁴¹. Les contrats futures sont liquidés par cash-settlement sur la base de l'indice CME Degree Day Index, calculé par EarthSat⁴² et comportent 12 échéances mensuelles. Les Degree Day Index de chaque ville sont des indices quotidiens de température moyenne calculés par rapport à une température de référence de 65° Fahrenheit. Le HDD permet de mesurer, en nombre de degrés Fahrenheit, la relative fraîcheur d'une journée par rapport à cette température de référence, le CDD la relative chaleur.

$$\text{HDD} = \text{Max} (0 , 65 - \text{température moyenne enregistrée sur la journée})$$

$$\text{CDD} = \text{Max} (0 , \text{température moyenne enregistrée sur la journée} - 65)$$

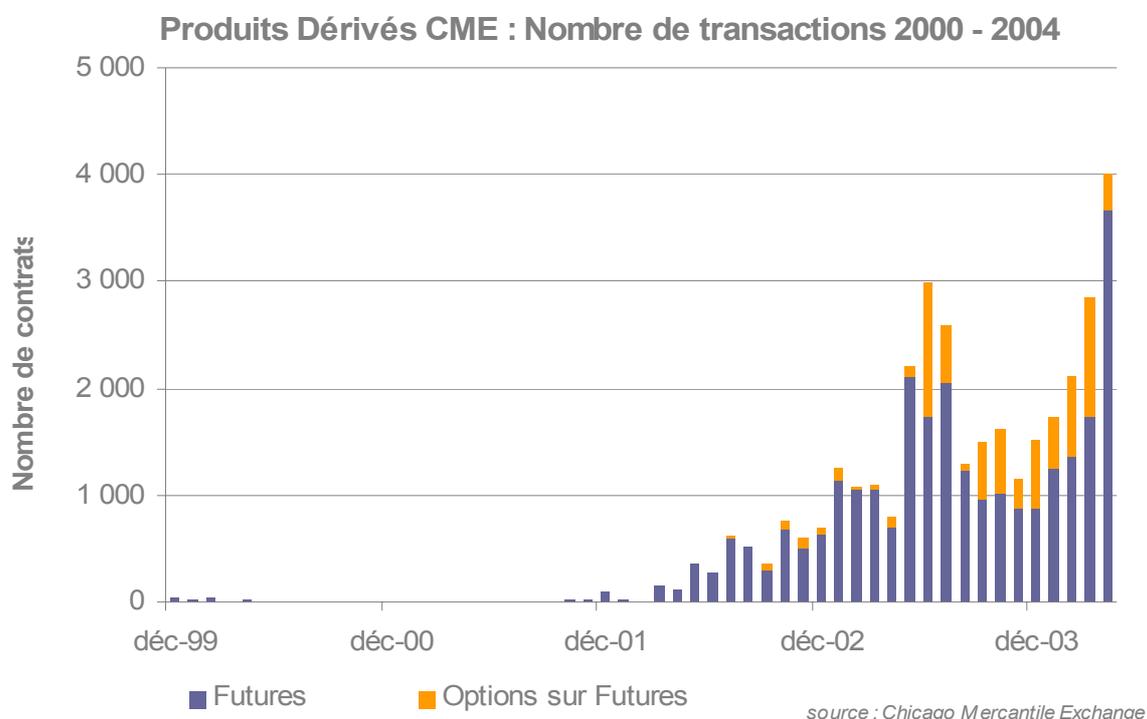
Les futures et options portent alors sur le cumul des unités de HDD ou de CDD sur une certaine période, généralement comprise entre le mois et l'année. Les futures sont cotées en points d'indices (HDD ou CDD), et leur notionnel équivaut à 100 USD multipliés par la valeur de l'indice, ce qui autorise dans la pratique des transactions de faible taille. Les transactions négociées sur le marché du CME ont connu un véritable essor à partir de 2002, et semblent aujourd'hui en pleine

⁴⁰ Selon les sources, les conditions météorologiques exerceraient une influence sur une part comprise entre 10% et 50% du PIB des Etats du G10 !

⁴¹ Les transactions des futures et options sur CDD furent lancées le 30 septembre 2000.

⁴² Earth Satellite Corporation.

progression, comme l'illustre l'historique du nombre mensuel de contrats entre décembre 1999 et avril 2004.



Porté par l'enthousiasme, le CME élargit progressivement la variété des produits proposés, et a lancé en octobre 2003 de nouveaux dérivés climatiques basés sur les températures quotidiennes moyennes enregistrées dans cinq villes européennes, Amsterdam, Berlin, Londres, Paris et Stockholm.

Des initiatives de création de marché secondaire de dérivés climatiques ont également été prises par d'autres institutions. L'Intercontinental Exchange (ICE), basé à Atlanta, créait en novembre 2001 un marché organisé de swaps basés sur des indices de température moyenne pour cinq grandes villes américaines⁴³. En décembre 2001, le LIFFE, à Londres, lançait également des futures sur des indices de température pour les villes de Paris, Londres et Berlin. Enfin, Euronext et Météo France sont partenaires pour développer le marché des produits dérivés climatiques en Europe via Euronext. Une série d'indices climatiques, baptisée NextWeather, a été conçue à ce propos, et devrait notamment permettre d'élargir le type de risque

⁴³ Chicago, Dallas, New York, Philadelphie et Sacramento. Ce marché de swaps vise plus particulièrement les acteurs du secteur de l'énergie.

couvert au delà de la température, grâce à des indices portant par exemple sur la pluie ou la neige.

Le marché des dérivés climatiques dans son ensemble suscite un vif intérêt et semble promis à un bel avenir. L'organisation progressive de marchés secondaires liquides et sans risque de contrepartie devrait contribuer au développement de ce nouveau type d'instruments, en rendant ces derniers accessibles pour un plus grand éventail d'acteurs. Les avancées en matière de tarification de ces produits seront également déterminantes pour leur avenir.

Si les dérivés climatiques ne constituent pas à proprement parler une alternative directe à la réassurance des grands risques naturels, ils représentent toutefois une nouvelle génération d'instruments de transfert du risque naturel. Ces produits revêtent d'ailleurs un caractère assez particulier, dans la mesure où ils se situent à la frontière entre contrats d'assurance et instruments financiers. Les deux formes juridiques, contrat d'assurance et instrument financier à terme, ont en effet été utilisées pour monter ce type de couverture.

Les dérivés climatiques sont jusqu'à présent essentiellement utilisés par des entreprises qui souhaitent couvrir des pertes de marge ou de chiffre d'affaires contre des variations de la température moyenne par rapport à des normales saisonnières. Mais le mécanisme peut être transposé pour structurer des protections, non plus contre des risques météorologiques « ordinaires », mais contre des événements naturels exceptionnels, créant ainsi une forme de couverture contre le risque naturel catastrophique, alternative à la réassurance traditionnelle. Ce type de couverture fait l'objet des chapitres suivants.

3 PRINCIPES DE LA COUVERTURE INDICIELLE CAT

3.1 L'origine de la couverture indicielle Cat

La couverture indicielle Cat est un dérivé climatique : l'actif sous-jacent sur lequel elle s'appuie est un indice dont la valeur dépend exclusivement de données climatiques, comme par exemple la température, la pluviométrie, la vitesse de vent ou l'ensoleillement.

3.1.1 Une vocation spécifique

La première « couverture indicielle Cat » a été conçue et placée en 2001 par le courtier en réassurance Benfield Paris. Elle visait à protéger le portefeuille d'assurance d'une cédante exposée au risque tempête en France, en se basant sur un indice calculé à partir des vitesses de vent maximales, enregistrées quotidiennement dans un panel de stations météorologiques du territoire français⁴⁴.

Les dérivés climatiques « classiques », c'est-à-dire du type HDD ou CDD, peuvent être décrits avant tout comme des instruments financiers à terme permettant la couverture du risque financier lié à la corrélation entre revenus et variations autour de la normale de paramètres météorologiques. La couverture indicielle Cat est quant à elle une protection de réassurance qui repose sur le transfert du risque financier lié à la corrélation entre la sinistralité d'un portefeuille d'assurance et les variations exceptionnelles de la météo. Elle vise à offrir aux acteurs du marché de l'assurance une protection contre des événements naturels catastrophiques, contre des événements rares, mais caractérisés par une grande intensité. Ceci correspond particulièrement bien au risque tempête à l'échelle de la réassurance.

En conséquence, la valeur de l'indice doit avant tout refléter la survenance de valeurs extrêmes des variables météorologiques, correspondant par hypothèse à la survenance de sinistres chez le détenteur d'un portefeuille de contrats d'assurance. L'indice climatique n'est donc pas, dans le cadre de la couverture indicielle Cat, défini

⁴⁴ « Benfield Greig, courtier en réassurance, et l'assureur Generali France ont mis en place un nouveau mode de couverture des risques climatiques. Le système consiste à indemniser à l'aide d'un indice météorologique, reposant sur des critères liés au vent, et non au sinistre réel.[...] » La Tribune, 12 mars 2002.

et construit afin de décrire l'écart par rapport au comportement moyen d'un paramètre météorologique, comme le sont par exemple les indices HDD et CDD vis-à-vis de la température. Il est structuré de façon à dépeindre les divergences exceptionnelles du comportement d'un paramètre météorologique.

3.1.2 Un produit de réassurance

La couverture indicielle Cat est une forme alternative ou complémentaire à la réassurance traditionnelle. Elle est substantiellement un produit de gré à gré, qui ne pourrait que difficilement faire l'objet d'un marché organisé. Ses caractéristiques, notamment la construction de l'indice climatique sous jacent, sont conçues pour répondre au besoin spécifique d'une cédante en matière de réassurance contre les périls naturels. Sa forme juridique est généralement celle d'un contrat de réassurance, puisqu'elle vise à transférer un risque de sinistralité pour la cédante. Elle peut néanmoins prendre la forme d'un contrat financier, par le biais du contrat cadre défini par l'ISDA⁴⁵.

La couverture indicielle est une protection de réassurance dont le prix dépend uniquement de la distribution d'un indice basé sur des variables climatiques. Ceci peut constituer, à la fois pour la cédante et pour ses réassureurs, un avantage par rapport à une protection traditionnelle, dont une partie de la prime repose généralement sur les incertitudes concernant l'exposition réelle du portefeuille de la cédante. Par ailleurs, les mécanismes de la couverture indicielle se révèlent très flexibles, et le principe de la protection peut permettre à une cédante d'obtenir une couverture contre des périls naturels difficilement réassurables sur le marché traditionnel⁴⁶.

Les principes de fonctionnement de la couverture indicielle permettent également d'éviter certains désagréments que peuvent parfois présenter les couvertures traditionnelles en excédent de sinistre par événement. Sont ainsi esquivés les éventuels litiges autour de la définition de l'événement en cas de survenance d'un sinistre. Les problèmes liés à l'estimation des sinistres à la charge de la réassurance et aux délais d'indemnisation de la cédante sont également évités.

⁴⁵ International Swaps and Derivatives Association

⁴⁶ Pour certains périls naturels, comme par exemple le gel ou la sécheresse, les cédantes peuvent rencontrer un réel manque d'appétit de la part des réassureurs.

Enfin, la nature de la couverture permet le double accès aux marchés des capitaux et au marché de la réassurance.

La couverture indiciaire Cat présente bien sûr une très forte analogie avec les dérivés climatiques évoqués dans le chapitre précédent : l'utilisation d'un paramètre météorologique, la construction d'un indice, et la mécanique d'un instrument à terme. Toutefois, sa vocation est différente, et elle comporte de ce fait des caractéristiques spécifiques. Le mécanisme de la couverture indiciaire est présenté plus en détail dans le sous chapitre suivant.

3.2 Fonctionnement de la couverture indiciaire Cat

3.2.1 Utilisation de séries de données climatiques

Tout comme les dérivés climatiques négociés sur les marchés financiers ces dernières années, la couverture indiciaire Cat est un contrat caractérisé par des flux financiers dont le montant est fonction de la valeur d'un indice construit à partir de l'observation de données météorologiques.

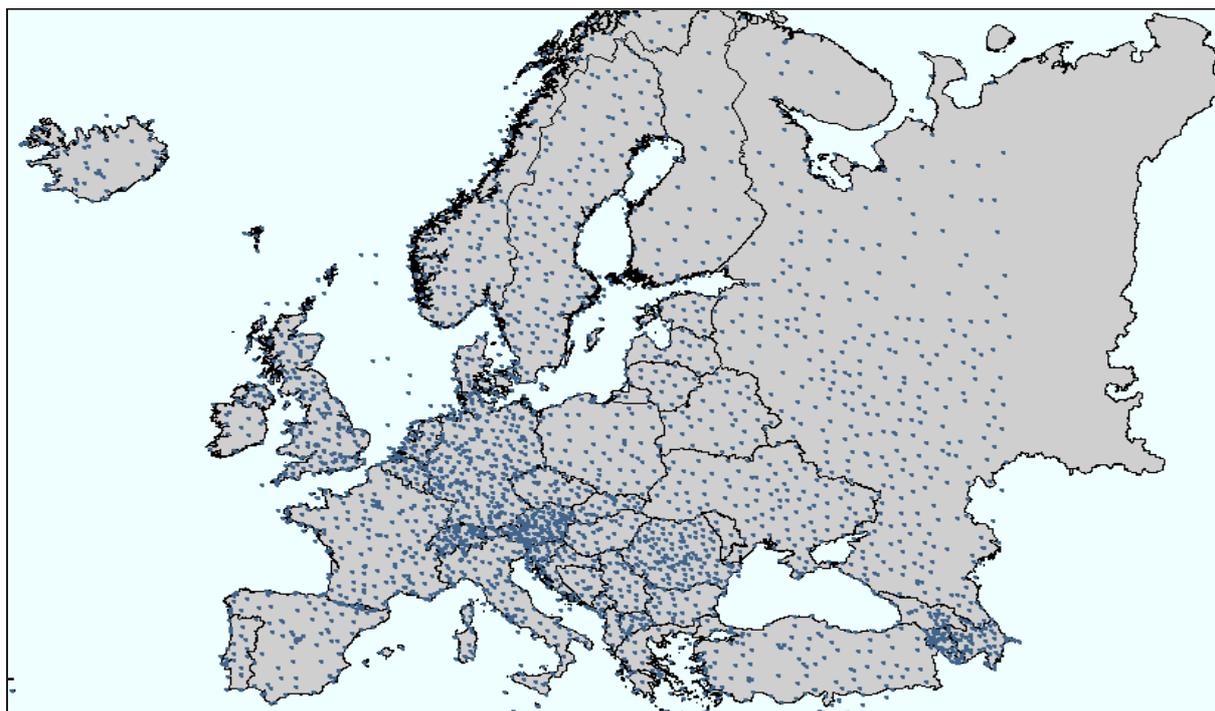
Ces données sont recueillies par les organismes météorologiques nationaux, disposant sur l'ensemble de leur territoire respectif de réseaux de stations assurant (entre autres)⁴⁷ la mesure des variables climatiques.

Ainsi, en France métropolitaine, Météo France dispose d'un réseau de stations d'observation comprenant environ 140 stations principales auxquelles s'ajoutent près de 3 000 stations auxiliaires (ne relevant que quelques paramètres). Au Royaume Uni, le UK Met Office s'appuie sur un réseau d'environ 200 stations principales, et il y a près de 5 000 stations auxiliaires. La carte ci-dessous permet de visualiser le réseau des stations européennes homologuées par l'Organisation Météorologique Mondiale⁴⁸ :

⁴⁷ Les stations météorologiques sont définies dans le « Dictionnaire du Climat », Larousse, comme des « *lieux et installations où sont régulièrement effectuées des mesures de l'atmosphère* ». Hormis cette fonction de mesure et d'observation, des travaux d'analyse et de compréhension, ainsi que de prévision des phénomènes météorologiques peuvent également y être réalisés.

⁴⁸ La WMO, World Meteorological Organization créée en 1950, utilise actuellement plus de 3 000 stations synoptiques en Europe.

Réseau météorologique européen – stations WMO



Les stations françaises proposent bien souvent des historiques de données météorologiques sur les cinquante dernières années, et les plus anciennes stations offrent des historiques longs de plus d'un siècle. Météo France possède à la fois une mission publique et une vocation commerciale. L'utilisation de la Climathèque est un service commercial proposé par Météo France, qui donne accès à une grande variété de données climatiques. Les données mesurées dans les stations sont enregistrées en temps réel, ce sont des données « synop » qui sont disponibles immédiatement. Un premier contrôle automatique permet de détecter certaines valeurs aberrantes, puis une vérification humaine peut avoir lieu. Si un problème est décelé pour une donnée, suite à une panne d'appareil par exemple, la valeur aberrante est effacée. Si les valeurs enregistrées sont validées, elles deviennent alors des données climatologiques. Les données climatologiques sont publiées par Météo France au bout de 48 heures. Selon les pays, ce délai peut être plus ou moins long, il est par exemple de deux mois pour les données britanniques, publiées par le UK Met Office.

Les données utilisées pour construire l'indice d'une couverture indicielle Cat sont des données climatologiques. Ce sont principalement trois caractéristiques de la donnée climatologique qui sont importantes dans le cadre d'une couverture indicielle Cat. Ces caractéristiques ont trait aux questions suivantes : Où ?, Quoi ? et Quand ?.

La localisation de la station où est mesuré le paramètre météorologique est un élément déterminant. La nature du paramètre météorologique observé en est bien évidemment un autre. La variété des paramètres météorologiques observés et enregistrés est grande : précipitations, température, force du vent, pression, humidité, insolation ou rayonnement par exemple. Enfin, la troisième caractéristique concerne l'unité temporelle de mesure, une donnée météorologique pouvant effectivement être appréciée à l'échelle d'une seconde, d'une heure, d'un jour ou d'un mois...⁴⁹

Les données climatologiques historiques sont utilisées en amont pour créer la structure d'une couverture indicielle, et les données climatologiques présentes sont utilisées pour connaître la valeur d'un indice en cours. La notion d'indice est définie plus avant dans la sous-partie suivante.

3.2.2 L'indice : définition, utilisation

La mise en place d'une couverture indicielle Cat s'articule autour de plusieurs grands axes :

- le choix du type de données météorologiques utilisées pour la définition de l'indice,
- le choix d'un panier de stations où les données météorologiques vont être étudiées et enregistrées,
- et la définition des modalités de calcul de l'indice lui-même et de l'indemnisation monétaire qui en découle.

a. Choix du type de données

Il s'agit là de déterminer quelle variable climatique permet de représenter de façon fidèle et adéquate l'occurrence et l'ampleur du péril naturel contre lequel la cédante souhaite se doter d'une protection de réassurance.

⁴⁹ Nous présentons en annexe 1 un exemples de données mesurées dans une station météorologique du réseau de Météo France.

Compte tenu des caractéristiques de mesure de la variable climatique (notamment ce qu'elle représente, l'unité temporelle de la mesure), l'éventail du type de données utilisables est virtuellement infini. Par exemple, si l'on s'intéresse au froid et aux basses températures, on peut choisir de considérer :

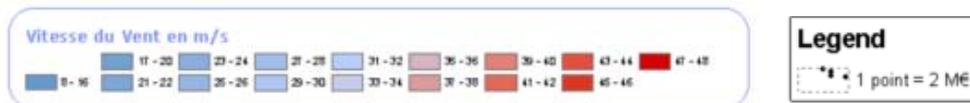
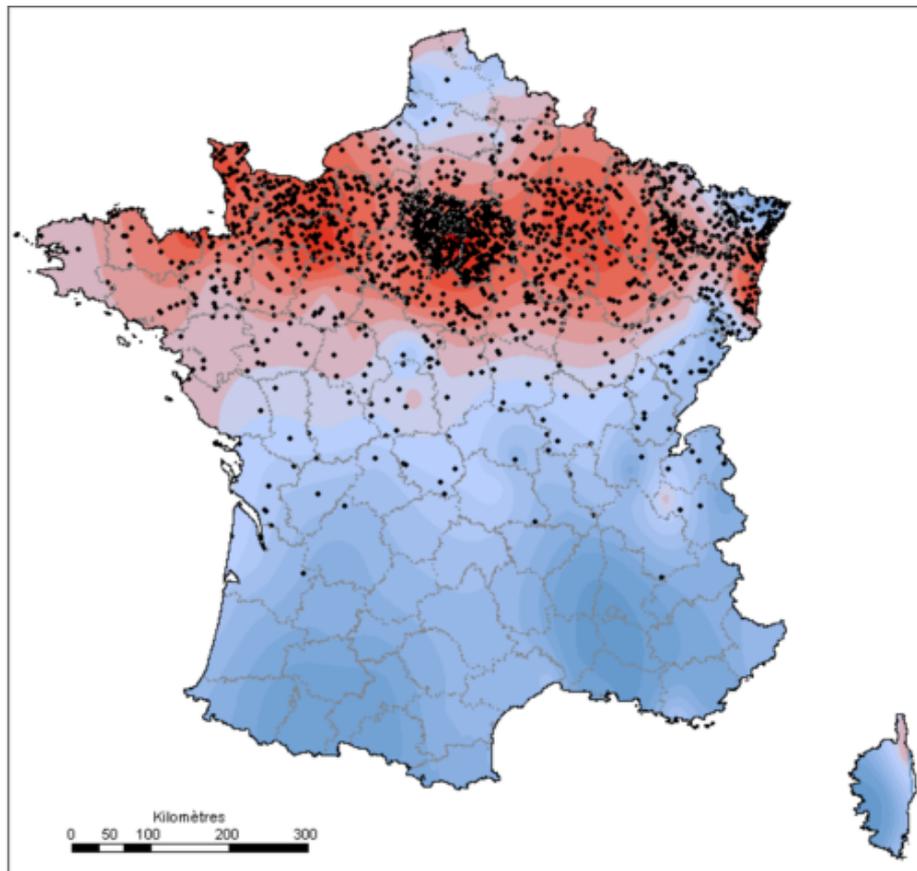
- la température minimale horaire sous abri,
- la température minimale horaire à l'air libre,
- la température minimale quotidienne,
- ou bien la moyenne des températures minimales horaires sur une journée,
- ou bien encore le nombre d'heures d'une journée où la température horaire est en dessous de 0°C,
- etc.

Toutefois, l'éventail des possibilités est tout d'abord et bien évidemment restreint par l'étendue de l'offre de données proposées par les organismes météorologiques. De plus, d'un point de vue pragmatique, il ne paraît pas judicieux de retenir un type de données qui, d'une part entraîne pour l'acteur en charge de la structuration de la couverture une complexité inutilement onéreuse (coût d'acquisition des données, mobilisation des ressources humaines et informatiques), et d'autre part s'accompagne d'une perception opaque et donc préjudiciable de la part de l'ensemble des acteurs engagés.

Aussi convient-il, afin de choisir de façon pertinente la variable climatique utilisée dans la construction de l'indice, d'étudier les liens entre cette dernière et la survenance de sinistres. Il est pour cela nécessaire d'effectuer, à l'aide de compétences météorologiques et statistiques, un travail qui permettra de déceler ou non des liens de corrélation ou de causalité, ainsi que d'évaluer leur intensité. En dépit de l'importance de cette étape dans la mise en place de la couverture indiciaire Cat, il ne sera malheureusement pas présenté ici de véritable méthodologie complète quant au choix du type de données. L'étude d'un exemple, détaillée plus loin dans ce même chapitre, permettra néanmoins d'introduire une approche possible dans le cadre de cette étape.

Il paraît ainsi intéressant d'étudier certains liens, qui intuitivement chez beaucoup d'entre nous semblent aller de soi. Il est effectivement aisé d'envisager des liens entre force du vent et sinistralité tempête, ou entre occurrence de

températures fortement négatives et sinistres dégâts des eaux dus au gel. La carte suivante illustre la sinistralité enregistrée par un panel de cédantes du marché français lors de la survenance de la tempête Lothar⁵⁰, parallèlement aux vitesses de vent maximales instantanées enregistrées par les stations de Météo France lors des 25 et 26 décembre 1999⁵¹.



Le tableau suivant dresse une liste non exhaustive des relations entre branches d'assurances exposées aux risques naturels et variables climatiques potentiellement explicatives.

⁵⁰ La sinistralité totale enregistrée par le panel représente près de 40% du sinistre Lothar marché.

⁵¹ Interpolation par krigeage du maximum entre les valeurs des 25 et 26 décembre 1999, sur une centaine de stations françaises.

Sinistres possibles	Péril Naturel	Variable climatique parente	Données météorologiques mesurées
Dommages aux biens (DAB), Dommage Auto	Tempête, Ouragan, Cyclone	Force du vent	Vitesse de vent maximale instantanée quotidienne, Vitesse de vent maximale moyenne quotidienne, Vitesse de vent horaire, Nombre de jours du mois avec vent $\geq 28\text{m/s}$...
DAB, Dommage Auto	Grêle	Précipitations	Occurrence de grêle
DAB, Dommage Auto	Poids de la neige	Précipitations	Occurrence de neige, Hauteur de neige fraîche, Hauteur de neige totale
Dégâts des eaux	Gel	Température	Température minimale quotidienne, Nombre de jours avec Température $< 0^{\circ}\text{C}$...
Avalanche, Cat Nat	Avalanche	Précipitations, Température	Hauteur de neige fraîche, Température maximale, Température moyenne
Inondations, Cat Nat	Inondation de plaine	Précipitations, Ensoleillement	Hauteur des précipitations quotidiennes, Cumul mensuel des hauteurs de précipitations, Durée mensuelle d'insolation
Inondations, Cat Nat	Inondation torrentielle	Précipitations	Hauteur des précipitations en 1h, en 3h, en 6h...
Subsidence, Cat Nat	Sécheresse	Température, Précipitations, Ensoleillement	
	Feux de Forêts	Température, Précipitations, Force du vent	

Il est toutefois dangereux d'ériger trop hâtivement une liste associant risques naturels en matière d'assurance et variables climatiques explicatives tant peuvent être complexes les comportements de certains phénomènes météorologiques et multiples les facteurs influençant la sinistralité d'un portefeuille d'assurances (nature du risque, type de construction, facteur humain). La seule valeur d'un indice climatique, quels que soient le nombre de paramètres météorologiques utilisés et la complexité de la construction, ne suffit pas à expliquer dans sa totalité le comportement de la sinistralité d'un portefeuille d'assurance. Toutefois, l'approche indiciaire offre un extraordinaire éventail de possibilités.

b. Utilisation d'un panier de stations

Afin que la couverture indiciaire Cat prenne en compte les spécificités géographiques du portefeuille d'assurances à couvrir et donc fournisse une réassurance appropriée, la structuration de l'indice climatique s'appuie sur l'utilisation de données météorologiques provenant d'un **panier de stations synoptiques**⁵². Ceci doit permettre à l'indice de refléter la survenance de phénomènes naturels localisés sur une partie du territoire à couvrir, et susceptibles d'entraîner un accroissement de la sinistralité du portefeuille d'assurances.

⁵² Les stations synoptiques sont définies dans le « Dictionnaire du Climat », Larousse, comme des « stations dont les observations sont collectées au niveau du réseau météorologique mondial, et qui sont utilisées pour l'établissement des cartes et la prévision du temps ».

En effet il serait plus que réducteur, dans le cas de la France, de mettre en correspondance la sinistralité d'un portefeuille d'assurance au niveau national - par exemple la sinistralité tempête sur l'ensemble du territoire français - avec l'évolution d'une variable climatique observée dans une seule station - par exemple la vitesse de vent maximale instantanée enregistrée quotidiennement dans la station de Paris Montsouris⁵³. Poursuivons cet exemple : imaginons qu'une cédante souscrive une couverture indicelle Cat, de période de couverture l'exercice 1999, pour laquelle l'indemnité perçue est, sans entrer dans les détails, croissante en fonction des vitesses de vent enregistrées dans la seule station de Paris Montsouris. Le 26 décembre 1999, la tempête Lothar apporte des vents de 169 km/h sur Paris et la cédante perçoit une indemnité a priori substantielle, le tout fonctionnant comme si la tempête était généralisée à l'ensemble de l'hexagone. Le 27 décembre 1999, Martin fait des ravages dans les départements de l'Ouest et du Sud de la France, provoquant un coût de plus de EUR 2 700 millions pour les assureurs⁵⁴, mais la station de Montsouris enregistre des vents de 72 km/h, et la cédante ne perçoit aucune indemnisation à ce titre.

Si le risque de base, c'est à dire le risque que l'indice climatique ne reflète pas fidèlement le montant de sinistres, est inhérent à la protection indicelle, le choix des stations utilisées peut contribuer à le réduire. Un échantillon de stations météo de taille suffisante est nécessaire si l'on veut appréhender correctement le lien entre sinistralité et valeurs de la variable climatique. D'un point de vue pragmatique, un nombre excessif de stations peut cependant être préjudiciable à la compréhension du contrat, et surtout s'avérer inutilement onéreux en matière d'achat de données.

c. Choix des stations, répartition du portefeuille entre les stations

Un portefeuille d'assurance possède une structure géographique qui lui est propre. L'amélioration des techniques informatiques rend aujourd'hui possible une localisation précise des risques assurés. Le développement de base de données détaillées pour de nombreuses cédantes permet de localiser à une échelle

⁵³ La station de Paris Montsouris existe depuis 1872, ce qui fait d'elle la plus ancienne station météorologique de France, et par conséquent une source précieuse de données climatiques historiques.

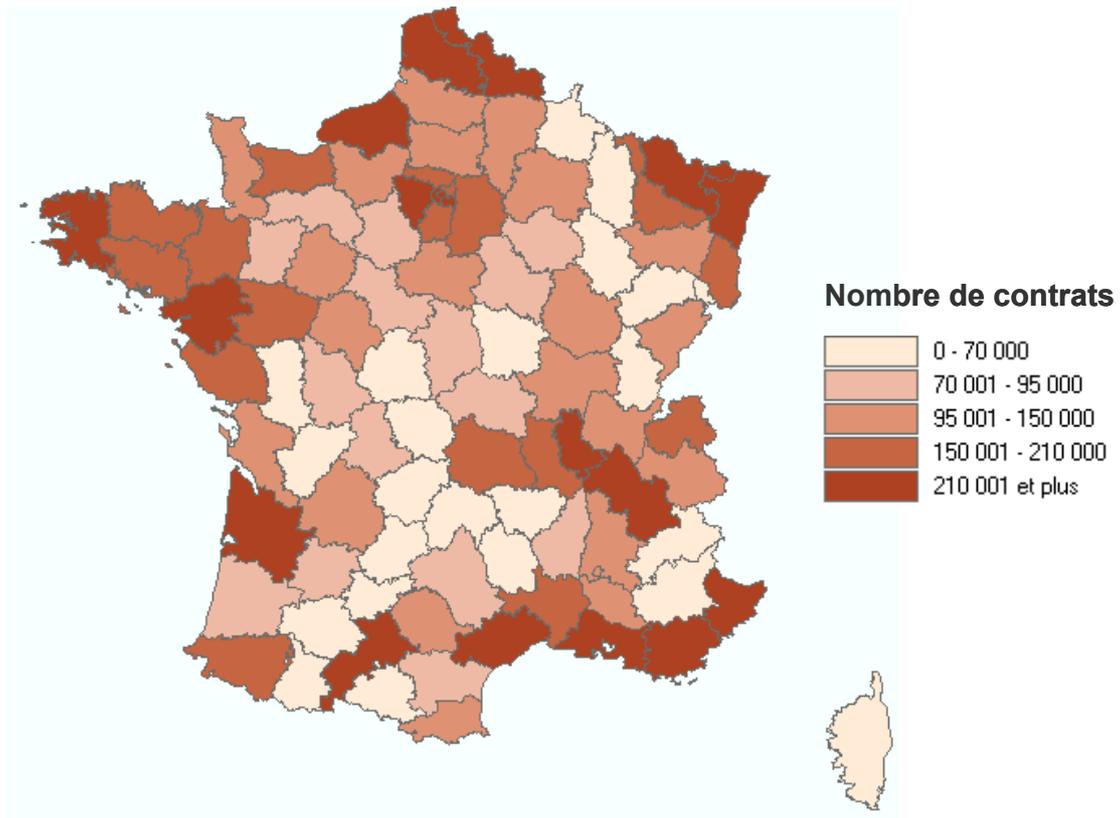
⁵⁴ « *Etude 2002 des renouvellements des protections tempête du marché français* », Benfield Paris.

géographique de plus en plus fine les paramètres d'exposition d'un portefeuille : montant des encaissements, nombre de contrats ou sommes assurées.

A titre d'exemple, le tableau suivant synthétise les données d'exposition, agrégées au département, d'un portefeuille d'une cédante fictive souscrivant sur le marché français.

Département	Cotisations (en K EUR)	Nombre de contrats	Valeurs assurées (en M EUR)
01	3 414	113 556	9 326
02	4 374	127 136	8 336
''	''	''	''
30	4 515	154 766	9 031
31	10 644	267 444	17 209
32	1 462	46 470	3 738
33	9 244	305 095	22 740
''	''	''	''
74	6 099	185 084	17 216
75	36 147	572 089	42 949
76	5 658	257 218	20 056
77	5 917	198 004	12 687
''	''	''	''
85	2 090	151 319	11 007
86	1 253	75 914	7 187
87	2 401	82 292	7 140
''	''	''	''
95	6 468	152 728	8 134
Total	399 252	13 638 469	979 584

La carte suivante représente, au niveau du département, la répartition du nombre de polices d'assurance de ce portefeuille.

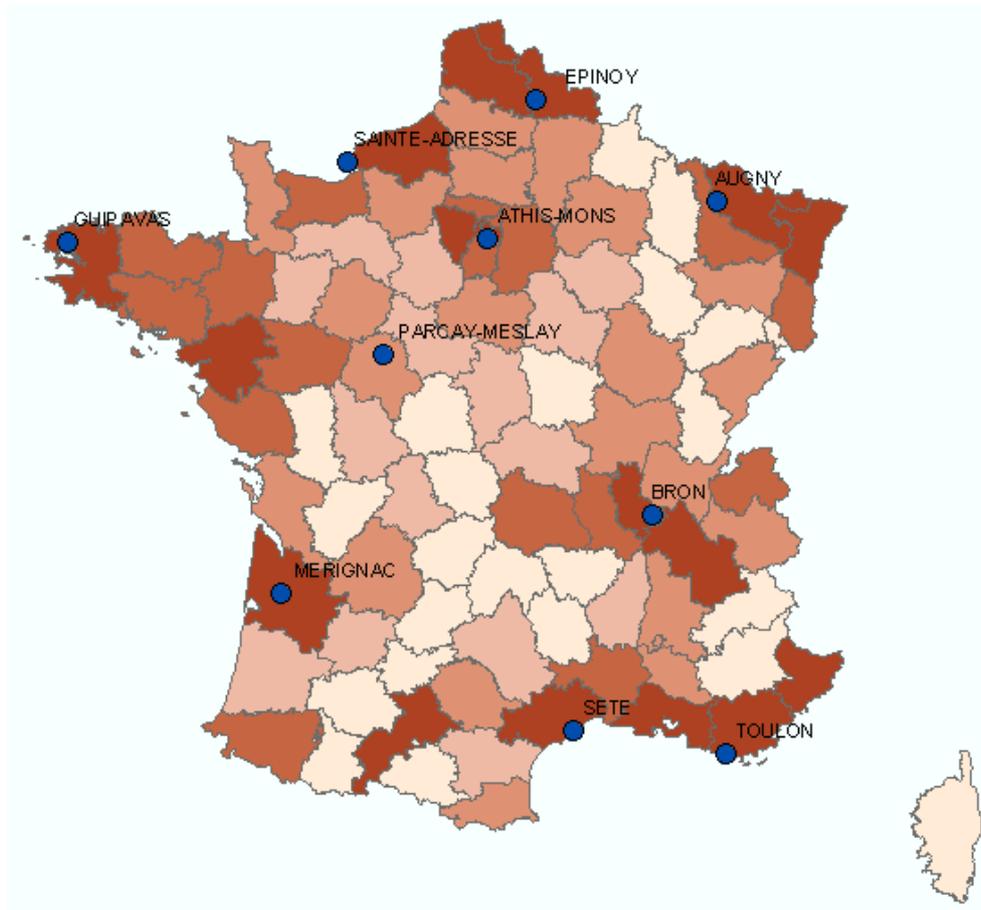


A partir de cette répartition géographique particulière des risques, il est procédé au choix d'un panier de stations approprié. Dans cet exemple illustratif, un réseau de $n = 10$ stations est retenu, dont voici la liste :

Numéro	Nom de la station	Département
1	GUIPAVAS	29
2	MERIGNAC	33
3	SETE	34
4	PARCAY-MESLAY	37
5	AUGNY	57
6	EPINOY	62
7	BRON	69
8	SAINTE-ADRESSE	76
9	TOULON	83
10	ATHIS-MONS	91

Plusieurs facteurs peuvent être pris en considération lors de la sélection des stations : la concentration de l'exposition dans certaines zones en particulier, les trajectoires des principales familles de tempêtes, la qualité du réseau de stations disponibles⁵⁵ notamment.

La localisation des stations retenues dans cet exemple est représentée sur la carte suivante :



Pour chaque station i , avec $i \in (1, \dots, n)$ et n le nombre de stations retenues, est calculé un facteur de pondération p_i qui représente l'importance de ladite station eu égard à son poids dans le portefeuille national, vis-à-vis d'une variable de référence. Ces facteurs de pondération seront par la suite utilisés dans la formule de calcul de l'indice. Cette opération vise à distinguer pour chaque station, la part de l'exposition globale du portefeuille qui lui est imputable, ceci par souci d'accorder au mieux l'éventuelle indemnisation future et la sinistralité potentielle de l'ensemble du

⁵⁵ La notion de qualité d'une station sera précisée ultérieurement.

portefeuille. La variable de référence doit donc être choisie parmi plusieurs agrégats représentatifs du risque total du portefeuille. On peut donc par exemple utiliser comme variable de référence le montant des sommes assurées du portefeuille, ou le nombre de contrats du portefeuille, ou encore le montant des primes émises du portefeuille.

Soit A la valeur de la variable de référence pour l'ensemble du portefeuille. Dans notre exemple, si la variable de référence choisie est le montant des primes encaissées, alors $A = 399\,252$ K EUR.

Plusieurs méthodes de répartition de l'exposition A entre les stations sont envisageables. L'échelle géographique retenue est un premier critère. Nous exposerons ici un exemple de méthodologie à l'échelle du département.

Soit A_j la valeur de la variable de référence pour le département j avec $j \in (1, \dots, m)$ et m le nombre de départements ($m=95$ pour la France métropolitaine). Dans notre exemple $A_{86} = 1\,253$ K EUR.

Cette méthode illustrative consiste à attribuer à chaque station une zone d'influence géographique (Z_i est la zone d'influence de la station i , avec $i \in (1, \dots, n)$). On accepte l'hypothèse que la valeur A_j du département j est concentrée dans le chef lieu de département (on pourrait également utiliser le centroïde⁵⁶ de département).

Tout d'abord, on calcule les $n.m$ distances séparant tous les chefs-lieux de département de toutes les stations retenues.

On définit X_j et Y_j comme respectivement la latitude et la longitude du chef lieu du département j , noté Dpt_j .

De même, X'_i et Y'_i sont respectivement définis comme la latitude et la longitude de la station météo retenue i .

La distance entre le chef lieu du département Dpt_j et la station i est alors :

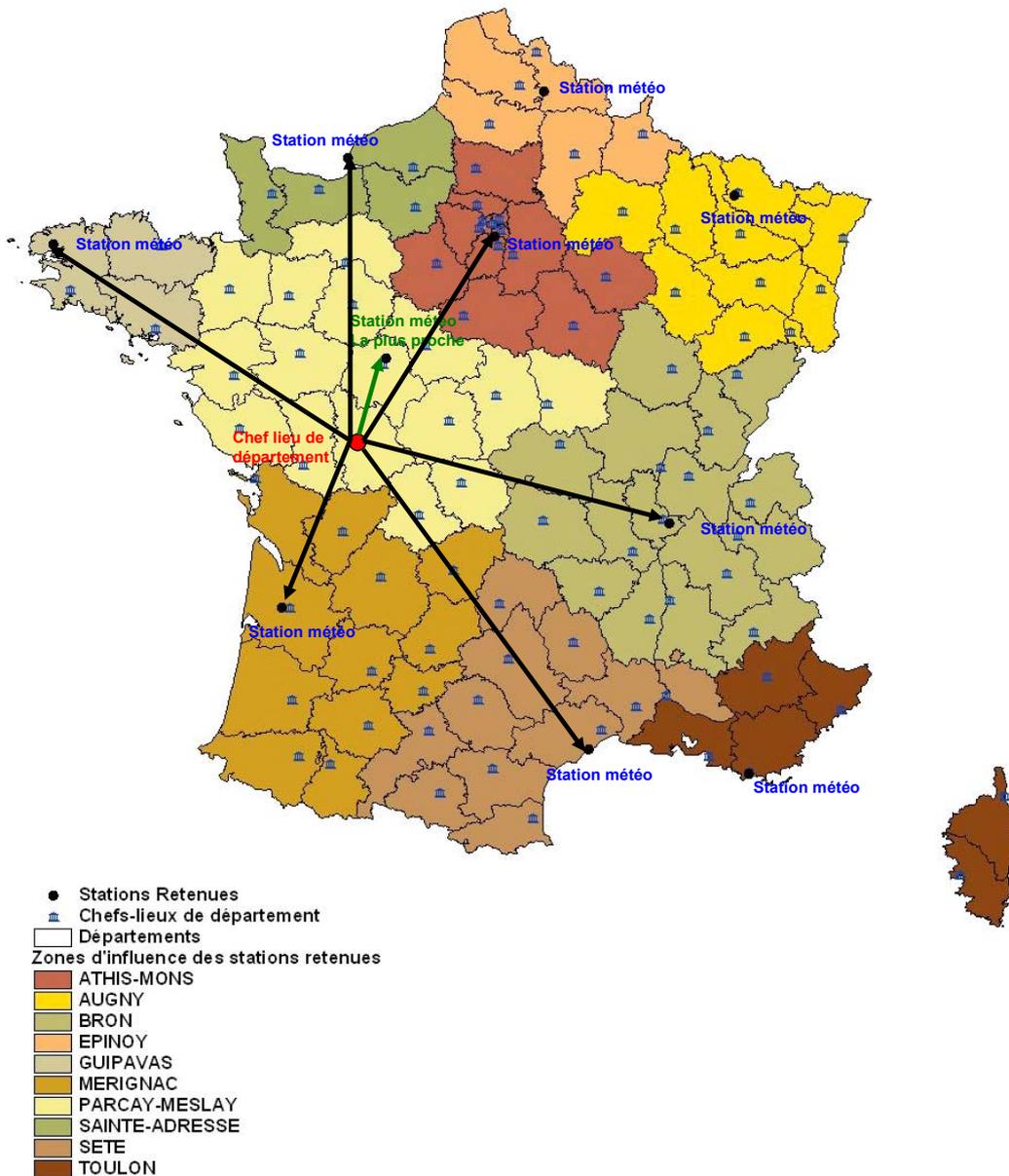
⁵⁶ On appelle centroïde d'une unité administrative un point quelconque situé approximativement au centre de la plus grande face composant cette unité.

$$d_{i,j} = \sqrt{(X'_i - X_j)^2 + (Y'_i - Y_j)^2} .$$

On considère que le chef-lieu concerné j , et a fortiori le département j tout entier, est dans la zone d'influence Z_i de la station la plus proche. C'est-à-dire :

$$(Dpt_j) \in Z_i \Leftrightarrow d_{i,j} = \min(d_{k,j}) \quad \forall k \in (1, \dots, n)$$

Chaque département est ainsi associé à la zone d'influence d'une station précise, ce qu'illustre la carte suivante :



Alors la valeur de la variable de référence pour la zone d'influence i est $A'_i = \sum_{\forall j \text{ tq } Dpt_j \in Z_i} A_j$. Dans notre exemple le montant des encaissements associés à la zone d'influence de Parçay-Meslay est $A'_4 = 32\,913$ K EUR.

Le facteur de pondération attribué à la zone d'influence i se calcule donc simplement comme $p_i = \frac{A'_i}{A} \times 100$, exprimé en pourcentage. Ici, $p_4 = 8,24\%$ est le poids de la zone d'influence de Parçay-Meslay en termes de primes encaissées.

$$\text{Les } p_i \text{ sont tels que } \frac{\sum_{i=1}^{i=n} p_i}{100} = \frac{\sum_{i=1}^{i=n} A'_i}{A} = 1.$$

La définition des poids p_i est importante dans la mesure où elle influence sensiblement la valeur de l'éventuelle indemnisation, dont le mécanisme de calcul est détaillé dans les sous parties suivantes. La méthode décrite précédemment est assez grossière, et on peut concevoir des approches bien plus fines de répartition de l'exposition du portefeuille d'assurance.

Il est par exemple possible de répartir l'exposition d'une même zone géographique entre plusieurs stations, au lieu de la rattacher totalement à une seule station. Dans un exemple où l'on souhaiterait répartir l'exposition de chaque département entre les k stations les plus proches, une façon de calculer les poids p_i pourrait être la suivante :

$$p_i = \frac{\sum_{j=1}^{j=m} \left(A_j \times \frac{d'^{-2}_{ij}}{\sum_{i=1}^{i=n} d'^{-2}_{ij}} \right)}{A} \times 100$$

avec $d'_{ij} = d_{ij}$ si d_{ij} est inférieure ou égale à la distance entre le chef lieu de Dpt_j et la $k^{\text{ième}}$ station la plus proche du chef lieu de Dpt_j ,

et $d'_{ij} = 0$ sinon.

Il est également envisageable de retraiter les variables d'exposition de chaque zone géographique, afin de mieux tenir compte des spécificités régionales en matière de vulnérabilité au péril naturel protégé. Enfin, si l'opération de répartition de l'exposition entre les stations peut être effectuée à partir de l'état présent du portefeuille d'assurance, elle peut aisément être réalisée en tenant compte d'une éventuelle évolution du portefeuille, comme par exemple une prévision de croissance dans certaines zones.

d. Définition d'un indice par station

Poursuivons la description des mécanismes de la couverture indiciaire Cat avec la présentation du calcul des indices par station. Encore une fois, il est possible de concevoir une grande variété de modes de construction des indices. Nous nous focaliserons dans ce mémoire sur l'exemple du mécanisme présenté ci-dessous, qui a généralement été utilisé pour la conception de couvertures indiciaires.

Pour chaque station i , on observe et on enregistre à chaque date t ($t \in [0, T]$, où T est la date de maturité de la couverture indiciaire Cat contractée en 0), les données climatiques sur un intervalle de temps $[t-1, t]$, de durée $t - (t-1)$ égale par exemple à une heure ou un jour.

Soit $X_i(t)$ la valeur en unités de mesure de la variable climatique, par exemple km/h ou degré Celsius, enregistrée dans la station i à la date t .

Pour chaque station i , le contrat de réassurance définit une valeur K_i en unité de mesure de la variable climatique, et peut définir une valeur L_i en unité de mesure de la variable climatique telle que $L_i > K_i$.

Est alors défini comme indice sur l'intervalle de temps $[t-1, t]$, à la date t :

$$I_i(t) = \min[(L_i - K_i), \max[(X_i(t) - K_i), 0]]$$

Ainsi par exemple : $I_i(t)$ est un indice horaire si la durée $t - (t - 1)$ égale une heure, et $I_i(t)$ est un indice journalier si la durée $t - (t - 1)$ égale un jour.

On définit l'indice de la station i à la date t comme le cumul sur $[0, t]$ des $I_i(t)$.

$$\text{Ainsi : } S_i(t) = \sum_{t'=1}^{t=t} I_i(t')$$

Le contrat de réassurance indiciel cat étant de maturité T (l'année généralement), l'indice final de la station i est alors :

$$S_i = S_i(T) = \sum_{t=1}^{t=T} I_i(t)$$

On remarque que $S_i(t)$ est croissant avec t , puisque $I_i(t) \geq 0$.

e. Mécanisme de l'indice final, indemnisation due au contrat

L'indice général de la couverture indicielle Cat est à la date t :

$$S(t) = \sum_{i=1}^{i=n} p_i \times S_i(t)$$

L'indice à maturité est :

$$S(T) = \sum_{i=1}^{i=n} p_i \times S_i(T)$$

Le principe d'indemnisation dans une couverture indicielle Cat est semblable à celui d'une option européenne classique ou à celui d'un traité de réassurance non proportionnel.

Ainsi, l'assureur ayant acheté une protection indicielle cat, recevra à la fin de la période de couverture (à la date T), une indemnité $C(T)$, si la valeur de l'indice

$S(T)$ est supérieure à une franchise K (en unités de la variable climatique) déterminée au contrat, telle que :

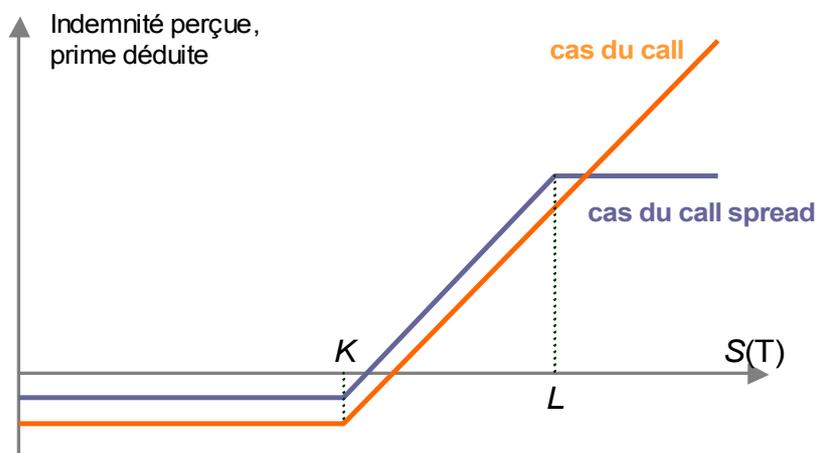
$$C(T) = N \times \max[(S(T) - K), 0]$$

où N est le montant nominal en unités monétaires correspondant à une unité d'indice supplémentaire au dessus de la franchise K , que l'on désigne également par « tick ».

L'indemnité perçue au moyen de la couverture indicielle Cat, correspond donc au payoff d'un call de maturité T sur le sous jacent S , avec un prix d'exercice K , et avec un facteur de conversion N .

Les preneurs de risques préfèrent généralement, face à ce type de couverture, et tout comme pour les traités de réassurance traditionnelle, limiter leur engagement. Ainsi les caractéristiques d'une couverture indicielle comportent bien souvent la définition d'une valeur L (en unité climatique) plafonnant la valeur de l'indice donnant droit à indemnisation, si bien que l'indemnité perçue au moyen de la couverture indicielle Cat s'apparente au payoff d'un callspread (ou capped call) ou encore à l'indemnisation d'une tranche d'un traité de réassurance en excédent de sinistre de priorité K et de portée $L - K$.

$$C(T) = N \times \max[\min[(L - K), (S(T) - K)], 0]$$



La construction de l'indice sous-jacent à la protection indicielle Cat peut revêtir une multitude de formes différentes. Les mécanismes présentés précédemment peuvent être rendus plus complexes, par exemple par extension à plusieurs pays, par utilisation de plusieurs variables climatiques, ou encore par définition des ticks au niveau de chaque station.

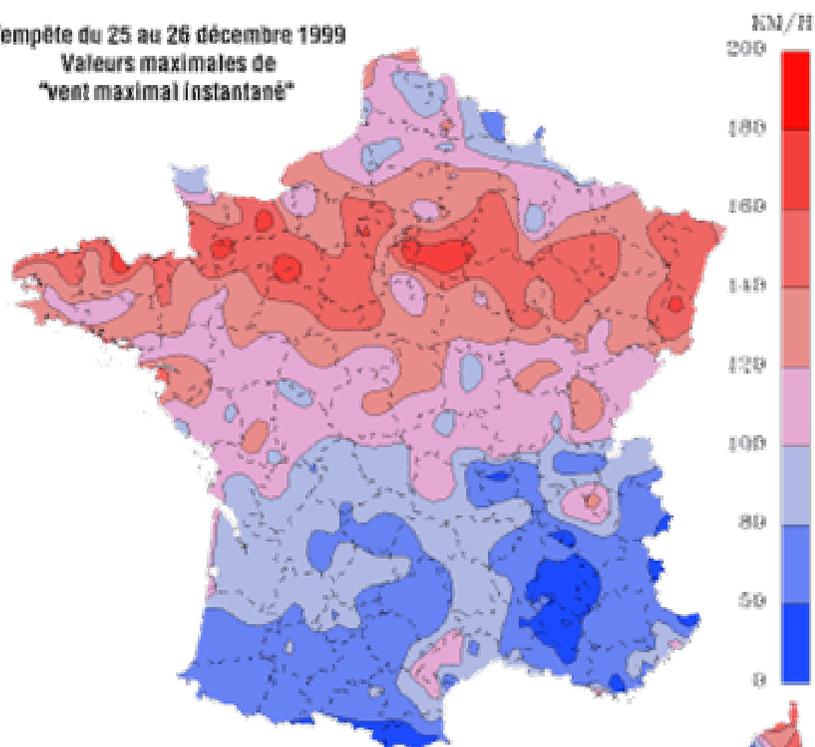
3.3 Etude d'un exemple

Ainsi, une couverture indicielle Cat ayant pour but de réassurer une cédante contre le risque tempête, peut être construite en utilisant un indice basé sur les vitesses de vent maximales instantanées enregistrées quotidiennement. C'est sur ce lien existant entre occurrence de vents violents et sinistralité d'un portefeuille d'assurance due à la tempête, qu'a été introduite la première couverture indicielle Cat évoquée précédemment. Les cartes suivantes visent à illustrer ce lien : pour les deux tempêtes de décembre 1999, on peut comparer visuellement les forces de vent maximales instantanées le jour de la tempête à la sinistralité déclarée à la suite de la tempête par un panel de cédantes du marché français. La simple observation de ces cartes semble indiquer une forte relation entre force du vent et sinistralité due à la tempête.

Vitesses de vent enregistrées

Lothar

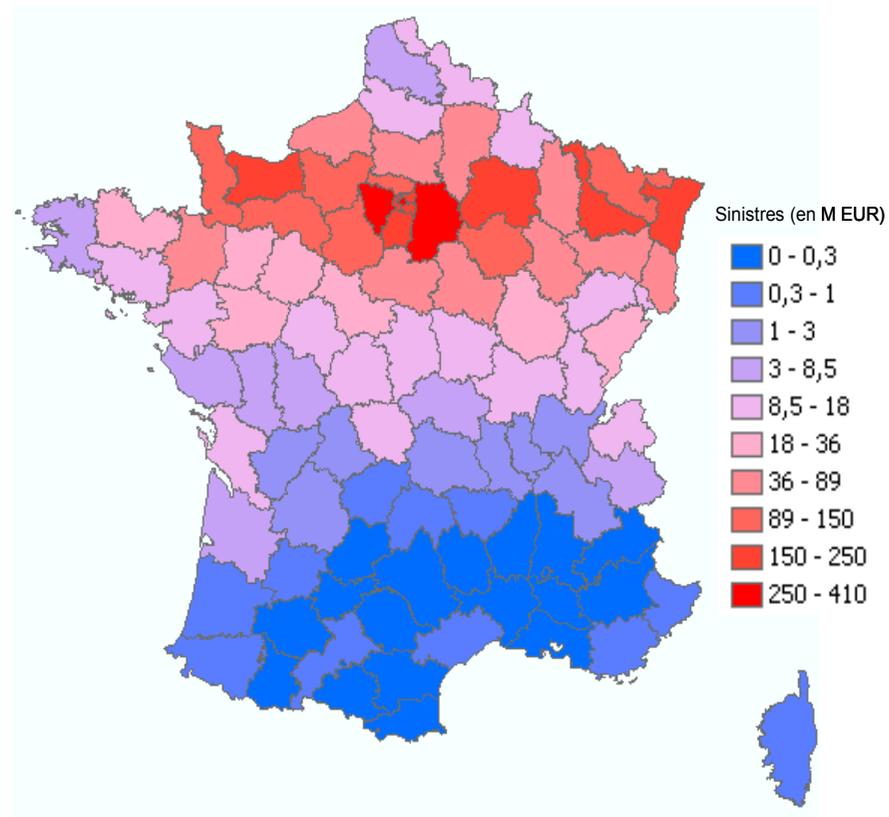
Tempête du 25 au 26 décembre 1999
Valeurs maximales de
"vent maximal instantané"



Stations dont l'altitude est inférieure ou égale à 500 mètres

Sinistralité par département d'un panel de cédantes

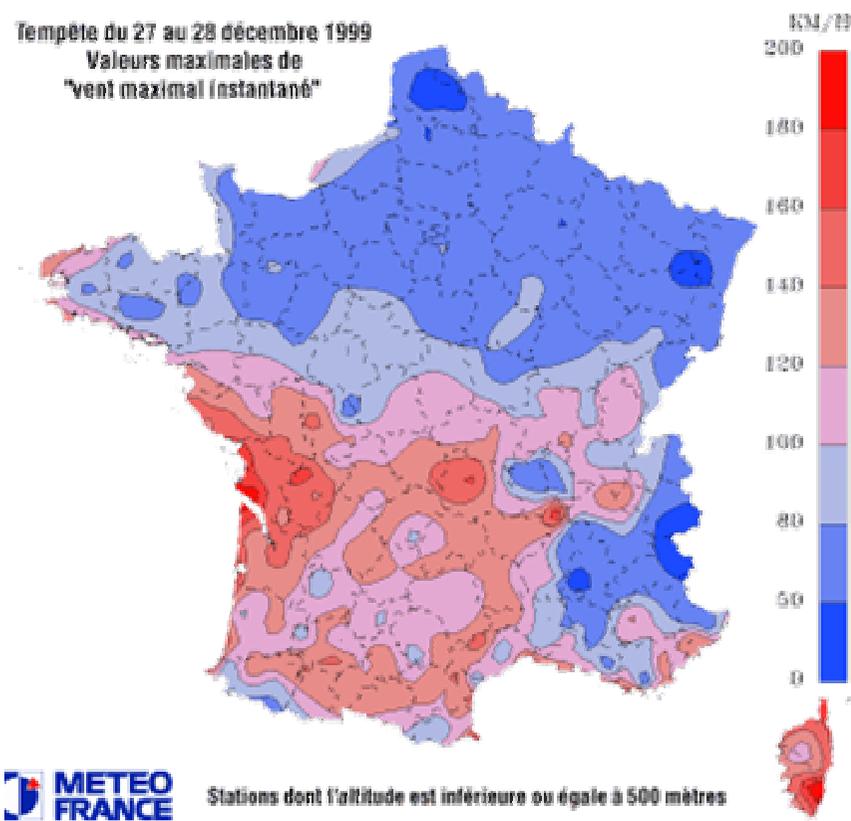
Lothar



Vitesses de vent enregistrées

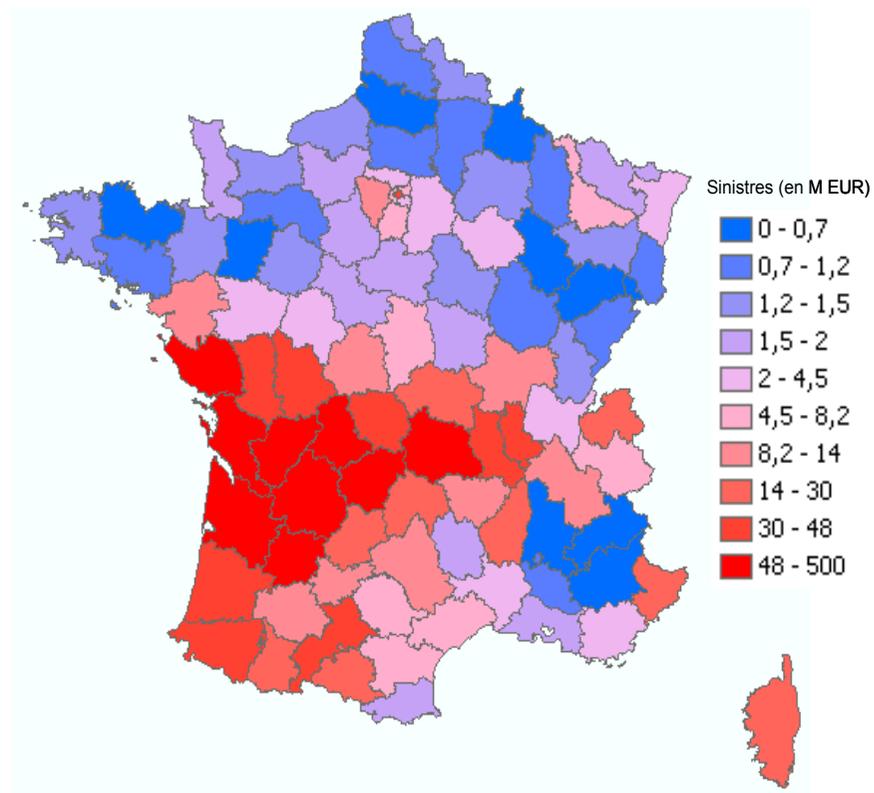
Martin

Tempête du 27 au 28 décembre 1999
Valeurs maximales de
"vent maximal instantané"



Sinistralité par département d'un panel de cédantes

Martin



Afin d'étudier plus précisément la relation entre la force du vent et la sinistralité tempête, il est par exemple possible de mettre en correspondance, au niveau de chaque département, une mesure de la sinistralité enregistrée lors de ces deux événements et les vitesses de vent interpolées aux centroïdes de département à partir des observations d'un échantillon de stations météorologiques.

Soit Y_{ij} , avec $i \in (1,2)$ et $j \in (1, \dots, 95)$, la mesure de la sinistralité enregistrée par le département j lors de la tempête i .

Dans notre exemple, $i = 1$ correspond à la tempête Lothar, et $i = 2$ correspond à la tempête Martin.

Nous utiliserons la mesure de sinistralité suivante :

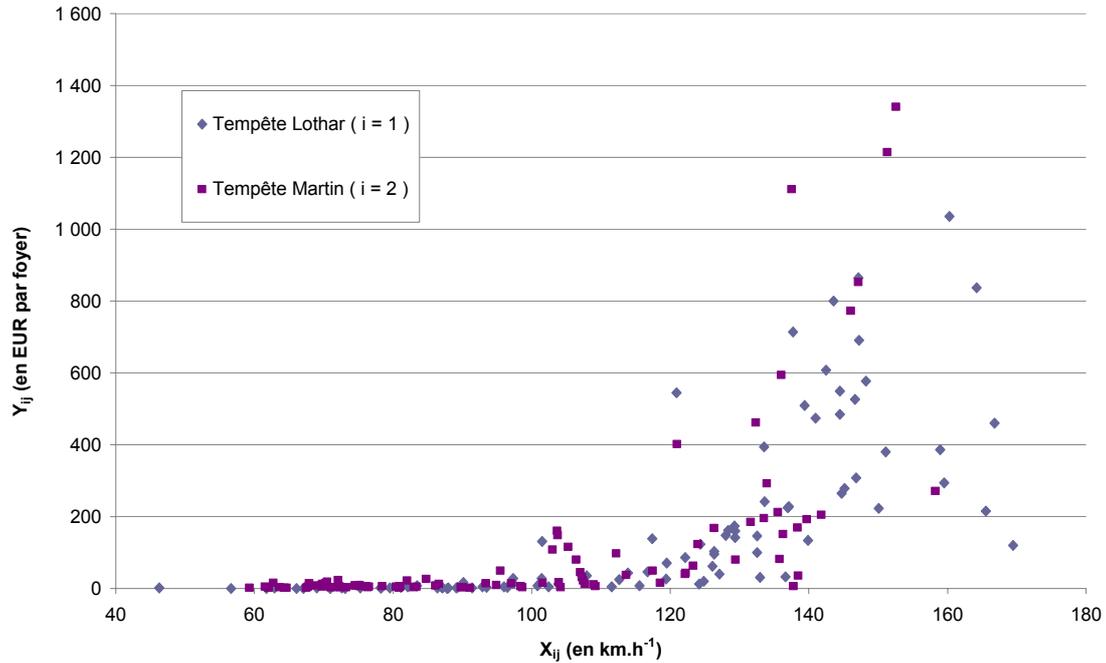
$$Y_{ij} = \frac{\text{Montant des sinistres enregistrés dans le département } j \text{ suite à la tempête } i}{\text{Nombre de foyers du département } j \text{ (recensement Insee 1999)}}$$

Soit X_{ij} , avec $i \in (1,2)$ et $j \in (1, \dots, 95)$, la vitesse de vent interpolée du département j lors de la tempête i .

Pour chaque département, la valeur des X_{ij} est calculée par interpolation des valeurs maximales de vent maximal instantané enregistrées dans les cinq stations les plus proches⁵⁷ du centroïde de département j lors de la survenance de la tempête i (25 et 26 décembre 1999 pour Lothar, 27 et 28 décembre 1999 pour Martin).

Le graphique suivant représente pour chaque couple (i,j) la valeur X_{ij} en abscisse, et la valeur Y_{ij} en ordonnée.

⁵⁷ Le calcul est effectué à partir d'un échantillon de 133 stations.



On constate que l'occurrence d'une sinistralité significative due aux tempêtes n'intervient qu'à partir d'un seuil de vitesse de vent compris entre 90 et 100 km/h.

En effectuant une régression linéaire des valeurs $\log(Y_{ij})$ par les X_{ij} , on obtient l'ajustement suivant :

$$\log(\hat{Y}_{ij}) = 0,0636 \times X_{ij} - 3,5977$$

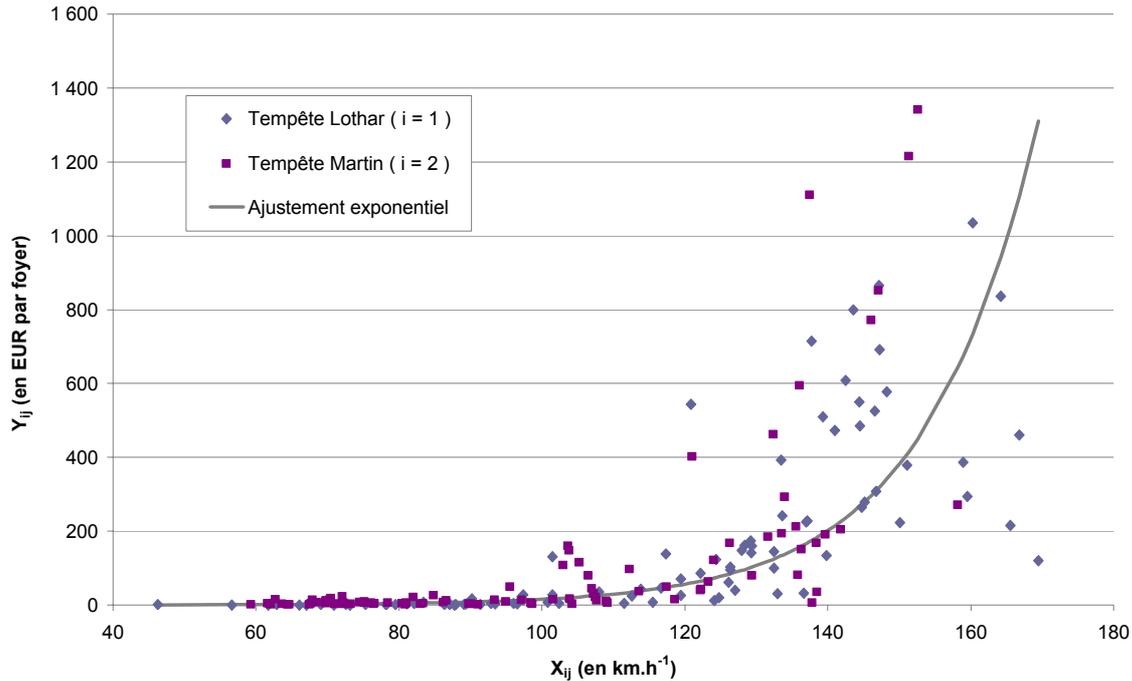
c'est à dire :

$$\hat{Y}_{ij} = 0,0274 \times \exp(0,0636 \times X_{ij})$$

Le coefficient R^2 de l'ajustement est égal à :

$$R^2 = \frac{\sum(Y_{ij} - \bar{Y})^2 - \sum Y_{ij} (Y_{ij} - \hat{Y}_{ij})^2}{\sum(Y_{ij} - \bar{Y})^2} = 0,747$$

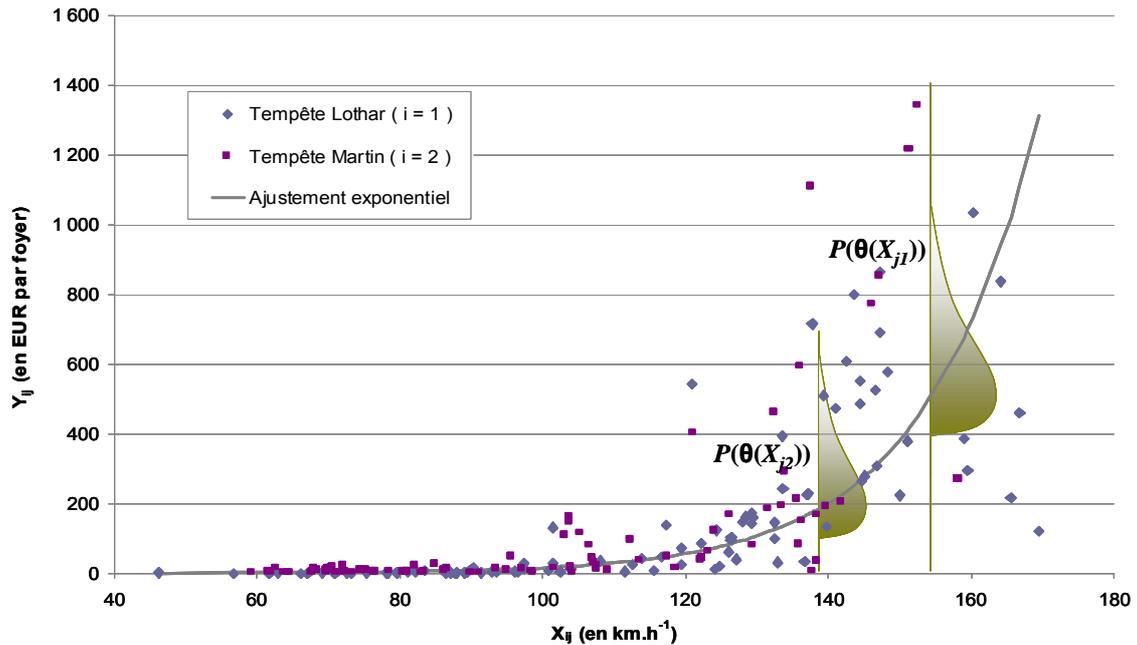
Voici la représentation graphique de l'ajustement :



L'ajustement est assez satisfaisant, compte tenu du R^2 et de la représentation graphique obtenus. Il est donc envisageable de représenter une mesure de sinistralité par département liée à une tempête (soit Y_j cette variable aléatoire) en fonction de la vitesse de vent interpolée pour le département (soit X_j cette variable aléatoire), c'est-à-dire définir une fonction f telle que : $Y_j = f(X_j)$. L'ajustement réalisé à partir des tempêtes historiques Lothar et Martin peut par exemple être utilisé comme estimation de f .

Toutefois, la sinistralité est très volatile pour les vitesses de vents extrêmes, et l'utilisation de cette fonction f ne prendrait pas suffisamment en compte l'incertitude qui porte sur le niveau de sinistralité une fois donnée la mesure de vitesse de vent. Cette incertitude est couramment désignée comme l'incertitude secondaire.

Afin de mieux appréhender les liens entre vitesses de vent et sinistralité, une alternative consiste ainsi à faire l'hypothèse que Y_j suit une loi de probabilité P de paramètre $\theta(X_j)$ (où $\theta(X_j)$ est un vecteur de paramètres qui dépendent de X_j), puis à chercher le meilleur ajustement possible. Le graphique suivant illustre cette approche :



L'étude de la relation entre une variable climatique et la sinistralité potentielle d'un portefeuille d'assurance mérite un travail d'analyse statistique approfondi, qui ne fait pas toutefois l'objet de ce mémoire.

La couverture indicielle Cat, comme indiqué précédemment, repose sur le lien existant entre les variations exceptionnelles de variables climatiques et la sinistralité d'un portefeuille d'assurance. L'indice est donc généralement conçu de telle façon que sa valeur croisse avec le caractère exceptionnel des observations de la donnée météorologique. Pour une couverture indicielle contre le risque de tempête, seules les vitesses de vent supérieures à un seuil entraîneront une hausse de l'indice. Dans la mesure où elles ne s'accompagnent pas de la survenance de sinistres, les vitesses de vent nulles, faibles ou moyennes, n'affectent pas la valeur de l'indice, tel qu'il est généralement construit.

Nous poursuivrons désormais avec l'étude d'un exemple précis d'une couverture indicielle vent, dont les caractéristiques sont détaillées ci-après.

Nous prendrons donc l'exemple d'une protection basée sur un indice construit à partir des vitesses de vent maximales instantanées enregistrées quotidiennement dans un panel volontairement limité de $n = 4$ stations, dont voici la liste :

Numéro i	Nom de la station	Code WMO
1	BORDEAUX-MERIGNAC	7510
2	METZ-FRESCATY	7090
3	NICE	7690
4	ORLY	7149

Dans cet exemple, nous utiliserons pour les quatre stations les facteurs de pondérations p_i suivants :

Numéro i	Nom de la station	Poids p_i
1	BORDEAUX-MERIGNAC	25 %
2	METZ-FRESCATY	18 %
3	NICE	20 %
4	ORLY	37 %

Par ailleurs, quatre seuils K_i ont été choisis pour le calcul des indices par station :

Numéro i	Nom de la station	Franchise K_i (en m/s)	Franchise K_i (en dam/h) ⁵⁸
1	BORDEAUX-MERIGNAC	29,5	10620
2	METZ-FRESCATY	30,0	10800
3	NICE	29,5	10620
4	ORLY	31,5	11340

La couverture porte sur la valeur de l'indice $S(T)$, en dam/h, construit comme la somme des produits des facteurs de pondération p_i et des indices $S_i(T)$ de chacune des quatre stations :

$$S(T) = \sum_{i=1}^{i=4} p_i \times S_i(T)$$

⁵⁸ Nous utiliserons souvent le décimètre par heure (dam/h) comme unité de mesure des vitesses de vent.

L'indice $S_i(T)$, en dam/h, de la station i est défini par :

$$S_i(T) = \sum_{t=t_0}^{t=T} I_i(t) = \sum_{t=t_0}^{t=T} \max((X_i(t) - K_i), 0)$$

avec $X_i(t)$ la valeur en dam/h de la vitesse de vent maximale instantanée enregistrée à la date t dans la station i ,

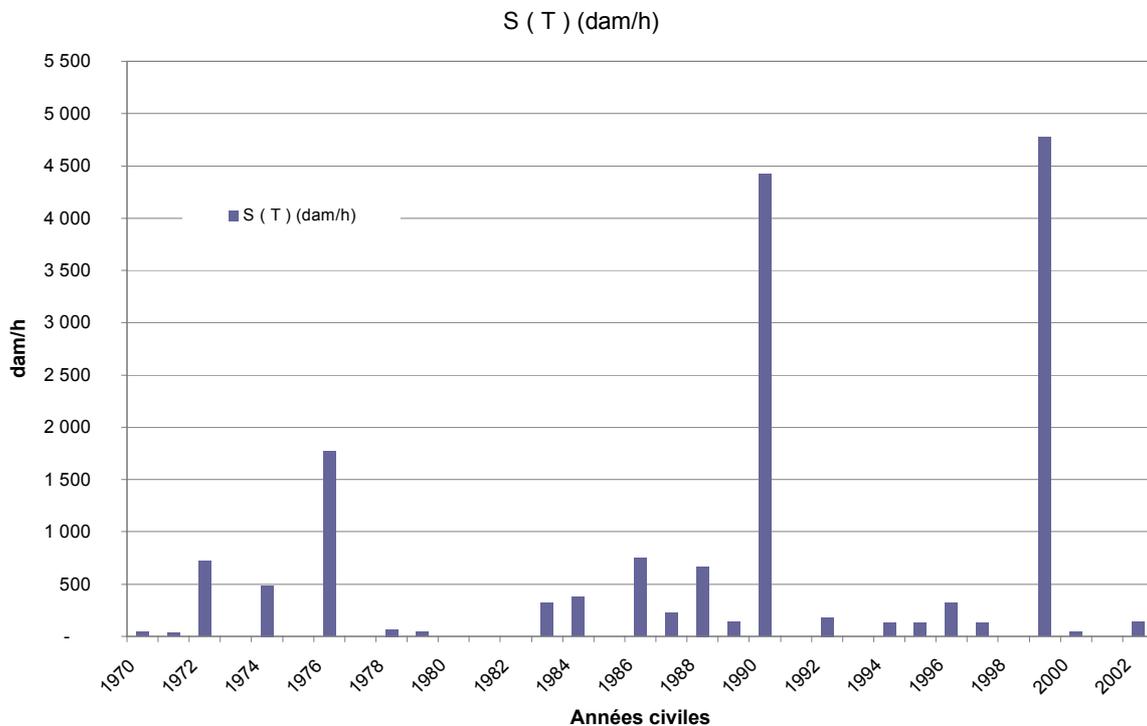
et avec t_0 le 1^{er} janvier de l'année N , et T le 31 décembre de l'année N . Le choix de l'année civile comme période de couverture n'est pas meilleur en soi, mais s'adapte mieux aux habitudes du marché de la réassurance.

Cet exemple très simplifié de couverture indicielle poursuit avant tout une fin illustrative. Il permettra de faciliter les travaux d'analyse qui sont exposés dans la suite du mémoire.

En utilisant les enregistrements historiques des vitesses de vent pour les quatre stations depuis le 01/01/1970, il est possible de calculer la valeur des indices S(T) de 1970 à 2002, ce qu'illustre le tableau suivant :

	BORDEAUX	METZ	NICE	ORLY	Total
Année civile	S ₁ (T) (dam/h)	S ₂ (T) (dam/h)	S ₃ (T) (dam/h)	S ₄ (T) (dam/h)	S(T) (dam/h)
1970	180	-	-	-	45,00
1971	-	-	180	-	36,00
1972	2 880	-	-	-	720,00
1973	-	-	-	-	-
1974	900	360	-	540	489,60
1975	-	-	-	-	-
1976	6 300	-	-	540	1 774,80
1977	-	-	-	-	-
1978	-	-	-	180	66,60
1979	180	-	-	-	45,00
1980	-	-	-	-	-
1981	-	-	-	-	-
1982	-	-	-	-	-
1983	540	-	900	-	315,00
1984	1 260	360	-	-	379,80
1985	-	-	-	-	-
1986	2 880	-	180	-	756,00
1987	900	-	-	-	225,00
1988	-	-	-	1 800	666,00
1989	540	-	-	-	135,00
1990	900	14 040	-	4 500	4 417,20
1991	-	-	-	-	-
1992	-	-	900	-	180,00
1993	-	-	-	-	-
1994	-	720	-	-	129,60
1995	-	720	-	-	129,60
1996	1 260	-	-	-	315,00
1997	-	720	-	-	129,60
1998	-	-	-	-	-
1999	3 780	2 520	900	8 640	4 775,40
2000	180	-	-	-	45,00
2001	-	-	-	-	-
2002	540	-	-	-	135,00

La représentation graphique suivante des valeurs historiques de $S(T)$ permet d'observer la variabilité historique de l'indice :



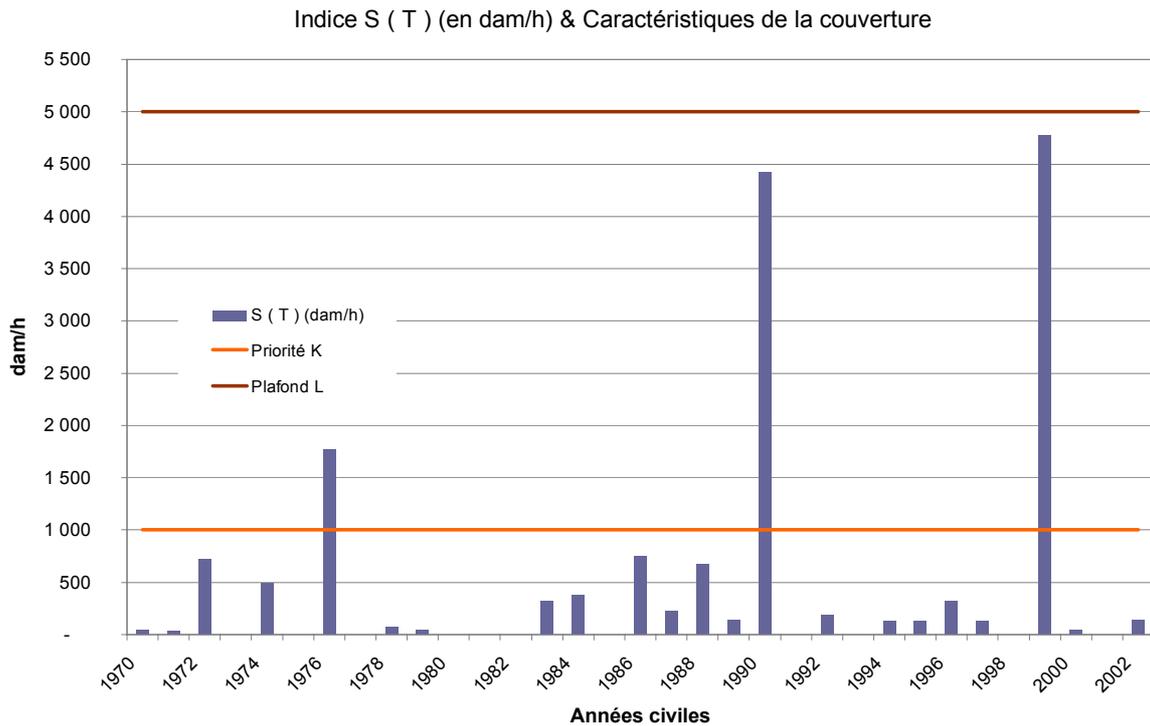
Enfin, la couverture sur l'indice $S(T)$ est caractérisée par la définition de deux bornes, K et L , semblables à une priorité et un plafond, et par la définition du tick N , qui entrent dans le calcul du montant de l'indemnisation $C(T)$:

$$C(T) = N \times \max[\min[(L - K), (S(T) - K)], 0]$$

Nous utiliserons les valeurs suivantes :

Priorité K	Plafond L	Tick N
1 000 dam/h	5 000 dam/h	EUR 1 000 par dam/h

Avec ces caractéristiques, la couverture aurait historiquement conduit au paiement d'une indemnisation à la cédante au cours de trois années, 1976, 1990, et 1999, comme on peut l'observer sur le graphique suivant.



Les montants des indemnisations pour ces trois années auraient alors été les suivants :

Année	Indice Annuel S (T) (en dam/h)	Franchise K	Indice à la charge de la couverture (en dam/h)	Montant de l'indemnisation C (T) (en EUR)
1976	1 774,80	1 000,00	774,80	774 800
1990	4 417,20	1 000,00	3 417,20	3 417 200
1999	4 775,40	1 000,00	3 775,40	3 775 400

Après une synthèse des principaux enjeux liés aux couvertures indicielles en général, les chapitres suivants aborderont les approches possibles afin de déterminer le montant de la prime correspondant à cette protection.

3.4 Limites et débat sur la couverture indiciaire Cat

Le mécanisme de la couverture indiciaire permet de concevoir un large éventail de protections d'un portefeuille d'assurance contre les risques naturels. L'exemple de couverture qui va être approfondi dans le cadre de ce mémoire est un cas très particulier de l'ensemble des applications que l'on peut imaginer pour les couvertures indiciaires. Le choix des paramètres météorologiques, la structuration de l'indice et la définition du mécanisme d'indemnisation sont autant de leviers qui permettent d'adapter la protection indiciaire à divers objectifs de couverture : nature du péril couvert, horizon de couverture (sinistralité survenant pour une certaine période de retour), couverture aggregate annuelle.

L'obstacle majeur au développement des couvertures indiciaires comme forme alternative de réassurance des périls naturels tient à l'existence structurelle du risque de base. Bien que pouvant être réduit, notamment par le choix judicieux des paramètres climatiques et des stations dans la construction de l'indice, ce risque d'écart entre le montant de la sinistralité réelle et le montant de l'indemnisation due au titre de la couverture indiciaire est inéluctable.

Par ailleurs, le développement des couvertures indiciaires nécessite un développement conjoint des méthodes de tarification de ces nouvelles couvertures. Le présent mémoire vise à participer à cette recherche de méthodes de tarification, et les prochaines parties détailleront certaines approches envisageables, dans le cadre de l'exemple décrit précédemment.

4 PREMIERES APPROCHES DE TARIFICATION DE LA COUVERTURE

Nous allons présenter dans cette partie principalement deux approches envisageables en vue d'estimer le prix d'une couverture indicieille. En prenant comme exemple la couverture indicieille vent décrite dans la partie précédente, nous estimerons tout d'abord le prix de la couverture en utilisant l'historique des valeurs de l'indice $S(T)$. Puis nous effectuerons une modélisation des variables climatiques sous jacentes elles-mêmes, à savoir les vitesses de vent maximales instantanées enregistrées quotidiennement dans les stations météorologiques retenues.

Les approches de tarification abordées dans ce mémoire sont basées sur l'hypothèse de la représentativité des valeurs climatiques historiques quant au comportement futur de ces variables climatiques. Dans notre exemple de couverture indicieille vent, nous n'utiliserons pas de modèle de prévision climatique afin d'estimer les probabilités des valeurs futures des observations de la variable climatique. Si cette hypothèse est très nettement discutable dans le cadre d'indices de température, il n'existe pas aujourd'hui de modèles scientifiques de prévision des tempêtes en Europe.

4.1 Hypothèses sur la qualité des séries de données

Les historiques de données climatiques des offices météorologiques nationaux sont généralement riches et longs. La plupart des stations météorologiques permettent de disposer de séries de données longues de plusieurs décennies. Dans la manipulation de données propres à un portefeuille d'assurance (historiques des cotisations, des sinistres), la longueur relativement faible de l'historique exploitable, généralement de cinq à dix ans, constitue souvent un inconvénient lors de la tarification⁵⁹.

L'utilisation de longues séries de données climatiques offre la possibilité d'effectuer une tarification sur des bases intéressantes. Toutefois, il existe certaines

⁵⁹ L'évolution constante d'un portefeuille d'assurance (croissance interne, croissance externe, changement du comportement de la sinistralité, etc.) rend souvent difficile l'interprétation des statistiques du passé.

difficultés liées à la qualité des séries de données climatiques, qui sont susceptibles d'altérer la crédibilité d'une approche de tarification.

4.1.1 Qualité de mesure au cours du temps

De nombreux facteurs peuvent affecter la qualité de mesure des paramètres météorologiques au cours du temps dans une station, parmi lesquels :

- l'évolution de l'environnement direct de la station, comme le développement urbain ou la modification de la végétation,
- le changement (le perfectionnement souvent) des instruments de mesure.

Ainsi, certaines évolutions exogènes au comportement de la variable climatique mesurée peuvent entraîner des changements singuliers, des ruptures dans les séries de données climatiques. A l'aide de techniques statistiques, les ruptures présentes dans les séries de données climatiques peuvent être détectées et éventuellement corrigées en vue de rendre les séries homogènes. Nous utiliserons dans le cadre de notre exemple les séries originelles de vitesses de vent des quatre stations retenues. Le choix de faire débiter l'historique exploité au 01/01/1970 nous permet d'esquiver l'éventuelle présence de données non homogènes.

4.1.2 Le problème des données manquantes

Les instruments d'observation météorologique n'étant pas infaillibles, le problème de l'absence de mesure de la variable climatique pour certaines dates peut se poser dans le calcul de l'indice sous-jacent à une couverture. Le problème se pose également dans l'étude des valeurs historiques en vue d'une tarification. Pour y remédier, la première approche tend à privilégier les séries de données les plus complètes lors de la sélection des stations météorologiques qui sont retenues pour la construction de l'indice. Enfin, en vue de reconstituer des séries de données complètes et d'assurer le calcul d'un indice en vigueur, chaque station sélectionnée se voit associer une ou plusieurs stations de repli. Celles-ci sont prioritairement choisies en fonction de leur proximité géographique des stations originales. Elles font

également l'objet d'études de corrélation avec les stations originales, afin de mesurer la pertinence de leur utilisation.

4.1.3 Conclusion

La qualité des séries de données climatiques constitue un enjeu de premier plan quant à l'évaluation du prix d'une couverture indiciaire. Nous utiliserons dans le cadre de ce mémoire les séries de données de vitesses de vent maximales instantanées constituées par la société CatiXL⁶⁰ pour les quatre stations retenues dans l'exemple. Les séries dont nous disposons comportent ainsi un historique complet de valeurs quotidiennes de vitesses de vent maximales instantanées jusqu'à la date du 31/12/2002. Les fiches descriptives des séries utilisées et réalisées par CatiXL se situent en annexe 2.

Nous choisissons, comme indiqué précédemment, de faire débiter les séries de vitesse de vent utilisées dans ce mémoire au 01/01/1970. Compte tenu de la modification des appareils de mesure de la force du vent au début des années 1970 dans les stations de Météo France⁶¹, il apparaît préférable de tronquer les séries afin de disposer de données homogènes.

⁶⁰ La société CatiXL recueille et vérifie des données météorologiques provenant d'un réseau international de stations météorologiques. Elle peut ainsi constituer des séries de données climatiques complètes pour un certain nombre de stations, en choisissant attentivement les stations de repli. CatiXL met ensuite à la disposition de ses clients des outils de structuration de couvertures sur base d'indices climatiques.

⁶¹ Il s'agissait d'un changement de capteur. Les stations disposent toutes des mêmes instruments de mesure, il n'y a donc pas de problème d'homogénéité entre les stations.

4.2 Tarifications sur les valeurs d'indice

En vue d'estimer le prix de la couverture, la démarche qui semble la plus évidente, en raison de sa simplicité, consiste à utiliser l'historique des valeurs d'indice annuel. Nous présentons dans cette sous-partie les différentes voies étudiées qui s'inscrivent dans cette démarche.

4.2.1 Burning Cost sur l'indice S(T)

Le Burning Cost correspond à la prime annuelle qu'un réassureur ayant souscrit la couverture indiciaire sur les n dernières années aurait dû demander à la cédante afin que les primes de réassurance perçues équilibrent le montant des indemnités qu'il aurait été amené à verser.

Si on note C_j et S_j respectivement la valeur de l'indemnité $C(T)$ et la valeur de l'indice $S(T)$ pour l'année j , le Burning Cost sur n années de la couverture indiciaire étudiée s'établit comme suit :

$$\bar{C}_n = \frac{\sum_{j=2002-n+1}^{j=2002} C_j}{n} \times N = \frac{\sum_{j=2002-n+1}^{j=2002} \max[\min[(L-K), (S_j-K)], 0]}{n} \times N$$

Il correspond donc à la moyenne des indemnités sur les n dernières années, et donne ainsi un estimateur de la prime pure de la couverture. Lorsque le Burning Cost est non nul, il constitue bien souvent dans la pratique un élément de référence pour la cotation d'une protection de réassurance.

Dans notre exemple, nous obtenons les résultats synthétisés dans le tableau ci-après :

Burning Cost \bar{C}_n (en EUR)	en dam/h	en EUR
sur 5 ans (période 1998-2002)	755,08	755 080
sur 10 ans (période 1993-2002)	377,54	377 540
sur 20 ans (période 1983-2002)	359,63	359 630
sur 33 ans (période 1970-2002)	241,44	241 440

Pour rémunérer sa prise de risque, le réassureur demande généralement à la cédante une prime commerciale supérieure à la prime pure, en incorporant un chargement de sécurité. Le chargement est le plus souvent le produit d'un coefficient que nous noterons a , par l'écart type empirique du montant annuel des indemnisations⁶². Nous notons s_n l'écart type des indemnisations annuelles sur les n dernières années. La prime commerciale s'écrit alors :

$$\Pi = \bar{C}_n + a \times s_n$$

En utilisant deux valeurs différentes de a , 20 % et 40 %, nous obtenons les estimations des primes commerciales suivantes :

Période	\bar{C}_n (en EUR)	s_n	Π avec $a = 20\%$ (en EUR)	Π avec $a = 40\%$ (en EUR)
sur 5 ans	755 080	1 688 410	1 092 762	1 430 444
sur 10 ans	377 540	1 193 886	616 317	855 095
sur 20 ans	359 630	1 108 442	581 318	803 007
sur 33 ans	241 436	876 919	416 820	592 204

⁶² De nombreux mémoires ont étudié ce point. Nous n'approfondirons pas ici les différentes méthodes de chargement envisageables.

Si la méthode, qui peut exploiter un historique long de 33 années dans notre exemple, a l'avantage de la simplicité et de la rapidité, son application présente toutefois des limites. En fonction de la longueur de l'historique retenu, les montants estimés des primes commerciales de la couverture varient du simple au triple.

Dans près de 91 % des années (30 sur 33), la couverture n'aurait pas conduit à une indemnisation de la cédante. Cette volatilité rend difficile l'interprétation du Burning Cost, et problématique l'appréhension du risque pris par le réassureur.

4.2.2 Modélisation de l'indice $S(T)$ au premier dam/h

Dans le but de déterminer une meilleure estimation de la prime de la couverture indiciaire vent, une alternative peut consister à ajuster une distribution probabiliste aux valeurs empiriques de l'indice $S(T)$ afin d'estimer $E(C(T))$, l'espérance de $C(T)$. De même que le Burning Cost, la recherche d'un ajustement statistique sur les valeurs d'indice constitue une approche immédiate en raison de sa simplicité et de son utilisation répandue.

Nous disposons donc de l'échantillon de $n = 33$ valeurs de l'indice $S(T)$, sur la période de 1970 à 2002 :

Année civile	$S_j(T)$ (dam/h)	Année civile	$S_j(T)$ (dam/h)
1999	4 775,40	1978	66,60
1990	4 417,20	1970	45,00
1976	1 774,80	1979	45,00
1986	756,00	2000	45,00
1972	720,00	1971	36,00
1988	666,00	1973	-
1974	489,60	1975	-
1984	379,80	1977	-
1983	315,00	1980	-
1996	315,00	1981	-
1987	225,00	1982	-
1992	180,00	1985	-
1989	135,00	1991	-
2002	135,00	1993	-
1994	129,60	1998	-
1995	129,60	2001	-
1997	129,60		

Nous utiliserons les ajustements réalisés à partir de quatre lois, la loi exponentielle, la loi gamma, la loi lognormale et la loi de Weibull. La loi Normale, étudiée également, ne sera pas présentée car elle n'offrirait pas un ajustement satisfaisant.

Afin d'appréhender la qualité des ajustements, nous examinerons pour chaque loi étudiée le graphique « quantile-quantile » associé, ou Q-Q plot, et réaliserons les tests e.d.f. de Kolmogorov-Smirnov et d'Anderson-Darling.

▪ Notations

Soit (X_1, \dots, X_n) une réalisation du n-échantillon de la variable aléatoire réelle de l'indice X , de fonction de répartition F , et soit $X_{(1)}, \dots, X_{(n)}$ l'échantillon ordonné correspondant tel que $X_{(1)} \leq \dots \leq X_{(n)}$.

Soit $F = \{F(x, \theta), \theta \in \Theta\}$ une famille paramétrée des fonctions de distribution, où Θ est un ouvert de \mathbb{R}^q .

F_n^* est la fonction de répartition empirique définie par :

$$F_n^*(x) = \frac{1}{n} \sum_{r=1}^n \mathbb{1}_{]-\infty; x[}(X_{(r)})$$

▪ Présentation des tests e.d.f.

Les tests e.d.f. (empirical distribution function) sont réalisés à partir de la fonction de répartition empirique F_n^* .

❖ *Le test de Kolmogorov-Smirnov :*

Soit $\hat{\theta}$ l'estimateur de maximum de vraisemblance de θ sous l'hypothèse $H_0 : F_X \in F$. Le test de Kolmogorov-Smirnov utilise la statistique D_n :

$$D_n = \sup_{x \geq 0} |F_n^*(x) - F(x, \hat{\theta})| = \max(D_n^+, D_n^-)$$

$$\text{ou } D_n^+ = \max_{r=1, \dots, n} \left(\frac{r}{n} - F(X_{(r)}, \hat{\theta}) \right) \text{ et } D_n^- = \max_{r=1, \dots, n} \left(F(X_{(r)}, \hat{\theta}) - \frac{r-1}{n} \right)$$

La statistique D_n permet de comparer en termes de distance plusieurs ajustements entre eux, le meilleur étant celui qui possède la distance D_n la plus faible.

❖ *Le test d'Anderson-Darling :*

Le test d'Anderson utilise la statistique A_n^2 suivante :

$$A_n^2 = n \times \int_0^{+\infty} \frac{(F_n^*(x) - F(x))^2}{F(x) \cdot (1 - F(x))} dF(x) = -n - \frac{1}{n} \times \left(\sum_{r=1}^n (2r+1) (\log(F(X_{(r)}, \hat{\theta})) + \log(1 - F(X_{(n-r+1)}, \hat{\theta}))) \right)$$

Nous utiliserons également la statistique A_n^2 en termes de distance afin de comparer les ajustements, le meilleur d'entre eux correspondant à la valeur la plus faible de A_n^2 .

Le test d'Anderson-Darling donne une importance plus grande aux queues de distribution, nous le préférons donc au test de Kolmogorov-Smirnov en cas de conclusions divergentes entre les deux tests.

▪ **Présentation des Q-Q plots**

Sous l'hypothèse $H_0 : F_X \in F$:

$$X_{(i)} \approx F^{-1} \left(\frac{i}{n+1} \right)$$

Un Q-Q plot est une technique d'analyse exploratoire des données qui permet d'étudier la qualité d'un ajustement en positionnant les points $[(X_{(i)})_{i=1, \dots, n} ; F^{-1}(F_n^*(X_{(i)}))_{i=1, \dots, n}]$ sur un graphique. On utilise souvent la définition suivante de la fonction de répartition empirique :

$$F_n^*(X_{(i)}) = \left(\frac{i-0,5}{n} \right)$$

Pour chaque loi étudiée, nous observerons ainsi l'alignement des points $(X_{(i)} ; F^{-1}\left(\frac{i-0,5}{n}\right))$ avec $i = (1, \dots, n)$. La qualité de l'ajustement sera jugée de façon subjective en comparant l'alignement des points à la droite d'équation $Y=X$.

Nous obtiendrons les valeurs $F^{-1}\left(\frac{i-0,5}{n}\right)$ par simulation des distributions étudiées, à l'aide d'un logiciel permettant l'application de la méthode de Monte Carlo⁶³.

En générant des nombres aléatoires r_i distribués uniformément sur $[0 , 1]$, on obtient des nombres aléatoires x_i distribués selon une loi de fonction de répartition F en transformant les nombres r_i par la fonction inverse de F : $x_i = F^{-1}(r_i)$.

- **La loi exponentielle**

Une loi exponentielle de paramètre a avec $a > 0$, a la fonction de répartition suivante :

$$F(x) = 1 - \exp(-ax)$$

L'estimation de a par la méthode du maximum de vraisemblance nous donne :

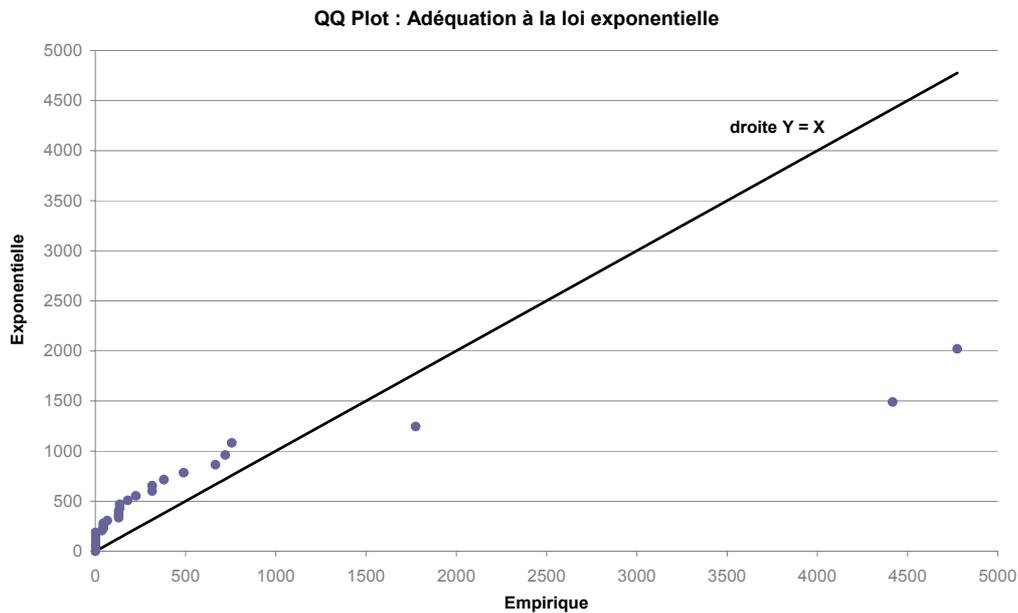
$$\hat{a} = \frac{n}{\sum_{j=1}^n S_j(T)} = \frac{33}{\sum_{j=1}^{33} S_j(T)} = 2,07 \times 10^{-3}$$

Nous simulerons la distribution exponentielle en utilisant :

$$x_i = -\frac{1}{\hat{a}} \log(1 - r_i)$$

⁶³ Il s'agit du logiciel de modélisation financière dynamique, ReMetrica, développé par Benfield Group. Une brève présentation de ReMetrica est située en annexe 3.

❖ *Q-Q plot*



L'observation du graphique indique que l'ajustement par une loi exponentielle n'est pas très approprié.

❖ *Test e.d.f.*

Le test de Kolmogorov-Smirnov vaut 0,39 et le test d'Anderson-Darling vaut 47,43. La loi exponentielle, nous le verrons, n'apparaît pas comme la plus appropriée au vu du résultat de ces tests.

▪ **La loi gamma**

Une loi gamma de paramètres a et b a la fonction de densité suivante :

$$f(x) = \frac{b^a}{\Gamma(a)} \cdot x^{a-1} \cdot \exp(-bx)$$

Nous obtenons les estimations suivantes de a et b :

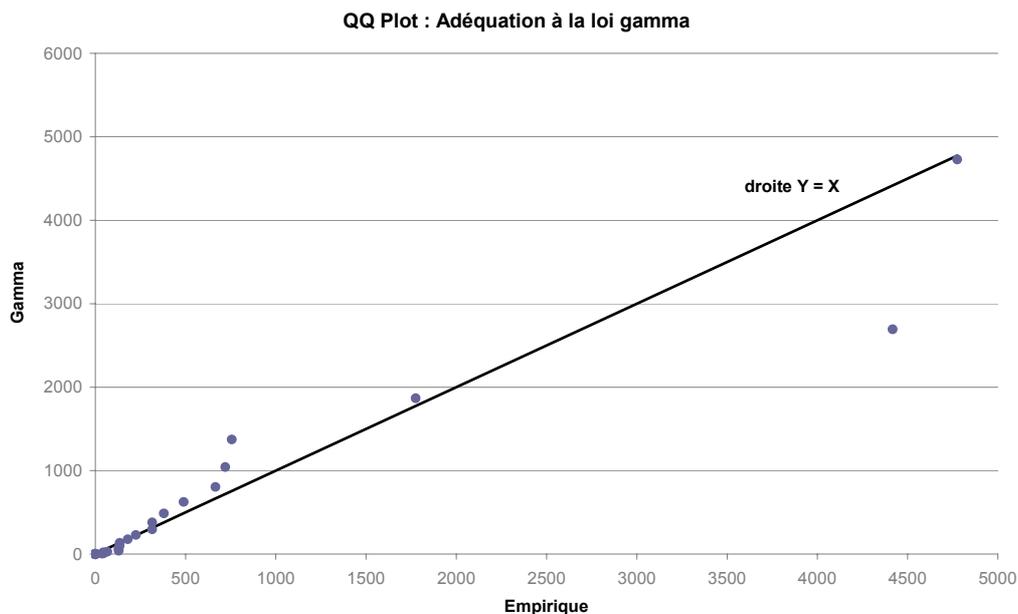
$$\hat{a} = \frac{\left(\sum_{j=1}^n S_j(T) \right)^2}{n \times \left(\sum_{j=1}^n (S_j(T) - \bar{S}(T))^2 \right)} = 0,18553$$

$$\hat{b} = \frac{\left(\sum_{j=1}^n S_j(T) \right)}{\left(\sum_{j=1}^n (S_j(T) - \bar{S}(T))^2 \right)} = 0,00038$$

Nous simulerons la distribution gamma en utilisant l'algorithme suivant :

- ▶ poser $m = [\hat{a}]$ et $q = \hat{a} - m$,
- ▶ générer $z = -\sum_{i=1}^m r_i$,
- ▶ générer w suivant une loi Beta de paramètres q et $1-q$,
- ▶ générer $y = \log(u)$ où u est distribuée uniformément sur $[0,1]$,
- ▶ on obtient alors : $x_i = \frac{(z+wy)}{\hat{b}}$.

❖ Q-Q plot



L'observation graphique indique un ajustement plus satisfaisant.

❖ *Test e.d.f.*

Le test de Kolmogorov-Smirnov vaut 0,26 et le test d'Anderson-Darling vaut 2,45. Ces valeurs sont, nous le verrons, les plus faibles pour les quatre lois étudiées.

▪ **La loi lognormale**

Une loi lognormale de paramètres μ et σ a la fonction de répartition suivante :

$$F(x) = \Phi\left(\frac{\log(x) - \mu}{\sigma}\right)$$

où Φ est la fonction de répartition d'une loi normale centrée réduite.

Nous obtenons les estimations suivantes de μ et σ :

$$\hat{\mu} = \frac{\sum_{j=1}^n \log(S_j(T))}{n} = 1,3916$$

$$\hat{\sigma} = \left(\frac{\sum_{j=1}^n (\log(S_j(T)) - \hat{\mu})^2}{n} \right)^{\frac{1}{2}} = 5,9743$$

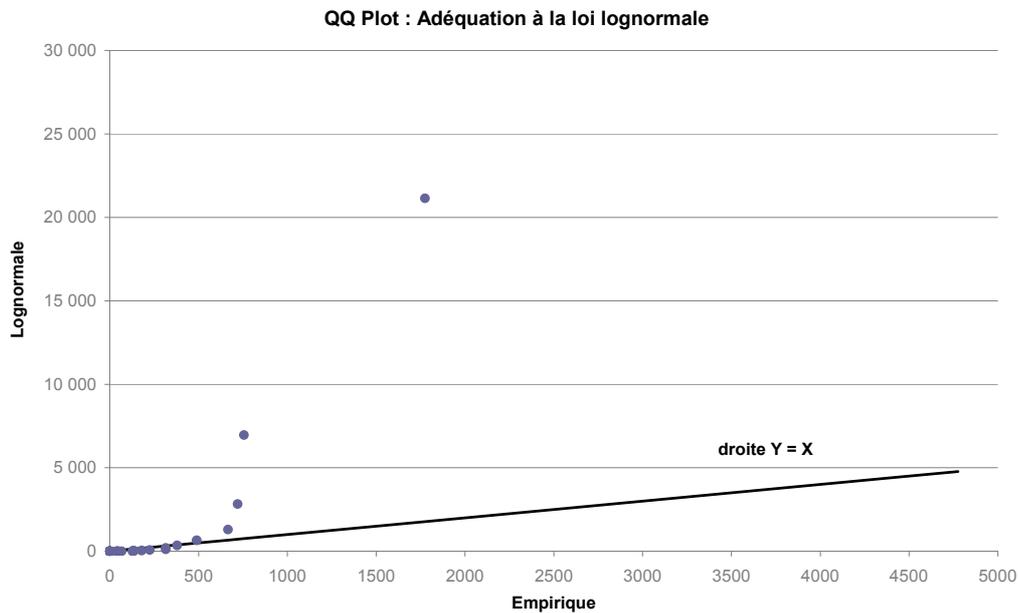
Nous simulerons la distribution lognormale de la façon suivante :

- ▶ Générer r_i , un nombre aléatoire uniforme sur $[0,1]$,
- ▶ calculer $\Phi(r_i)$ où Φ est la fonction de répartition d'une loi normale centrée réduite⁶⁴,
- ▶ on obtient alors :

$$x_i = \exp(\Phi(r_i) \cdot \hat{\sigma} + \hat{\mu})$$

⁶⁴ Voir par exemple « *Practical Risk Theory for Actuaries* » C.D. Daykin, T. Pentikäinen and M. Pesonen pour la méthode de calcul numérique.

❖ Q-Q plot



L'observation du graphique nous indique que l'ajustement n'est pas approprié.

▪ La loi de Weibull

Une loi de Weibull de paramètres a et m a la fonction de répartition suivante :

$$F(x) = 1 - \exp\left(-\left(\frac{x}{a}\right)^m\right)$$

Nous obtenons les estimations de a et m par la méthode du maximum de vraisemblance, par l'intermédiaire de la méthode de Newton, en maximisant sur \hat{a} et m la vraisemblance :

$$EMV = n \times \log\left(\frac{\hat{m}}{\hat{a}}\right) + (\hat{m} - 1) \times \sum_{j=1}^n \log\left(\frac{S_j(T)}{\hat{a}}\right) - \sum_{j=1}^n \left(\frac{S_j(T)}{\hat{a}}\right)^{\hat{m}}$$

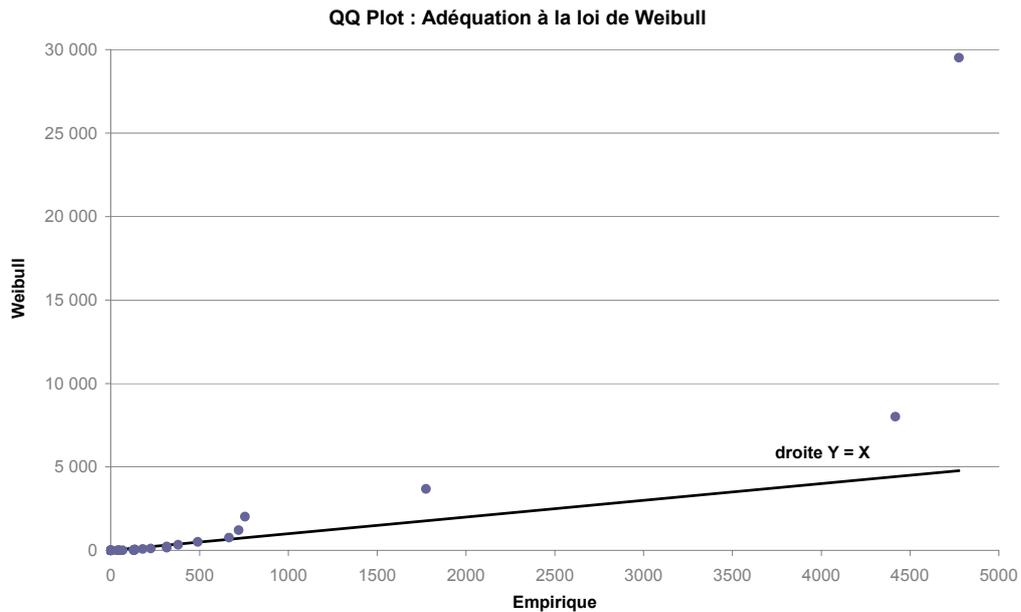
$$\hat{a} = 62,91$$

$$\hat{m} = 0,2329$$

Nous simulerons la distribution de Weibull en utilisant :

$$x_i = \hat{a} \cdot \left(\log \left(\frac{1}{1-r_i} \right) \right)^{1/\hat{m}}$$

❖ Q-Q plot

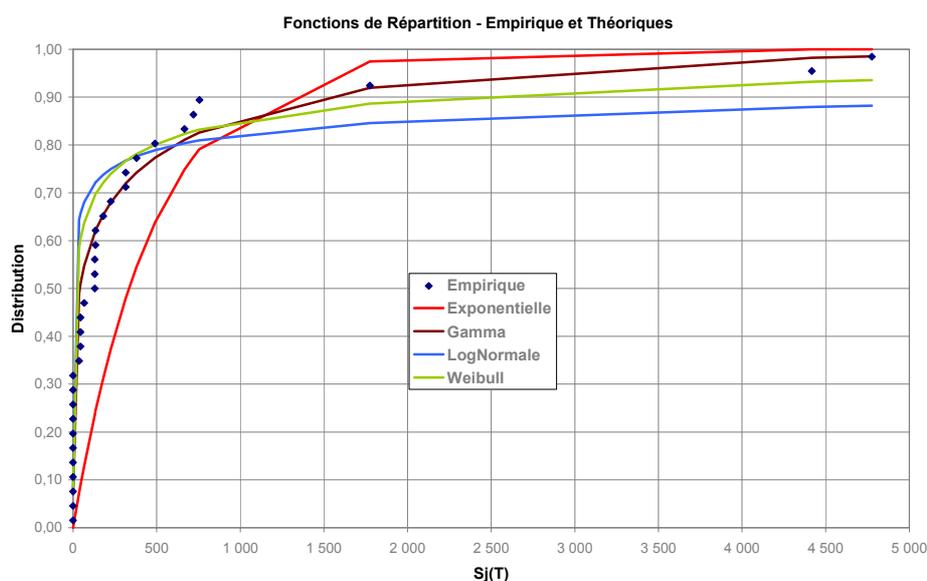


L'ajustement n'apparaît visuellement pas adéquat.

▪ Estimation de la prime de la couverture

Le tableau et le graphique suivants synthétisent les ajustements réalisés :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling	P-Value Kolmog-Smir
Exponentielle	$\hat{c} = 2,07 \cdot 10^{-3}$	0,39	47,43	1%
Gamma	$\hat{a} = 0,18553$ $\hat{b} = 3,8 \cdot 10^{-4}$	0,26	2,45	46%
lognormale	$\hat{\mu} = 1,3916$ $\hat{\sigma} = 5,9743$	0,31	4,04	10%
Weibull	$\hat{a} = 62,91$ $\hat{m} = 0,2329$	0,26	3,33	29%



Au vu des tests réalisés, nous décidons d'estimer la prime de la couverture indicielle vent en utilisant la seule distribution gamma, qui apparaît comme la moins inappropriée. La P-Value associée au test de Kolmogorov-Smirnov pour la loi gamma est effectivement la plus élevée des lois étudiées. Nous estimons la distribution de l'indemnisation $C(T)$ en effectuant 100 000 simulations de l'indice $S(T)$ à l'aide de ReMetrica⁶⁵.

Pour chaque simulation, nous effectuons le calcul de $C(T)$:

$$C(T)=1000 \times \max[\min(4000, S(T)-1000), 0]$$

Soient \bar{C} et s les estimations respectives de la moyenne et de l'écart type de l'indemnisation annuelle de la couverture. En appliquant des chargements de sécurité sur écart type de 20 % et 40 %, nous obtenons les estimations respectives $\Pi_{20\%}$ et $\Pi_{40\%}$ de la prime commerciale. Les résultats obtenus sont décrits dans le tableau suivant :

⁶⁵ Une brève présentation du logiciel de DFA ReMetrica est proposée en annexe 2

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost sur 33 ans	241 436	876 919	416 820	592 204
Modélisation loi Gamma	198 129	679 733	334 076	470 022

Nous obtenons des estimations de la prime de la couverture inférieures à celles basées sur le Burning Cost. Toutefois, les ajustements effectués sur l'ensemble des valeurs historiques de $S_j(T)$ ne sont pas très satisfaisants, dans la mesure où il ne reproduisent qu'imparfaitement le comportement de l'indice.

Il pourrait être judicieux d'étudier des ajustements à l'aide de lois supplémentaires, mais ceci n'a pas été réalisé dans le cadre ce mémoire.

4.2.3 Modélisation de l'indice $S(T)$ avec introduction d'un seuil

L'aspect atypique de la série des valeurs $S_j(T)$, comportant 11 valeurs nulles sur 33, incite à étudier l'introduction d'un seuil dans les ajustements. Toutefois, le choix d'un seuil est problématique, et conduit dans tous les cas à une réduction de la taille déjà faible de l'échantillon.

Nous choisissons cependant d'étudier l'introduction d'un seuil de modélisation $z = 100$ dam/h.

Nous disposons donc, sur la période de 1970 à 2002, de l'échantillon des $n = 17$ valeurs de l'indice $S(T)$ qui sont supérieures à 100 dam/h :

Année civile	$S_j(T)$ (dam/h)
1999	4 775,40
1990	4 417,20
1976	1 774,80
1986	756,00

Année civile	$S_j(T)$ (dam/h)
1972	720,00
1988	666,00
1974	489,60
1984	379,80
1983	315,00
1996	315,00
1987	225,00
1992	180,00
1989	135,00
2002	135,00
1994	129,60
1995	129,60
1997	129,60

Nous estimerons la probabilité annuelle d'occurrence d'un indice $S(T)$ de valeur supérieure à 100 dam/h à :

$$p = \frac{\text{nombre d'observations d'indices } S_j(T) \text{ supérieurs à 100 dam/h}}{\text{nombre total d'observations}} = \frac{17}{33} = 0,515$$

Nous modéliserons la valeur des indices $S(T)$ supérieurs à 100 dam/h en utilisant six lois : la loi exponentielle, la loi de Pareto, la loi lognormale, la loi de Weibull, la loi de Pareto Généralisée, et la loi dite « Generalized Extreme Value Distribution ».

Afin d'appréhender la qualité des ajustements, nous examinerons, comme pour les ajustements précédents, les graphiques « quantile-quantile », ou Q-Q plots, et réaliserons les tests e.d.f. de Kolmogorov-Smirnov et d'Anderson-Darling.

- **La loi exponentielle tronquée**

Une loi exponentielle tronquée de paramètre a et z avec $a > 0$ et $z > 0$, a la fonction de répartition suivante :

$$F(x) = 1 - \exp(-a(x-z)) \text{ avec } x \geq z$$

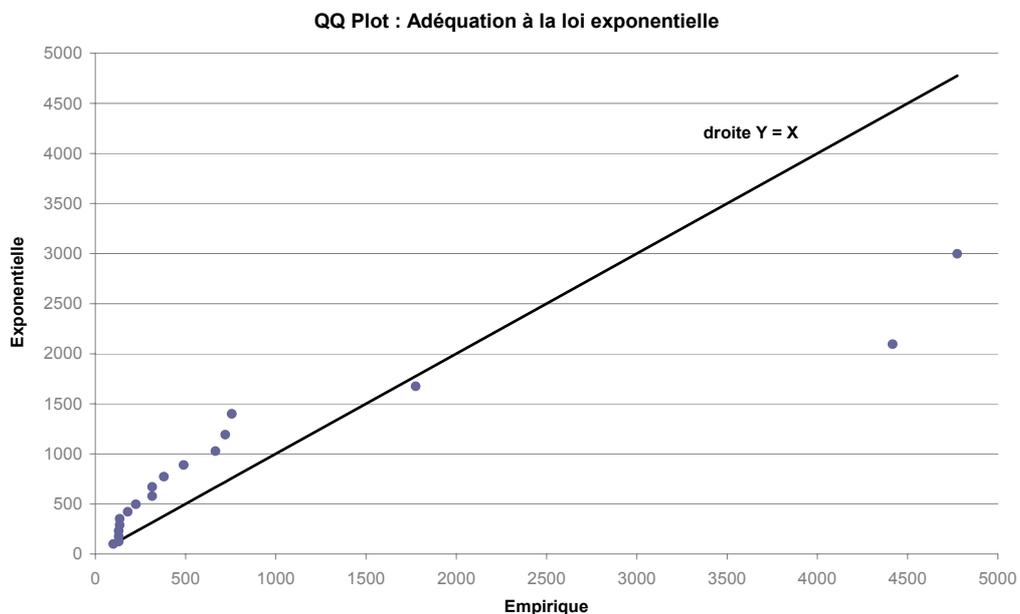
L'estimation de a par la méthode du maximum de vraisemblance nous donne :

$$\hat{a} = \frac{n}{\sum_{j=1}^n (S_j(T) - z)} = \frac{17}{\sum_{j=1}^{17} (S_j(T) - 100)} = 1,22 \times 10^{-3}$$

Nous simulerons la distribution exponentielle tronquée en utilisant :

$$x_i = z - \frac{1}{a} \log(1 - r_i)$$

- ❖ *Q-Q plot*



L'observation du graphique indique que l'ajustement par une loi exponentielle n'est pas très approprié.

- **La loi de Pareto**

Une loi de Pareto de paramètres a et z a la fonction de répartition suivante :

$$F(x) = 1 - \left(\frac{z}{x}\right)^a \quad \text{avec } x \geq z$$

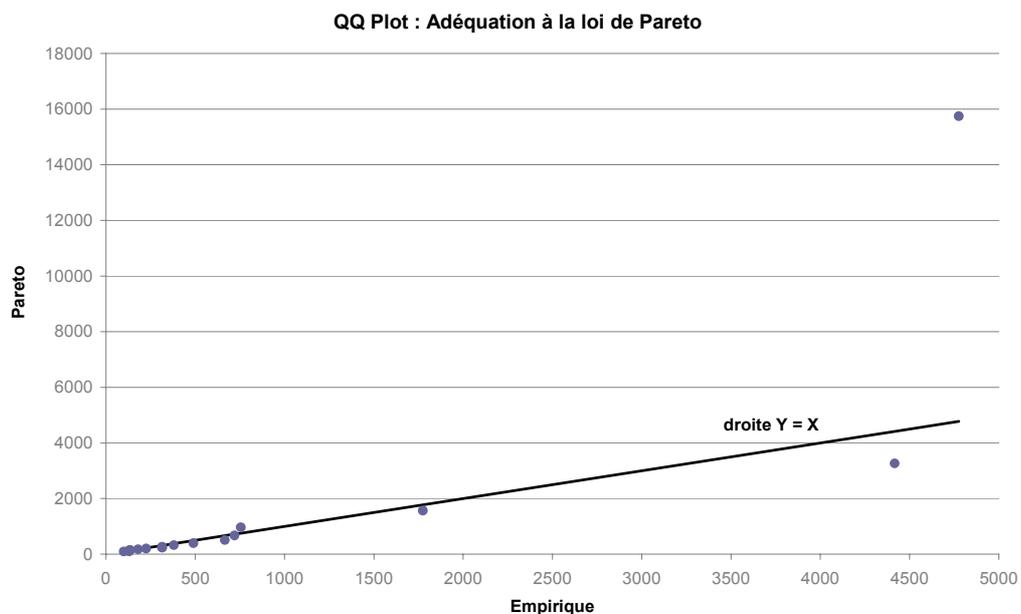
L'estimation de a par la méthode du maximum de vraisemblance :

$$\hat{a} = \frac{n}{\sum_{j=1}^n \log\left(\frac{S_j(T)}{z}\right)} = 0,696$$

Nous simulerons la distribution de Pareto en utilisant :

$$x_i = \frac{z}{(1-r_i)^{1/\hat{a}}}$$

- ❖ **Q-Q plot**



L'ajustement ne semble pas complètement approprié, en particulier pour la queue de distribution.

- **La loi lognormale tronquée**

Une loi lognormale tronquée de paramètres μ et σ et de seuil z a la fonction de répartition suivante :

$$F(x) = \frac{\Phi\left(\frac{\log(x)-\mu}{\sigma}\right) - \Phi\left(\frac{\log(z)-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{\log(z)-\mu}{\sigma}\right)} \quad \text{avec } x \geq z$$

où Φ est la fonction de répartition d'une loi normale centrée réduite.

Nous estimons μ et σ de la façon suivante :

$$\hat{\mu} = \frac{\sum_{j=1}^n \log(S_j(T))}{n} - \left(\frac{\sum_{j=1}^n \log(S_j(T))}{n} - \log(z) \right) \times \left(\frac{Z(\hat{u})}{Z(\hat{u}) - \hat{u}} \right)$$

$$\hat{\sigma} = \left[\frac{\sum_{j=1}^n (\log(S_j(T)) - \mu)^2}{n} + \left(\frac{\sum_{j=1}^n \log(S_j(T))}{n} - \log(z) \right)^2 \times \left(\frac{Z(\hat{u})}{Z(\hat{u}) - \hat{u}} \right)^2 \right]^{1/2}$$

où Z est une fonction définie sur \mathbb{R} telle que $Z(x) = \varphi(x) / (1 - \Phi(x))$, φ et Φ respectivement les fonctions de densité et de répartition de la loi normale centrée réduite.

et où $\hat{u} = \frac{\log(z) - \hat{\mu}}{\hat{\sigma}}$ est l'unique solution de l'équation :

$$\frac{1 - Z(u) \cdot [Z(u) - u]}{[Z(u) - u]^2} = \frac{\sum_{j=1}^n (\log(S_j(T)) - \mu)^2}{n \times \left(\frac{\sum_{j=1}^n \log(S_j(T))}{n} - \log(z) \right)^2}$$

Après avoir déterminé la valeur de $\hat{\mu}$ par la méthode de Newton (estimation de la tangente fournie par le solveur Excel), nous obtenons les estimations suivantes de μ et σ :

$$\hat{\mu}=3,7316$$

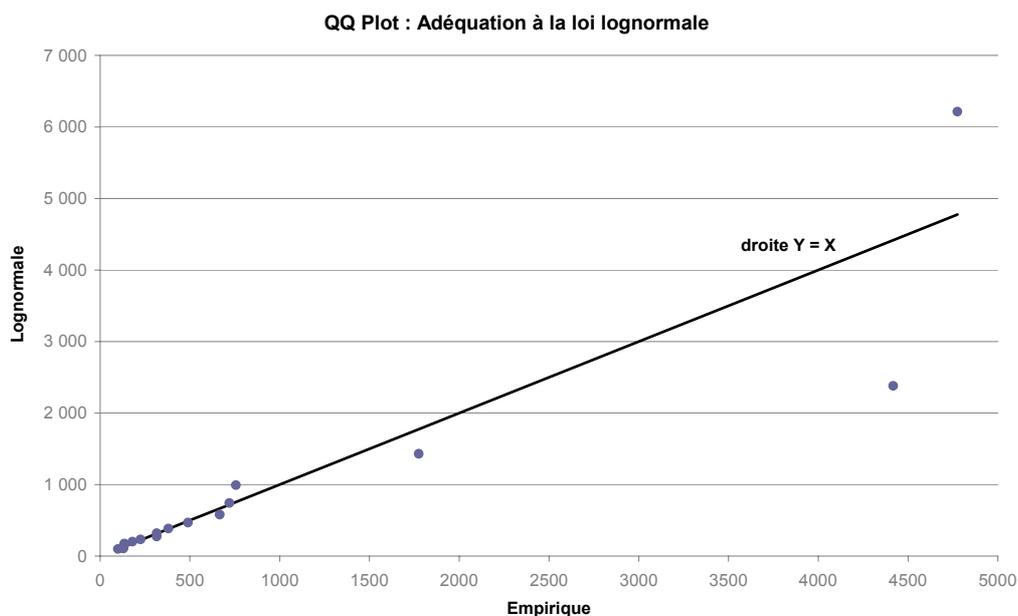
$$\hat{\sigma}=2,15307$$

Nous simulerons la distribution lognormale de la façon suivante :

- ▶ générer r_i , un nombre aléatoire uniforme sur $[0,1]$,
- ▶ calculer $l_i = \Phi\left(r_i ; \left(1 - \Phi\left(\frac{\log(z) - \hat{\mu}}{\hat{\sigma}}\right)\right) + \Phi\left(\frac{\log(z) - \hat{\mu}}{\hat{\sigma}}\right)\right)$ où Φ est la fonction de répartition d'une loi normale centrée réduite⁶⁶,
- ▶ on obtient alors :

$$x_i = \exp(l_i \cdot \hat{\sigma} + \hat{\mu})$$

❖ Q-Q plot



⁶⁶ Voir par exemple « Practical Risk Theory for Actuaries » C.D.Daykin, T.Pentikäinen and M.Pesonen pour la méthode de calcul numérique.

L'ajustement semble plus satisfaisant.

❖ *Test e.d.f.*

Le test de Kolmogorov-Smirnov vaut 0,15 et le test d'Anderson-Darling vaut 0,40. Ces valeurs sont, nous le verrons, les plus faibles parmi celles des six lois étudiées.

▪ **La loi de Weibull tronquée**

Une loi de Weibull tronquée de paramètres a et m , et de seuil z , a la fonction de répartition suivante :

$$F(x) = 1 - \exp\left(-\left(\frac{x-z}{a}\right)^m\right) \text{ avec } x \geq z$$

Nous obtenons les estimations de a et m par la méthode du maximum de vraisemblance, par l'intermédiaire de la méthode de Newton, en maximisant sur \hat{a} et \hat{m} la vraisemblance :

$$EMV = n \times \log\left(\frac{\hat{m}}{\hat{a}}\right) + (\hat{m}-1) \times \sum_{j=1}^n \log\left(\frac{S_j(T)-z}{\hat{a}}\right) - \sum_{j=1}^n \left(\frac{S_j(T)-z}{\hat{a}}\right)^{\hat{m}}$$

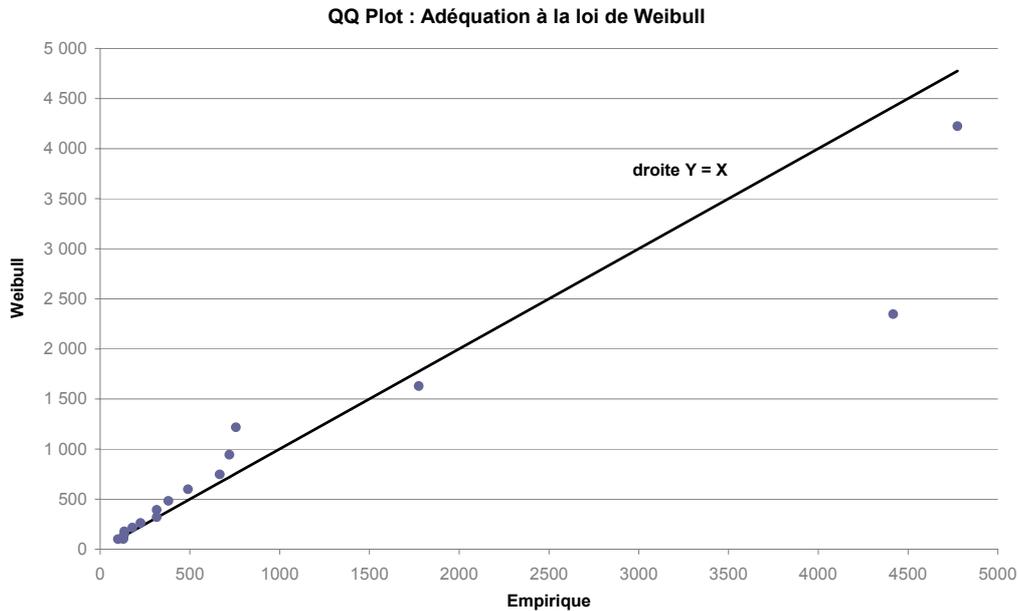
$$\hat{a} = 531$$

$$\hat{m} = 0,615$$

Nous simulerons la distribution de Weibull tronquée en utilisant :

$$x_i = z + \hat{a} \cdot \left(\log\left(\frac{1}{1-r_i}\right) \right)^{1/\hat{m}}$$

❖ Q-Q plot



L'ajustement semble satisfaisant, hormis la queue de distribution.

▪ **La loi de Pareto Généralisée**

La loi de Pareto Généralisée a été fréquemment utilisée, par l'intermédiaire de la méthode POT, ou « Peaks Over Threshold », pour l'étude du comportement des valeurs extrêmes des variables climatiques.

Une loi de Pareto Généralisée de paramètres a et c , et de seuil z , a la fonction de répartition suivante :

$$F(x) = 1 - \left(1 + c \cdot \frac{(x-z)}{a} \right)^{-1/c} \quad \text{avec } x \geq z \text{ et } c \neq 0$$

Nous obtenons les estimations de a et c par la méthode du maximum de vraisemblance, par l'intermédiaire de la méthode de Newton, en maximisant sur \hat{a} et c la log-vraisemblance :

$$\ln EMV = \sum_{j=1}^n \left(-\ln(\hat{a}) - \left(1 + \frac{1}{\hat{c}} \right) \cdot \ln \left(1 + \hat{c} \left(\frac{S_j(T) - z}{\hat{a}} \right) \right) \right)$$

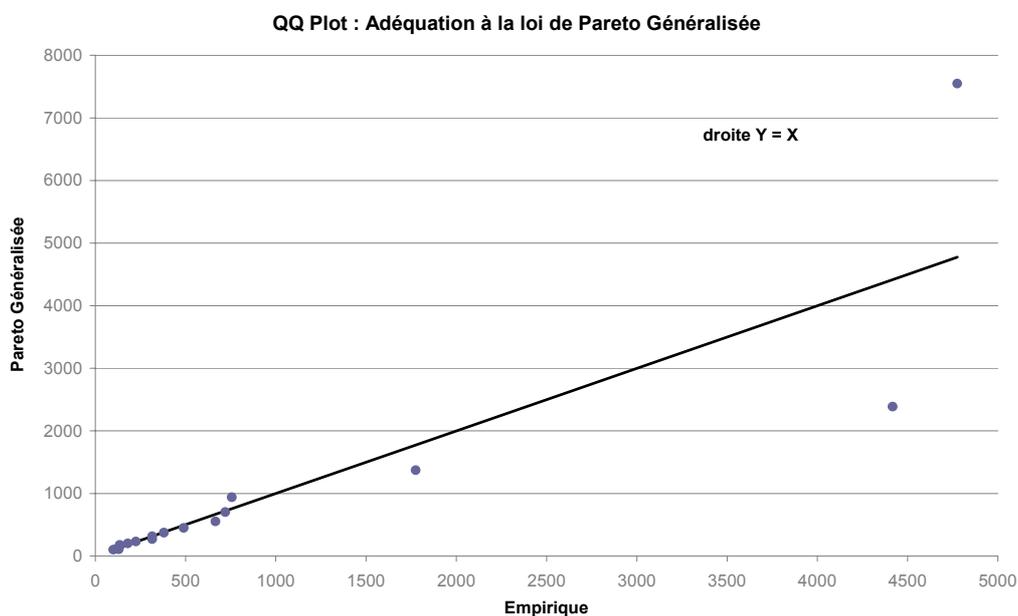
$$\hat{a}=212,87$$

$$\hat{c}=1,0227$$

Nous simulerons la distribution de Pareto généralisée en utilisant :

$$x_i = z + \hat{a} \cdot \frac{(1-r_i)^{-\hat{c}} - 1}{\hat{c}}$$

❖ *Q-Q plot*



L'ajustement semble plus satisfaisant.

▪ **La loi GEV**

La loi dite « Generalized Extreme Value Distribution », ou GEV, a, comme la loi de Pareto Généralisée, été fréquemment utilisée pour l'étude du comportement des valeurs extrêmes des variables climatiques, en se fondant sur les maxima d'une série d'observation de variables aléatoires.

Soient X_1, \dots, X_n n variables aléatoires i.i.d., la théorie des valeurs extrêmes consiste, principalement, à étudier la loi de $Y_n = \max (X_1, \dots, X_n)$.

Le théorème de Fisher-Tippet montre que, pour n suffisamment grand, il existe à la limite une distribution $F(X)$ de Y_n , qui peut être de trois types :

❖ Gumbel : $F(x) = \exp(-\exp(-x))$ avec $-\infty < x < +\infty$

❖ Fréchet : $F(x) = \exp(-x^{-\alpha})$ pour $x \geq 0$ et avec $\alpha > 0$
 $F(x) = 0$ pour $x \leq 0$

❖ Weibull : $F(x) = \exp(-(-x^\alpha))$ pour $x \leq 0$ et avec $\alpha > 0$
 $F(x) = 1$ pour $x \geq 0$

Ces trois lois peuvent être regroupées par la loi dite GEV, qui a pour fonction de répartition :

$$G(x) = \exp\left(-\left(1 + \xi \frac{x - \mu}{\psi}\right)^{-\frac{1}{\xi}}\right) \text{ avec } 1 + \xi \frac{x - \mu}{\psi} > 0 \text{ et } \psi > 0$$

Quand $\xi \rightarrow 0$, on retrouve la distribution de Gumbel,

Quand $\xi > 0$, on retrouve la distribution de Fréchet,

Quand $\xi < 0$, on retrouve la distribution de Weibull.

Nous obtenons les estimations de μ , ψ et ξ par la méthode du maximum de vraisemblance en maximisant la log-vraisemblance :

$$\ln EMV = \sum_{j=1}^n \left(-\ln(\psi) - \left(1 + \frac{1}{\xi}\right) \ln\left(1 + \xi \left(\frac{S_j(T) - \mu}{\psi}\right)\right) - \left(1 + \xi \left(\frac{S_j(T) - \mu}{\psi}\right)\right)^{-\frac{1}{\xi}} \right)$$

$$\hat{\mu} = 260,6$$

$$\hat{\psi} = 212,87$$

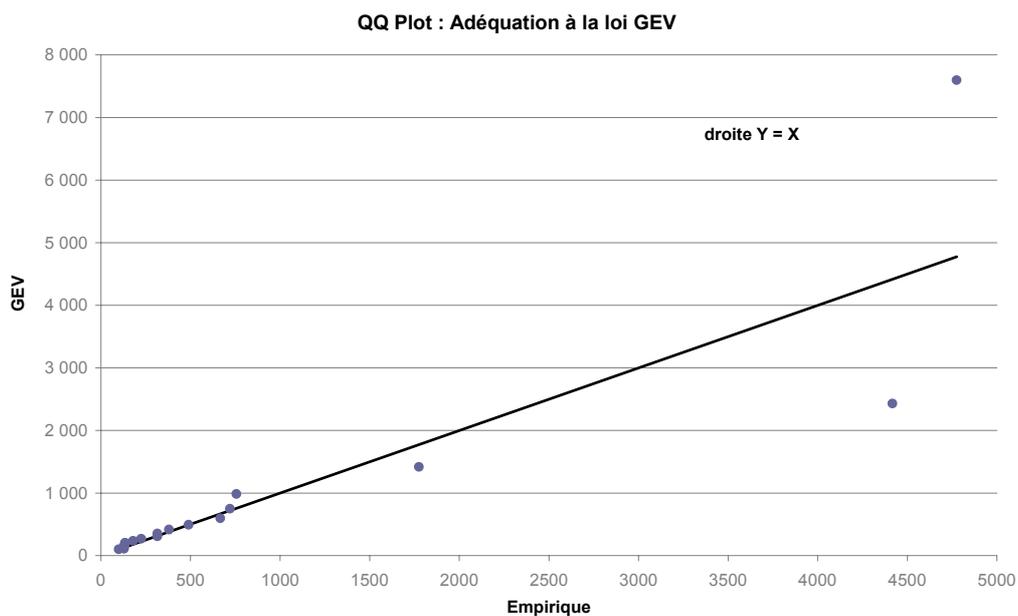
$$\hat{\xi} = 1,0227$$

L'optimisation n'étant pas toujours aisée, nous avons confronté lorsque cela été possible les estimations obtenues à partir du solveur d'Excel aux estimations obtenues notamment par l'utilisation sous le logiciel *R* de la fonction *gev* écrite par Alec Stephenson. Les estimations des paramètres de la loi de Pareto généralisée peuvent également servir de bonnes valeurs de départ pour l'optimisation.

Nous simulerons la distribution GEV en utilisant :

$$x_i = \hat{\mu} + \hat{\eta} \cdot \frac{(-\log(r_i))^{-\frac{1}{\hat{\xi}}} - 1}{\hat{\xi}}$$

❖ *Q-Q plot*

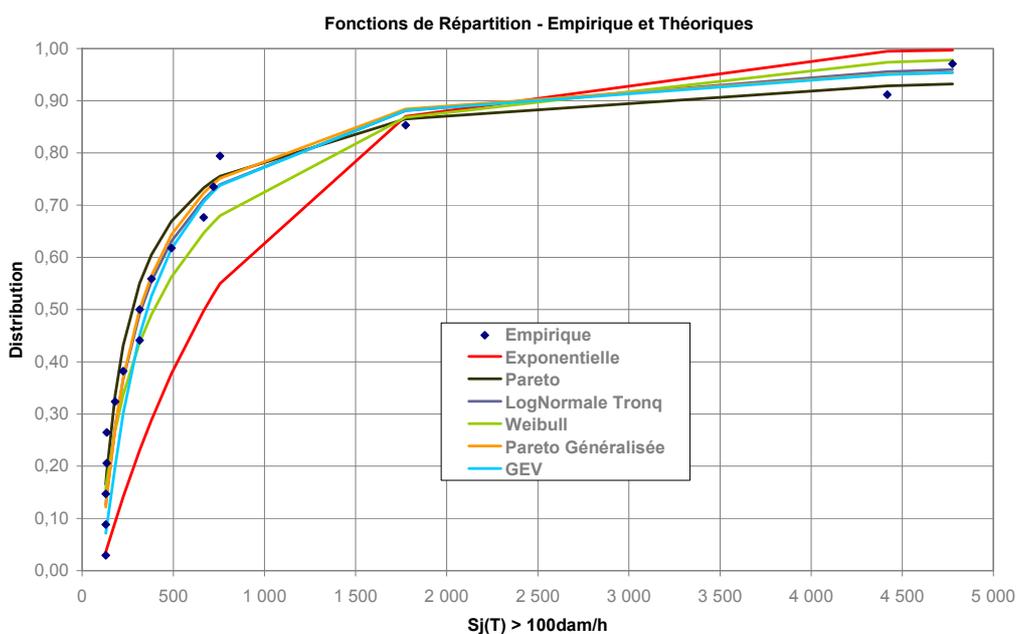


L'ajustement apparaît comme satisfaisant.

▪ **Estimation de la prime de la couverture**

Le tableau et le graphique suivants synthétisent les ajustements réalisés :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling	P-Value Kolmog-Smir
Exponentielle	$\hat{c} = 1,22 \cdot 10^{-3}$	0,30	4,13	47%
Pareto	$\hat{a} = 0,696$	0,17	0,48	+ de 95%
lognormale	$\hat{\mu} = 3,7316$ $\hat{\sigma} = 2,1531$	0,15	0,40	+ de 95%
Weibull	$\hat{a} = 531$ $\hat{m} = 0,615$	0,16	0,59	+ de 95%
Pareto Généralisée	$\hat{a} = 212,87$ $\hat{c} = 1,0227$	0,15	0,41	+ de 95%
GEV	$\hat{\mu} = 260,6$ $\hat{\psi} = 212,9$ $\hat{\xi} = 1,0227$	0,21	0,77	75%



Au vu des tests réalisés, nous décidons d'estimer la prime de la couverture indicielle vent en utilisant les distributions de Pareto, LogNormale, Weibull, Pareto Généralisée et GEV. En effet les observations des QQ-Plots ainsi que les résultats des tests de Kolmogorov-Smirnov et d'Anderson-Darling (statistiques les plus faibles et p-value plus élevées) nous indiquent que ces distributions sont appropriées.

Nous estimons la distribution de l'indemnisation $C(T)$ en effectuant 100 000 simulations. Pour chaque simulation, nous déterminons tout d'abord l'occurrence éventuelle d'un indice $S(T)$ supérieur à 100 dam/h, en réalisant un tirage sur $(0 ; 1)$ avec $p_0 = 1-p = 48,5 \%$ et $p_1 = p = 51,5 \%$, puis en cas d'occurrence, simulons la valeur de $S(T)$.

Pour chaque simulation, nous effectuons le calcul de $C(T)$:

$$C(T)=1000 \times \max[\min(4000, S(T)-1000), 0]$$

Soient \bar{C} et s les estimations respectives de la moyenne et de l'écart type de l'indemnisation annuelle de la couverture. En appliquant des chargements de sécurité sur écart type de 20 % et 40 %, nous obtenons les estimations respectives $\Pi_{20\%}$ et $\Pi_{40\%}$ de la prime commerciale. Les résultats obtenus sont décrits ci-après :

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Modélisation loi de Pareto	217 762	811 985	377 159	539 556
Modélisation loi lognormale	171 193	686 175	308 428	445 663
Modélisation loi de Weibull	165 354	611 614	287 677	410 000
Modélisation loi de Pareto Généralisée	172 808	703 864	313 581	454 356
Modélisation loi GEV	176 692	708 879	318 468	460 243

Les lois utilisées conduisent à nouveau à des estimations plutôt basses de la prime de la couverture, quand on les compare aux estimations basées sur le Burning Cost. Les lois Lognormale, Weibull, Pareto Généralisée et GEV, bien que conduisant à des ajustements satisfaisants dans l'ensemble, tendent à sous-estimer le poids de la queue de distribution de l'indice $S(T)$, notamment au delà de 1000 dam/h. La loi de Pareto permet une meilleure modélisation du comportement de l'indice $S(T)$ entre 1000 et 5000 dam/h, mais est caractérisée par une queue de distribution épaisse au-delà. Les estimations de la prime à partir de la loi de Pareto correspondent assez fidèlement aux estimations Burning Cost sur l'ensemble de l'historique (33 ans).

L'approche de modélisation de l'indice $S(T)$ par des lois de coût permet une estimation satisfaisante de la prime de la couverture étudiée en exemple. Il est à noter que l'étude de lois supplémentaires pourrait être intéressante.

Toutefois, la méthode peut présenter des limites si l'on souhaite étudier le prix de couvertures aux caractéristiques différentes, notamment en matière de niveau de priorité et de plafond. Le choix des distributions et des seuils de troncature appropriés pourrait rapidement se présenter comme un problème épineux. De plus, la méthode pourrait manquer de souplesse et s'avérer fastidieuse dans le cadre de l'étude d'une multitude de couvertures, reposant sur des panels de stations différents.

4.2.4 Modélisation des indices $S_i(T)$

En vue de gagner en rapidité lors de la cotation de couvertures indiciaires aux caractéristiques différentes, une modélisation individuelle des indices par station $S_i(T)$ pourrait constituer une approche intéressante. En effet, la modélisation des indices de chaque station permettrait d'étudier plus aisément l'impact sur le prix de l'ajout ou du retrait de stations dans la couverture.

Cependant, cette approche multiplie les problèmes de choix des seuils et des distributions. Elle ne permettrait pas non plus d'étudier facilement l'impact de modifications des franchises K_i par station. Enfin, les séries d'indices $S_{ij}(T)$ comportent souvent une très grande majorité de valeurs nulles, ce qui rend difficile le choix de lois pour les modéliser. Nous décidons donc de ne pas aller plus avant dans cette voie.

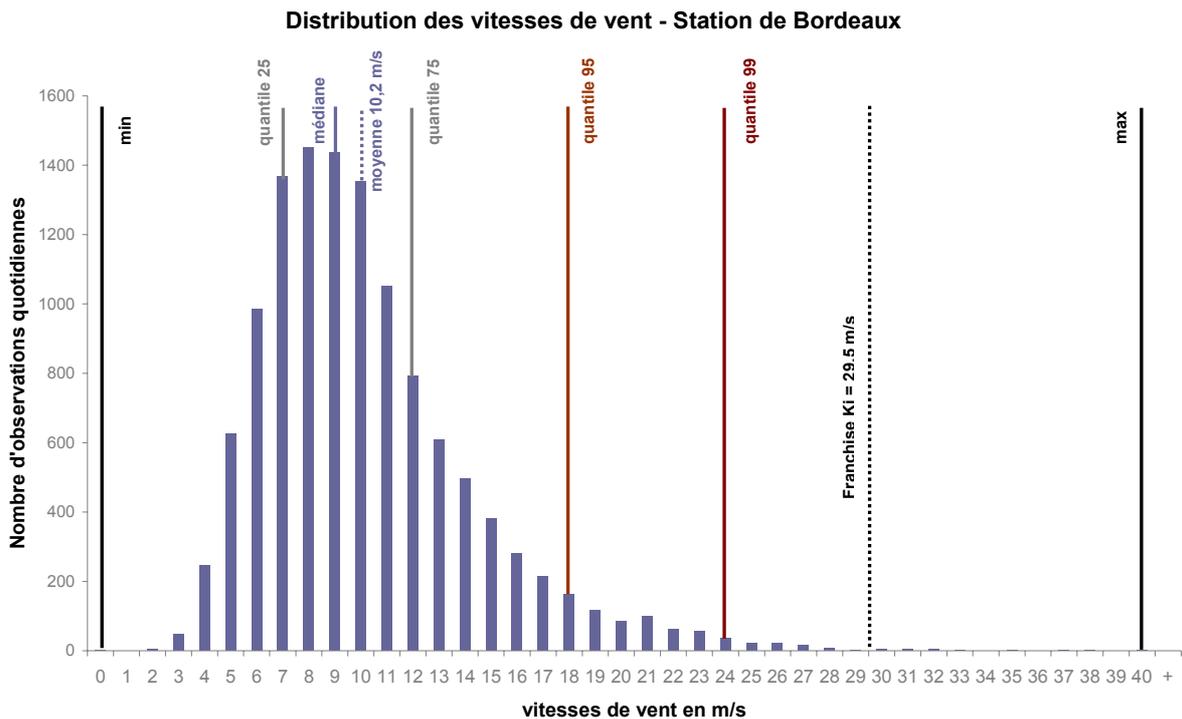
4.3 Tarifications sur les valeurs quotidiennes de vitesse de vent

Les approches abordées précédemment présentent des avantages de simplicité et de rapidité dans le cadre de l'étude d'une couverture indiciaire aux caractéristiques bien précises. Elles présentent cependant l'inconvénient de fournir des estimations qui paraissent très sensibles à un historique de taille relativement réduite de valeurs d'indices, qui comporte une très grande volatilité.

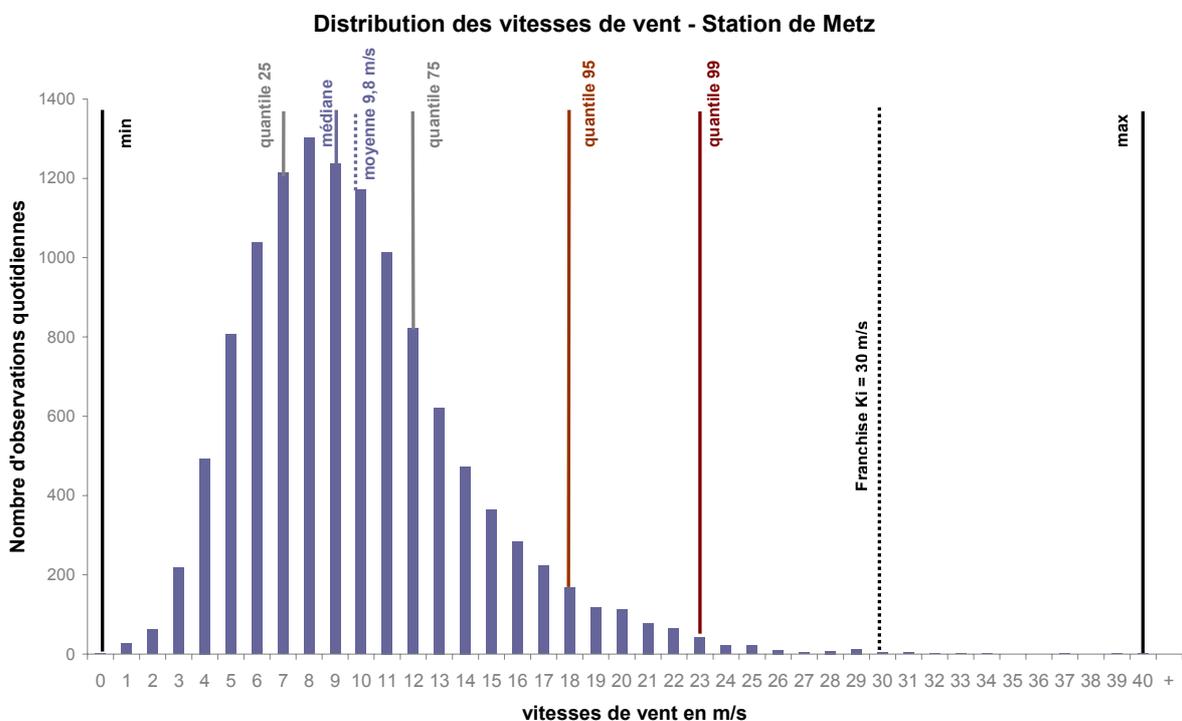
Il paraît donc intéressant d'exploiter les historiques des valeurs de la variable climatique sous-jacente elle-même. L'obtention d'ajustements fiables des séries de données climatiques pour chaque station pourrait être une aide utile dans le cadre de l'étude du prix des couvertures indiciaires en général.

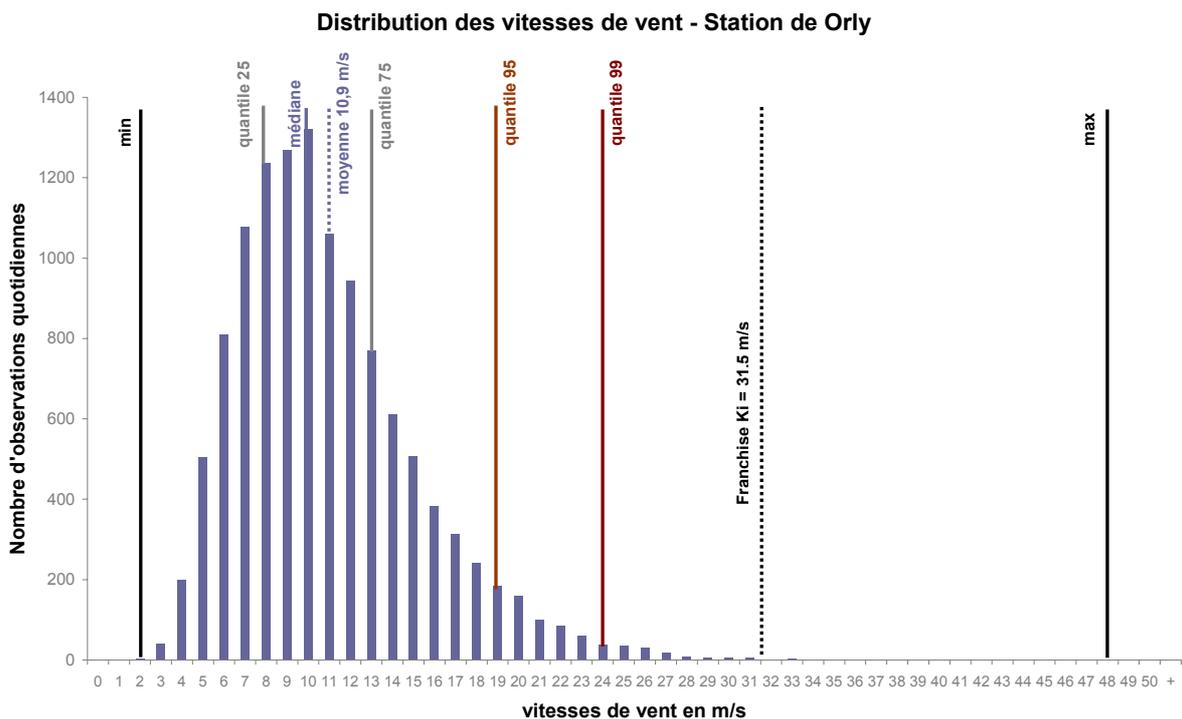
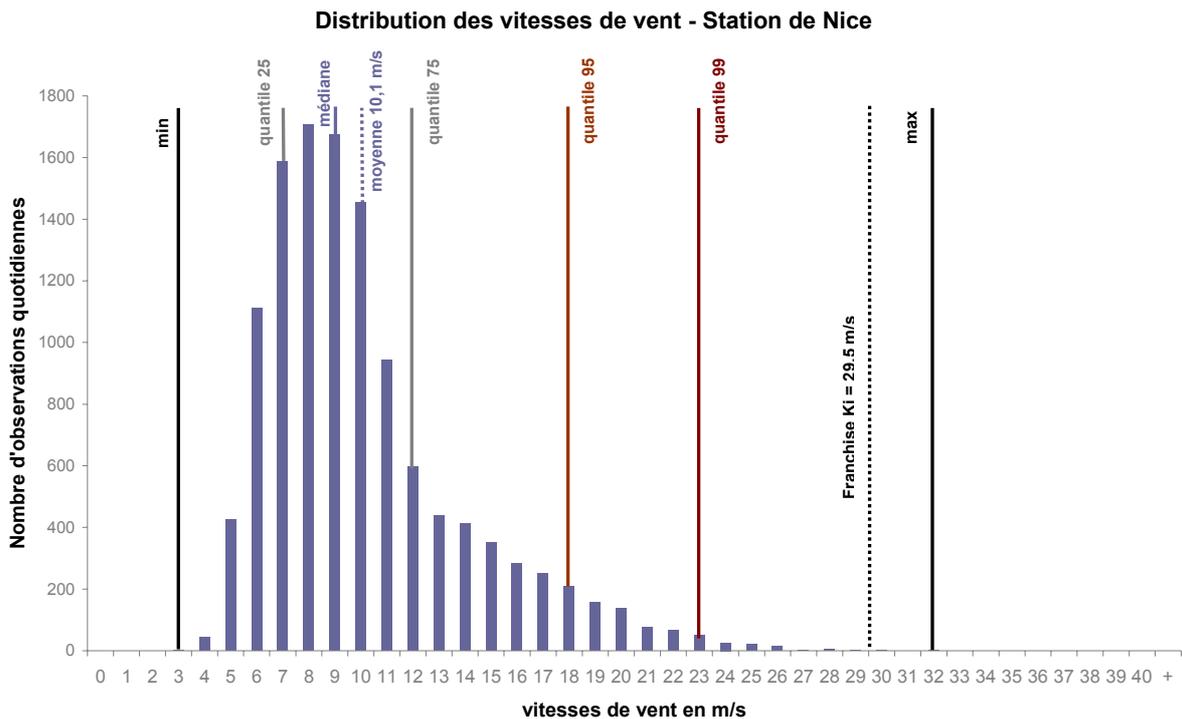
4.3.1 Présentation de la démarche

Nous allons poursuivre l'étude de l'exemple de la couverture indiciaire vent. Nous disposons des historiques de vitesses de vent maximales instantanées pour les quatre stations retenues dans notre exemple. Ces historiques sont composés de 12 053 valeurs quotidiennes de vitesses de vent, mesurées en m/s. Il s'agit de valeurs entières. Les histogrammes suivants représentent la distribution des vitesses de vent quotidiennes pour chacune des quatre stations.



Pour la station de Bordeaux, on peut noter que 50 % des valeurs sont comprises entre 7 m/s et 12 m/s. Nous indiquons la franchise retenue pour Bordeaux (29,5 m/s), et on observe qu'elle est supérieure au quantile 99 %. Ceci est également le cas pour les trois autres stations.





Dans un premier temps, nous étudierons les distributions de chaque station dans leur intégralité, puis nous nous focaliserons plus spécifiquement dans un second temps sur les observations de vitesses de vents supérieures à un seuil.

4.3.2 Ajustements au premier m/s

Soit $X_{i,t}$ la variable aléatoire représentant la valeur de la vitesse de vent enregistrée dans la station i à la date t . Nous modéliserons $X_{i,t}$ pour chaque station en utilisant quatre lois : la loi gamma, la loi lognormale, la loi de Weibull, et la loi GEV. Les lois exponentielle et Pareto généralisée ont été testées, mais les ajustements n'étaient clairement pas appropriés.

Afin d'appréhender la qualité des ajustements, nous examinerons, comme pour les ajustements précédents, les graphiques « quantile-quantile », ou Q-Q Plots, et réaliserons les tests e.d.f. de Kolmogorov-Smirnov et d'Anderson-Darling.

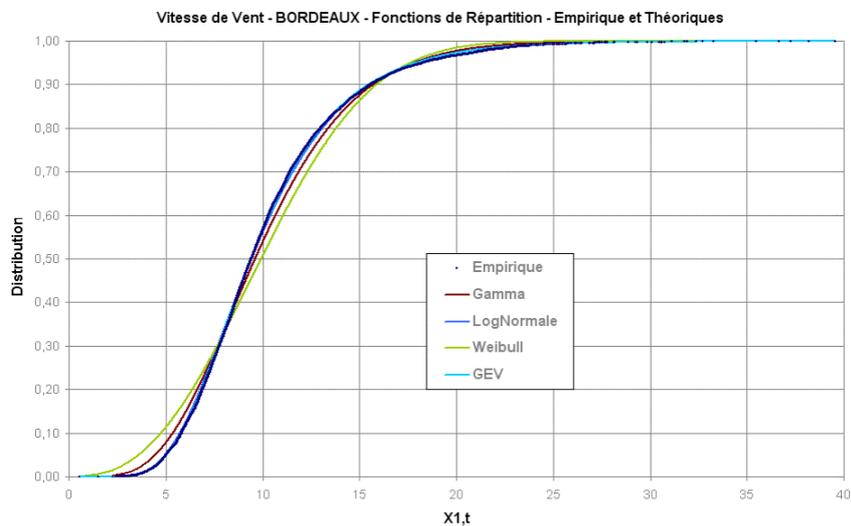
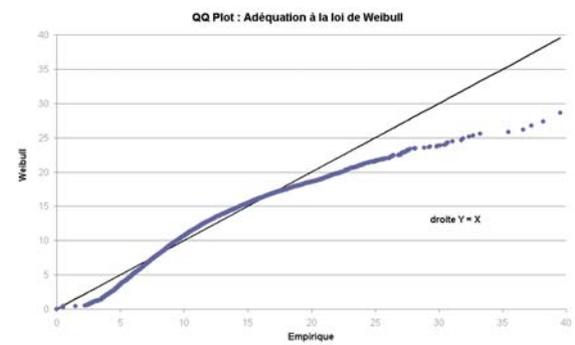
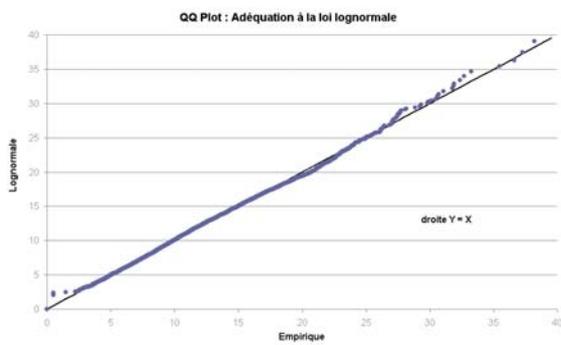
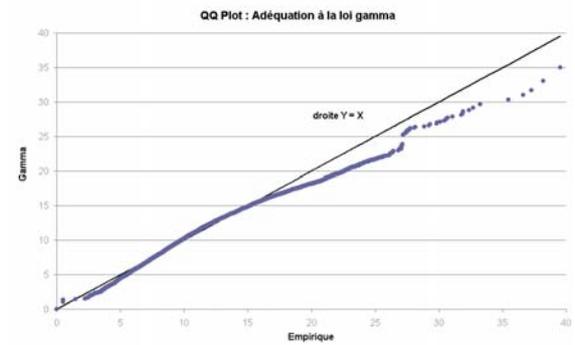
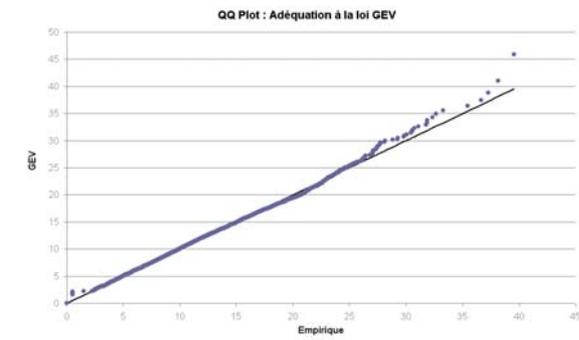
Si les estimations des paramètres des lois étudiées sont effectuées à partir des échantillons originaux des $X_{i,t}$, la réalisation des tests e.d.f. et les représentations graphiques sont fondées sur des échantillons légèrement modifiés, afin de transformer les séries de valeurs entières en séries de valeurs réelles. A chaque $X_{i,t}$ est ajouté une valeur ε choisie aléatoirement de façon uniforme sur l'intervalle $[-0,5 ; +0,5]$.

- **Station de Bordeaux – Ajustements des $X_{1,t}$**

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Gamma	$\hat{a} = 6,059$ $\hat{b} = 1,68$	0,040	46,15
Lognormale	$\hat{\mu} = 2,243$ $\hat{\sigma} = 0,389$	0,016	3,85
Weibull	$\hat{a} = 11,45$ $\hat{m} = 2,557$	0,075	178,78
GEV	$\hat{\mu} = 8,289$ $\hat{\psi} = 3,038$ $\hat{\xi} = 0,039$	0,010	1,29

Les graphiques suivants permettent d'observer la qualité des ajustements :



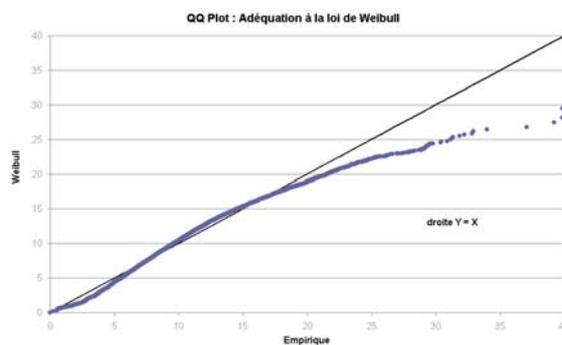
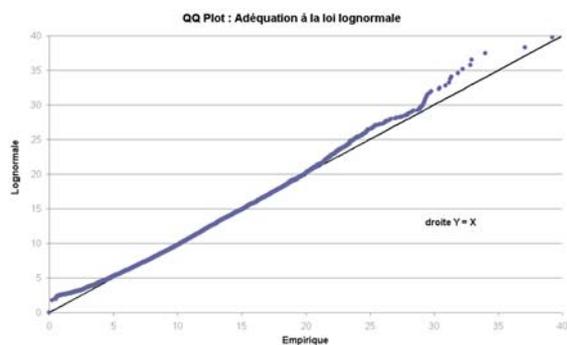
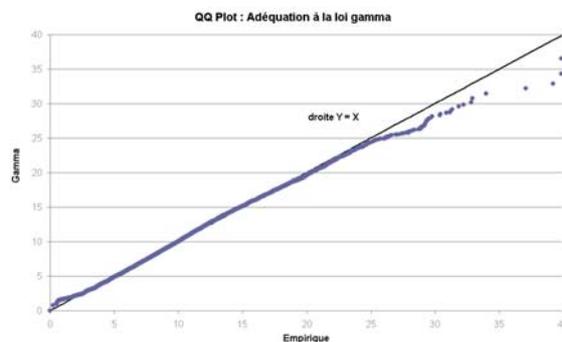
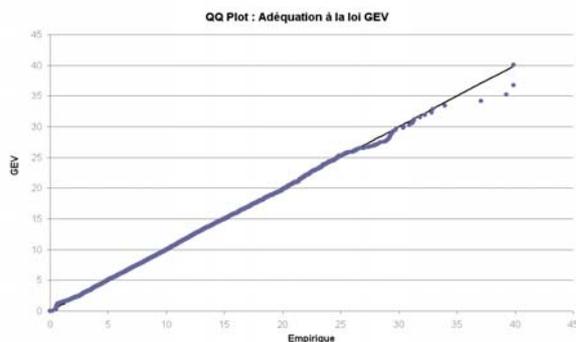
Nous retiendrons pour la suite de l'analyse les ajustements par les lois lognormale et GEV, qui paraissent les plus satisfaisants au vu des tests réalisés (observation graphique et valeurs les plus faibles des tests statistiques).

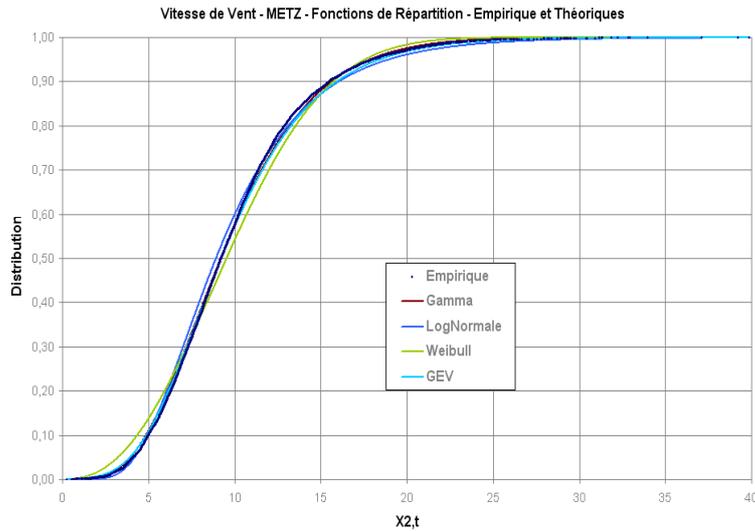
▪ **Station de Metz – Ajustements des $X_{2,t}$**

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Gamma	$\hat{a} = 5,173$ $\hat{b} = 1,90$	0,019	7,35
Lognormale	$\hat{\mu} = 2,188$ $\hat{\sigma} = 0,458$	0,034	27,09
Weibull	$\hat{a} = 11,10$ $\hat{m} = 2,396$	0,049	75,45
GEV	$\hat{\mu} = 7,908$ $\hat{\psi} = 3,414$ $\hat{\xi} = -0,014$	0,008	0,824

Les QQ Plots suivants permettent d'observer la qualité des ajustements :





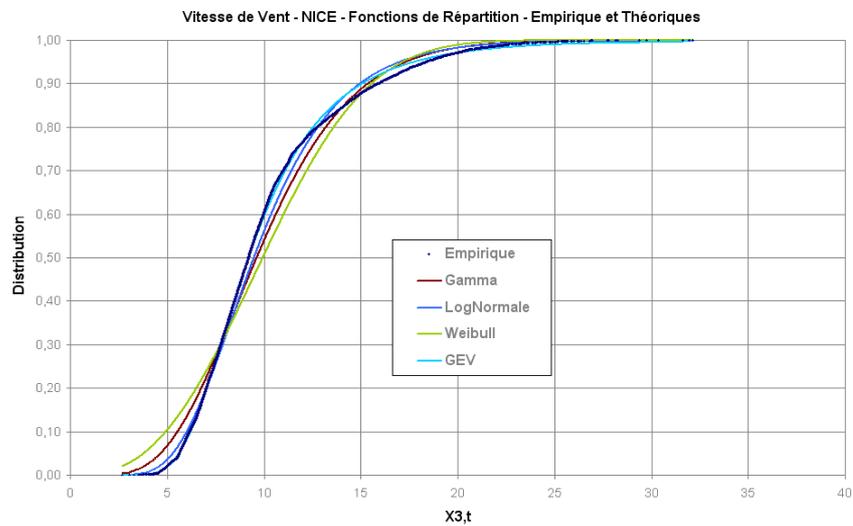
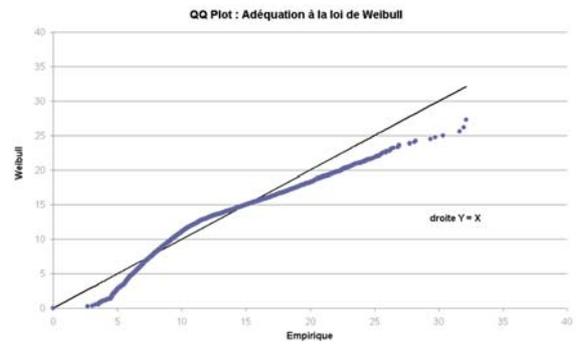
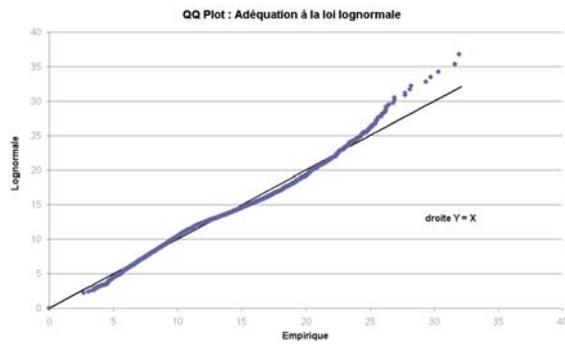
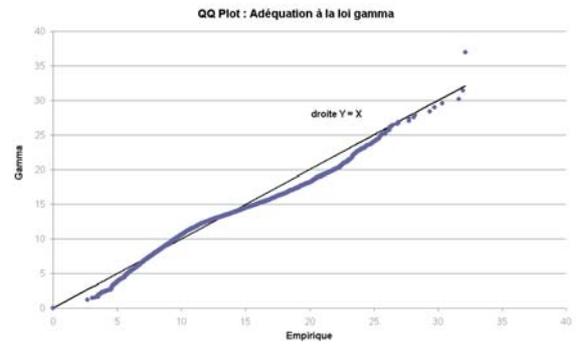
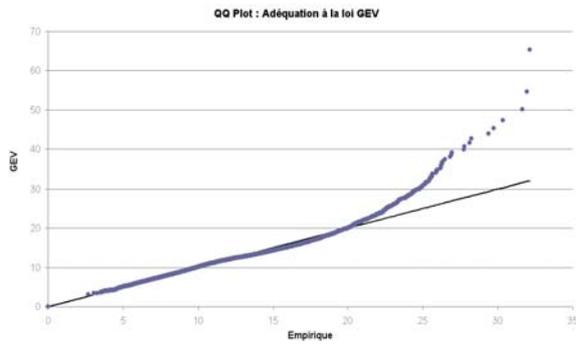
Nous retenons les ajustements par les lois lognormale et GEV plutôt que celui de la loi gamma, en raison des estimations plus conservatrices que devrait entraîner la plus grande épaisseur de leur queue de distribution.

- **Station de Nice – Ajustements des $X_{3,t}$**

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Gamma	$\hat{a} = 6,687$ $\hat{b} = 1,51$	0,076	131,82
Lognormale	$\hat{\mu} = 2,250$ $\hat{\sigma} = 0,354$	0,051	53,31
Weibull	$\hat{a} = 11,39$ $\hat{m} = 2,684$	0,112	308,99
GEV	$\hat{\mu} = 8,237$ $\hat{\psi} = 2,578$ $\hat{\xi} = 0,14$	0,022	9,205

Les graphiques suivants permettent d'observer la qualité des ajustements :



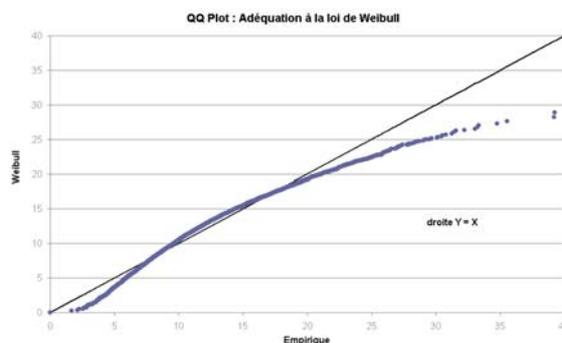
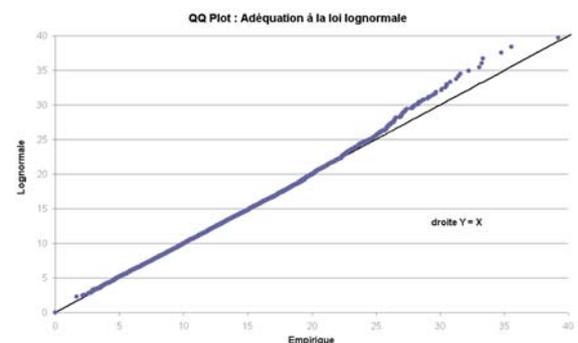
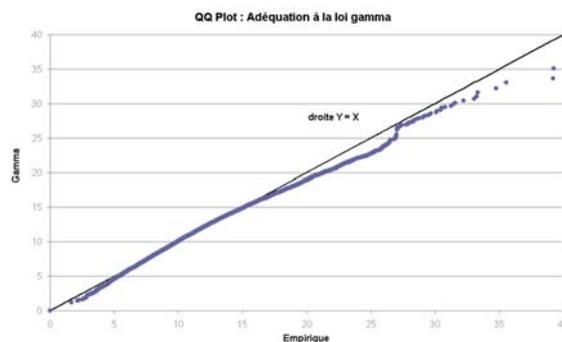
Compte tenu des résultats des tests, nous retenons uniquement l'ajustement par une loi lognormale dans un premier temps.

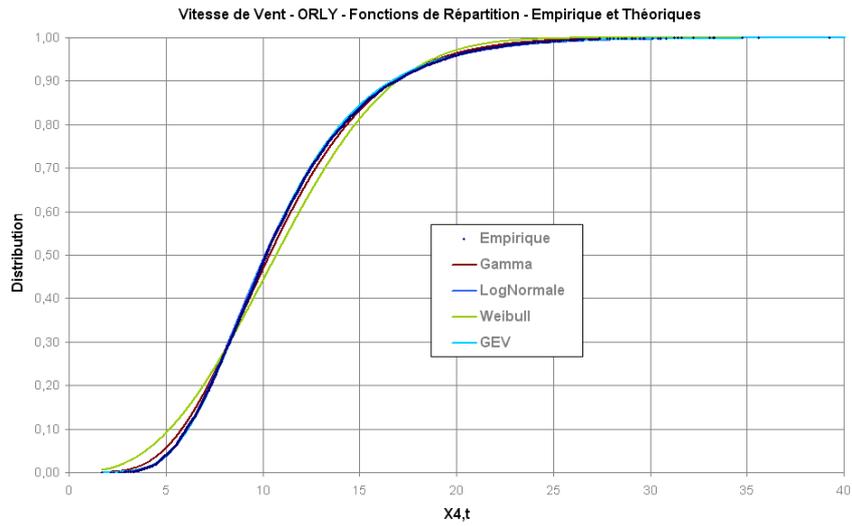
- Station de Orly – Ajustements des $X_{4,t}$

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Gamma	$\hat{a} = 6,226$ $\hat{b} = 1,76$	0,023	15,68
Lognormale	$\hat{\mu} = 2,315$ $\hat{\sigma} = 0,394$	0,010	1,25
Weibull	$\hat{a} = 12,31$ $\hat{m} = 2,615$	0,059	115,43
GEV	$\hat{\mu} = 8,936$ $\hat{\psi} = 3,361$ $\hat{\xi} = 0,016$	0,008	1,504

Les QQ Plots suivants permettent d'observer la qualité des ajustements :





Compte tenu du résultat des tests, nous retenons pour cette station les ajustements par les lois lognormale et GEV.

- **Estimation de la prime de la couverture**

Dans un premier temps, nous modélisons, en utilisant la méthode de Monte Carlo, les valeurs des $X_{i,t}$ de chaque station en utilisant les lois lognormale. Afin de déterminer les estimations des indices annuels $S_i(T)$, nous générons également dans notre modèle une fréquence annuelle moyenne de 365,25 valeurs $X_{i,t}$. La valeur de cette fréquence est réalisée par un tirage sur (365 ; 366) avec $p_{365} = 75 \%$ et $p_{366} = 25 \%$.

En effectuant pour chacune des 100 000 simulations réalisées le calcul $S_i(T) = \sum_{t=1}^{t=365 \text{ ou } 366} \max(0; X_{i,t} - K_i)$, nous obtenons les résultats indiqués dans le tableau suivant :

	Bordeaux S₁(T)	Metz S₂(T)	Nice S₃(T)	Orly S₄(T)
Rappel moyenne historique	704	589	93	491
Rappel écart type historique	1 378	2 461	263	1 685
Moyenne (dam/h)	855	1141	511	862
Ecart type (dam/h)	1363	1650	1034	1407

Hormis pour la station de Bordeaux, les ajustements effectués conduisent à des estimations des $S_i(T)$ assez éloignées des observations historiques. Les ajustements réalisés des vitesses de vent au premier dam/h par des distributions lognormale ne permettent pas une modélisation correcte des valeurs extrêmes de vitesse de vent.

Nous poursuivons tout de même, et effectuons pour chaque simulation le calcul de $C(T)$:

$$C(T) = 1000 \times \max \left[\min \left(4000, \sum_{i=1}^{i=4} p_i \times S_i(T) - 1000 \right), 0 \right]$$

Soient \bar{C} et s les estimations respectives de la moyenne et de l'écart type de l'indemnisation annuelle de la couverture. En appliquant des chargements de sécurité sur écart type de 20 % et 40 %, nous obtenons les estimations respectives $\Pi_{20\%}$ et $\Pi_{40\%}$ de la prime commerciale. Les résultats obtenus sont décrits dans le tableau suivant :

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Modélisation lois lognormale	216 249	472 244	310 698	405 175

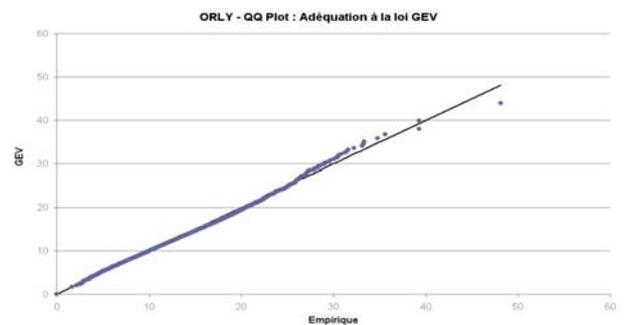
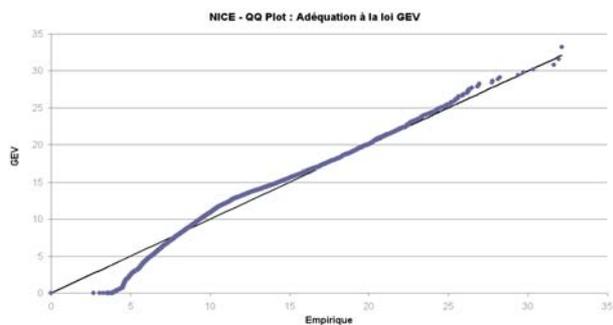
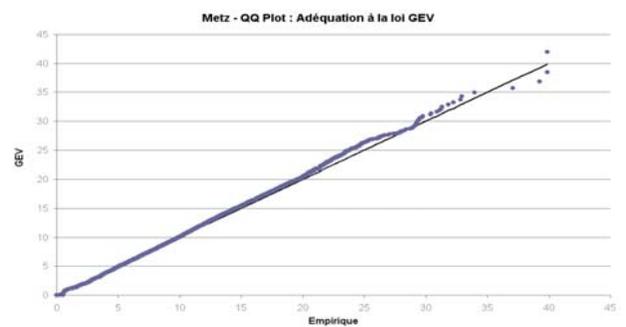
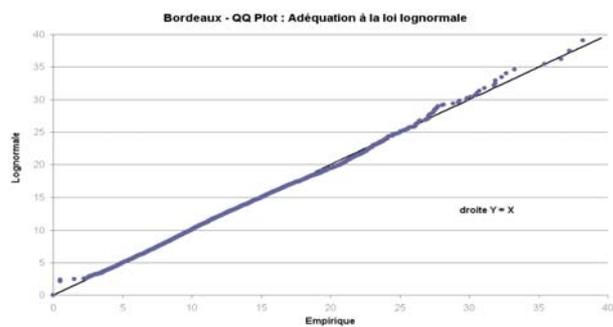
Bien que les estimations du prix de la couverture ne soient que légèrement inférieures à celles basées sur la moyenne historique, les modélisations des indices individuels par station tendent globalement à surestimer les valeurs des indices $S_j(T)$ par rapport aux observations historiques.

Les ajustements réalisés par les lois GEV sont globalement satisfaisants, mais ils présentent également des problèmes en ce qui concerne les queues de distribution pour les stations de Nice et Orly, conduisant après simulations à des surestimations des indices $S_j(T)$. En décidant de relâcher l'objectif du meilleur ajustement sur l'ensemble des observations de $X_{i,t}$, nous avons utilisé la flexibilité des paramètres des lois GEV afin d'obtenir de meilleurs ajustements sur les queues de distribution. Les paramètres ψ et ξ , respectivement appelés « scale » et « shape » en anglais, ont une forte influence sur l'allure de la distribution. Après des séries d'essais sur les valeurs de ces paramètres, nous avons retenu les paramétrages qui offraient, empiriquement et visuellement sur les QQ plots, de meilleurs ajustements pour les queues de distribution.

Nous construisons un nouveau modèle en conservant la loi lognormale pour la station de Bordeaux, et en utilisant trois nouveaux paramétrages pour les stations de Metz, Nice et Orly :

	Bordeaux	Metz	Nice	Orly
Loi utilisée	Lognormale	GEV	GEV	GEV
Paramètres	$\hat{\mu} = 2,243$	$\hat{\mu} = 7,91$	$\hat{\mu} = 8,24$	$\hat{\mu} = 8,94$
	$\hat{\sigma} = 0,389$	$\hat{\psi} = 3,61$	$\hat{\psi} = 4,00$	$\hat{\psi} = 3,20$
		$\hat{\xi} = -0,014$	$\hat{\xi} = -0,104$	$\hat{\xi} = 0,016$

Les Q-Q Plots suivants permettent de visualiser ces ajustements, notamment sur les queues de distribution :



La méthode est certes moins rigoureuse, et conduit à des valeurs plus élevées des statistiques des tests e.d.f., mais elle offre une meilleure estimation des valeurs extrêmes.

Après 100 000 simulations, nous obtenons globalement des estimations de la moyenne et de la volatilité des indices $S_j(T)$ plus conformes à la réalité historique :

	Bordeaux $S_1(T)$	Metz $S_2(T)$	Nice $S_3(T)$	Orly $S_4(T)$
Rappel moyenne historique	704	589	93	491
Rappel écart type historique	1 378	2 461	263	1 685
Moyenne (dam/h)	855	716	94	601
Ecart type (dam/h)	1363	1137	302	1137

Toutefois, nous obtenons par la suite des estimations du prix de la couverture encore plus éloignées des moyennes historiques :

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Modélisation finale	103 081	317 307	166 542	230 004

Cette première approche de modélisation des vitesses de vent au premier m/s permet donc de reproduire assez fidèlement, bien que difficilement, le comportement des indices de chaque station, mais ne fournit pas dans notre exemple une estimation du prix de la couverture qui nous paraisse acceptable.

4.3.3 Ajustements avec introduction d'un seuil

Afin de pallier les problèmes d'estimation des queues de distribution que nous avons rencontrés lors des ajustements au premier m/s, nous décidons d'introduire des seuils dans les ajustements.

Nous décidons donc de modéliser uniquement les vitesses de vent supérieures à un seuil z_i pour chaque station i .

Les échantillons historiques de vitesses de vent ont donc une taille plus réduite, et nous notons $X'_{i,t}$ les valeurs des vitesses de vent de la station i supérieures au seuil z_i .

Nous avons pu voir dans la troisième partie que des vitesses de vent inférieures à 90 km/h (25 m/s) avaient une probabilité presque nulle d'engendrer des sinistres pour un portefeuille d'assurance. C'est pourquoi les couvertures indicielles vent utilisent généralement des franchises par station supérieures ou égales à 100 km/h. Les seuils z_i seront donc choisis en dessous des valeurs des franchises K_i retenues dans notre exemple, mais également en dessous de 100 km/h, ce qui s'avère par ailleurs nécessaire si l'on souhaite garder des échantillons de taille suffisante pour les ajustements. Enfin, notre démarche consistera également à choisir des seuils suffisamment élevés afin de trouver les meilleurs ajustements possibles des queues de distribution.

Nous modéliserons les $X'_{i,t}$ pour chaque station en utilisant les lois exponentielle, Pareto, lognormale, Weibull, et Pareto Généralisée. Les ajustements par des lois GEV ne se sont pas avérés concluants (l'estimation des paramètres est compliquée et ne nous a pas conduit à des ajustements satisfaisants).

Parallèlement, nous effectuerons une modélisation du nombre annuel de jours N_i où les vitesses de vent sont supérieures au seuil z_i pour la station i . Nous utiliserons la loi de Poisson et la loi binomiale négative, et mettrons en œuvre le test d'adéquation du Chi-Deux.

Soient $N_{i,j}$ avec $j = (1, \dots, 33)$ les $n = 33$ observations historiques de N_i sur la période 1970 – 2002. Nous notons \bar{N}_i et s^2 les valeurs respectivement de la moyenne et de la variance de l'échantillon des $N_{i,j}$.

❖ *Loi de Poisson de paramètre λ :*

$$P(N_i=k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

Estimation de λ :

$$\hat{\lambda} = \bar{N}_i = \frac{\sum_{j=1}^{33} N_{i,j}}{n}$$

❖ *Loi binomiale négative de paramètres a et p :*

$$P(N_i=k) = \frac{\Gamma(a+k)}{\Gamma(a) \cdot k!} \cdot p^a \cdot (1-p)^k$$

Estimation de a :

$$\hat{a} = \frac{\bar{N}_i^2}{s^2 - \bar{N}_i}$$

Estimation de p :

$$\hat{p} = \frac{\bar{N}_i}{s^2}$$

❖ *Simulation des lois de fréquence :*

Afin de simuler la loi de Poisson et la loi Binomiale Négative, nous utiliserons la technique assez classique qui consiste à :

- ▶ simuler un nombre aléatoire r_i distribué uniformément sur $[0, 1]$,
- ▶ trouver le plus petit k tel que $F(k) > r_i$, où F est la fonction de répartition de la loi de fréquence simulée,
- ▶ retenir la valeur k , qui est donc un nombre généré aléatoirement selon la loi souhaitée.

❖ *Test du Chi-Deux*

Ce test sera utilisé pour mesurer la qualité des ajustements et choisir la loi la plus appropriée.

Afin de le mettre en œuvre, nous réaliserons des partitions des échantillons $N_{i,j}$ en C classes. La valeur de C sera choisie telle que :

- la fréquence théorique de chaque classe c soit supérieure à 1
- 80 % des fréquences théoriques de classes soient supérieures à 5
- $n^{1/3} \leq C \leq n^{2/3}$

Soient f_c le nombre d'observations de la classe c et F_c le nombre d'observations attendu de la classe c selon la loi étudiée.

La statistique du test est :

$$\chi^2_{obs} = \sum_{c=1}^C \frac{(f_c - F_c)^2}{F_c}$$

Nous utiliserons la p-value associée à la loi étudiée, et qui est définie par $P(\chi^2 > \chi^2_{obs})$, où χ^2 est une distribution du Chi-deux à $(C - 1 - \text{nombre de paramètres estimés})$ degrés de liberté.

- **Station de Bordeaux**

Nous choisissons pour la station de Bordeaux un seuil de modélisation $z_1 = 24,5$ m/s.

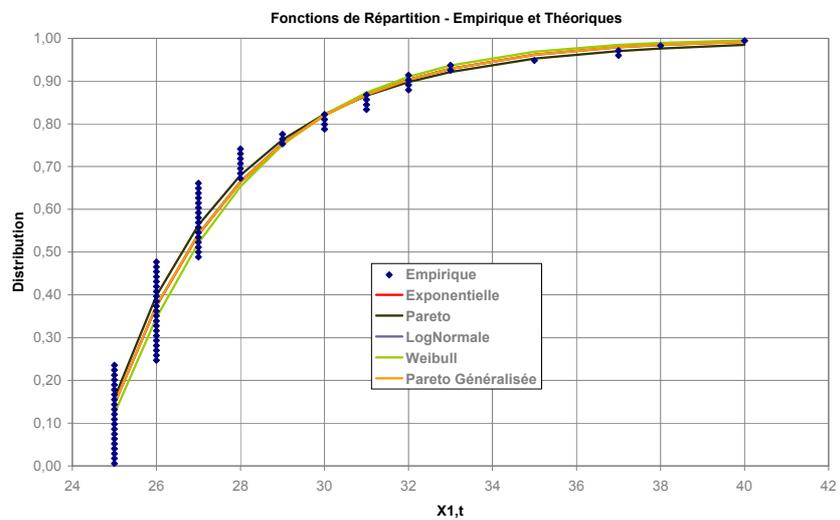
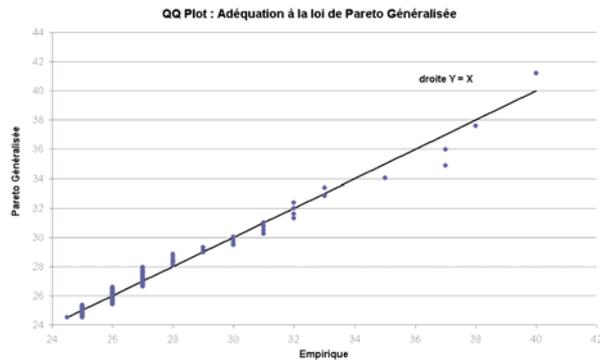
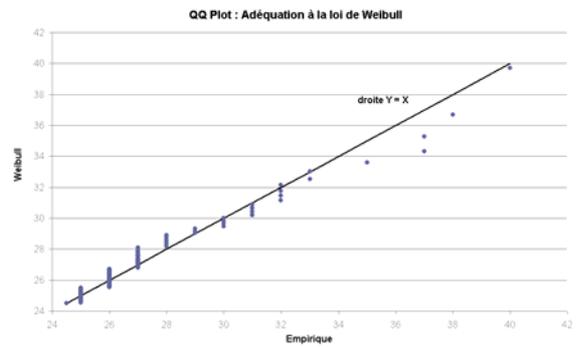
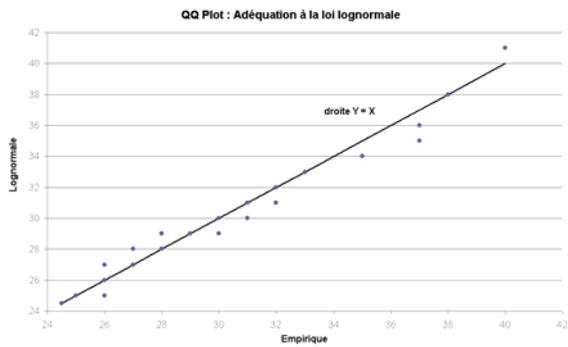
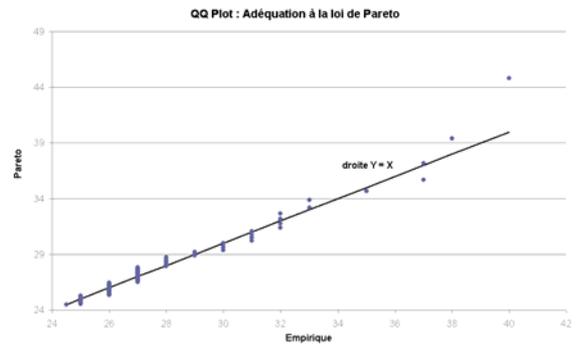
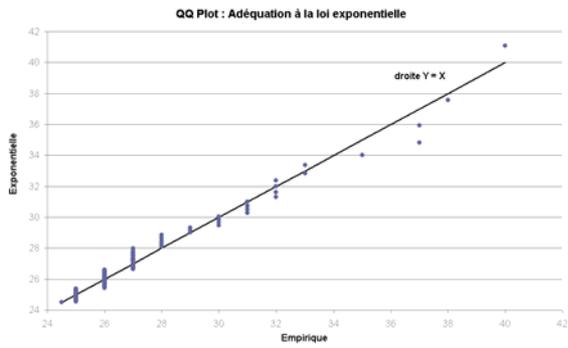
Ajustements des valeurs $X'_{1,t}$:

La taille de l'échantillon des valeurs $X'_{1,t}$ est de 87.

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Exponentielle	$\hat{a} = 3,11E-01$	0,144	2,06
Pareto	$\hat{a} = 8,535$	0,158	2,19
Lognormale	$\hat{\mu} = 2,457$ $\hat{\sigma} = 0,335$	0,145	2,05
Weibull	$\hat{a} = 3,32$ $\hat{m} = 1,077$	0,145	2,15
Pareto Généralisée	$\hat{a} = 3,196$ $\hat{c} = 0,00515$	0,145	2,06

Les graphiques suivants permettent d'observer la qualité des ajustements :

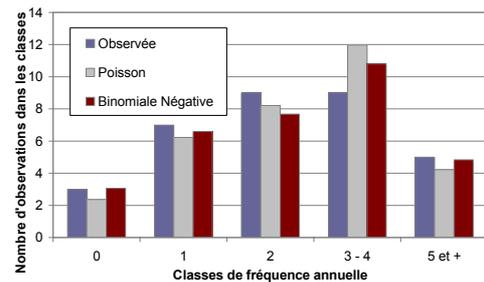


Ajustements de la fréquence N_1 :

Nous obtenons les ajustements suivants :

Loi étudiée	Paramètres	P-Value Test du Chi-deux
Poisson	$\hat{\lambda} = 2,64$	74,7 %
binomiale négative	$\hat{a} = 11,54$ $\hat{p} = 0,814$	75,6 %

Classes de fréquence annuelle	Observée	Poisson	bin neg
0	3	2,36	3,07
1	7	6,23	6,59
2	9	8,21	7,69
3 - 4	9	11,98	10,81
5 et +	5	4,22	4,84
valeur du χ^2_{obs} :		1,23	0,56



Nous retiendrons la loi binomiale négative (p-value du test du Chi-deux légèrement plus élevée).

- **Station de Metz**

Nous choisissons pour la station de Metz un seuil de modélisation $z_2 = 23,5$ m/s.

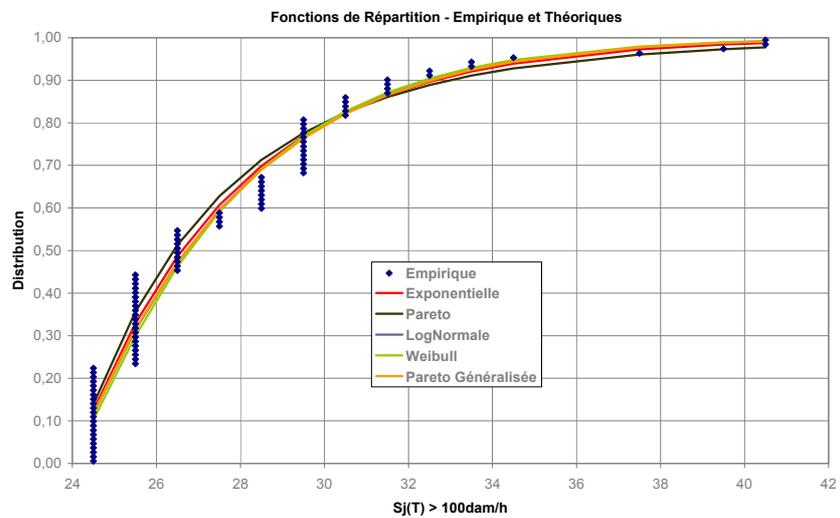
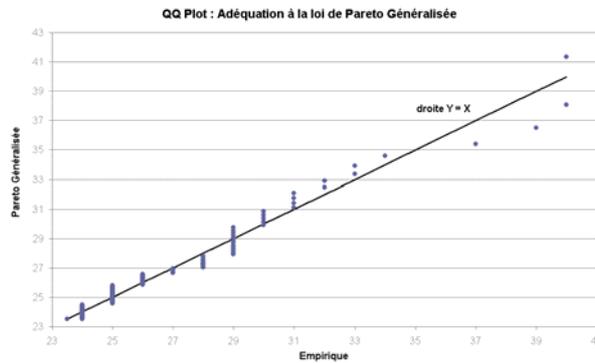
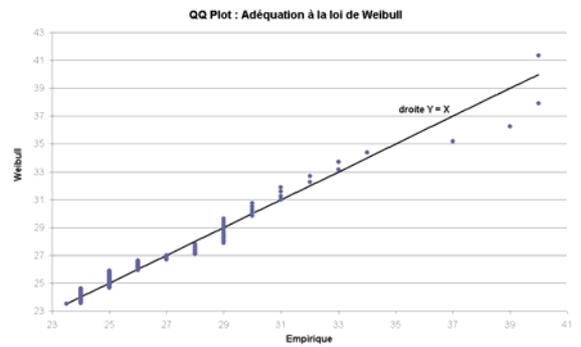
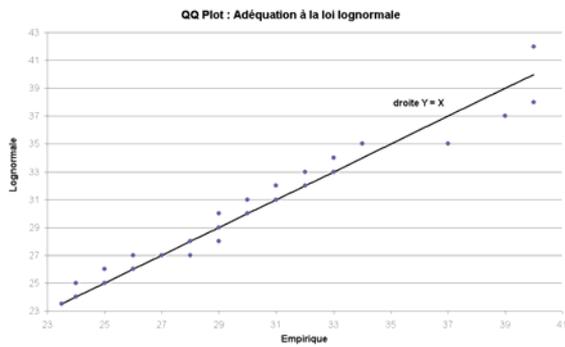
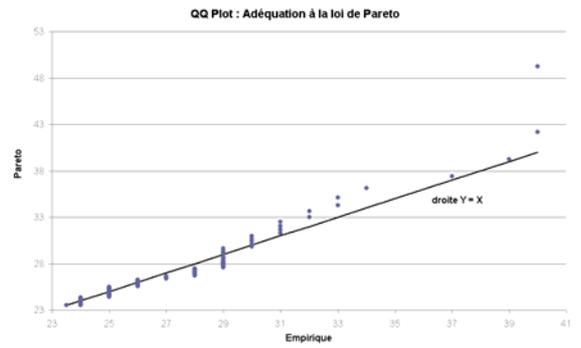
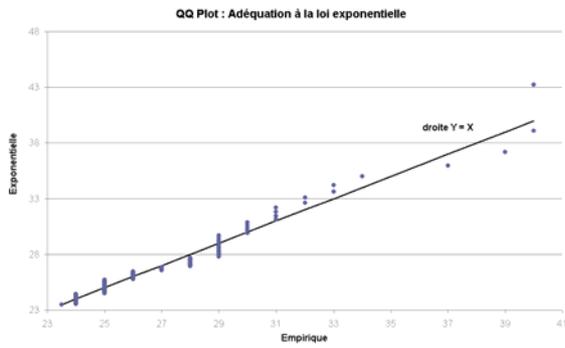
Ajustements des $X'_{2,t}$

La taille de l'échantillon des valeurs $X'_{2,t}$ est de 96.

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Exponentielle	$\hat{a} = 2,67E-01$	0,125	1,95
Pareto	$\hat{a} = 7,120$	0,139	2,24
Lognormale	$\hat{\mu} = 2,857$ $\hat{\sigma} = 0,276$	0,133	1,95
Weibull	$\hat{a} = 3,88$ $\hat{m} = 1,087$	0,148	2,09
Pareto Généralisée	$\hat{a} = 3,991$ $\hat{c} = -0,0641$	0,131	1,93

Les QQ Plots suivants permettent d'observer la qualité des ajustements :

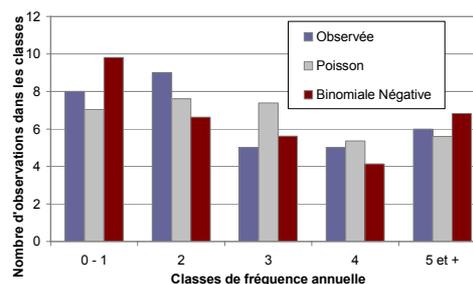


Ajustements de la fréquence N_2 :

Nous obtenons les ajustements suivants :

Loi étudiée	Paramètres	P-Value Test du Chi-deux
Poisson	$\hat{\lambda} = 2,91$	75,1 %
binomiale négative	$\hat{a} = 4,26$ $\hat{p} = 0,594$	46,6 %

Classes de fréquence annuelle	Observée	Poisson	bin neg
0 - 1	8	7,03	9,81
2	9	7,61	6,63
3	5	7,38	5,61
4	5	5,37	4,13
5 et +	6	5,60	6,82
valeur du χ^2_{obs} :		1,21	1,53



Nous retiendrons la loi de Poisson (p-value du test du Chi-deux plus élevée).

- **Station de Nice**

Nous choisissons pour la station de Nice un seuil de modélisation $z_3 = 23,5$ m/s.

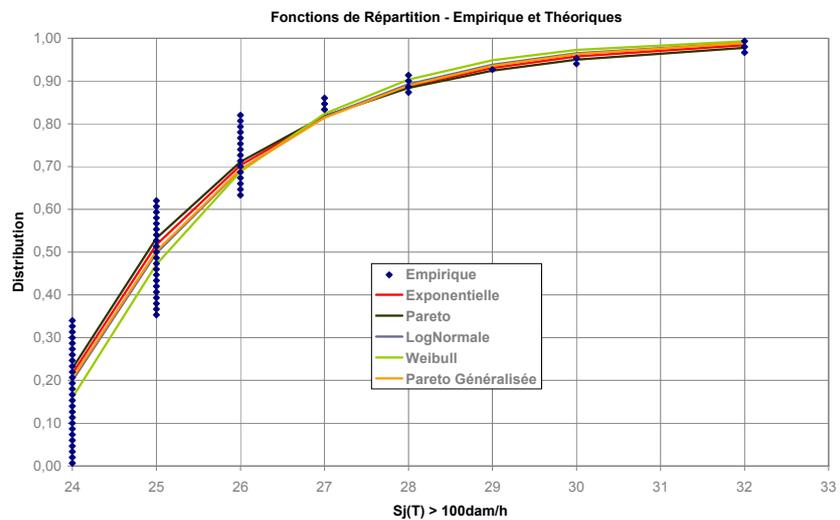
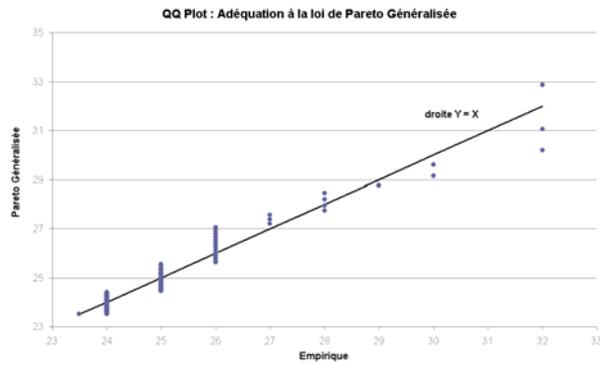
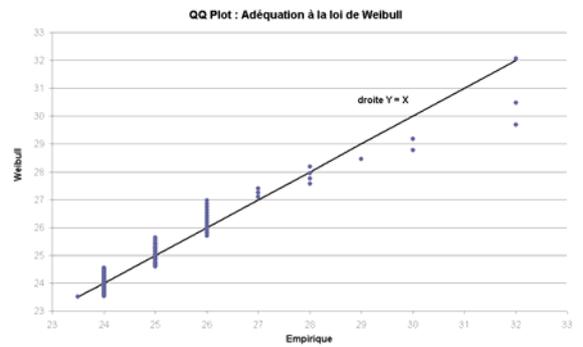
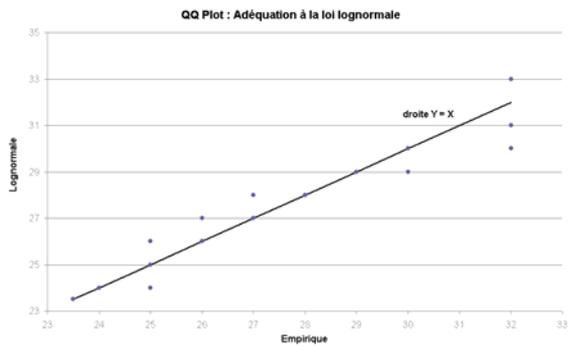
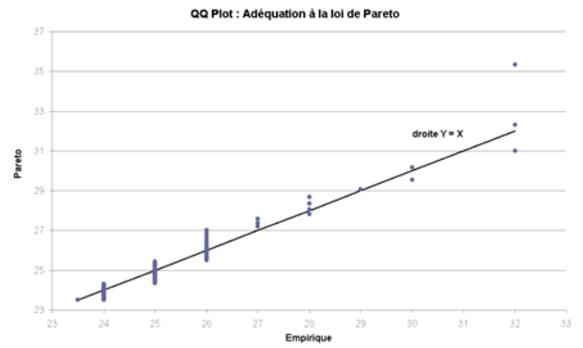
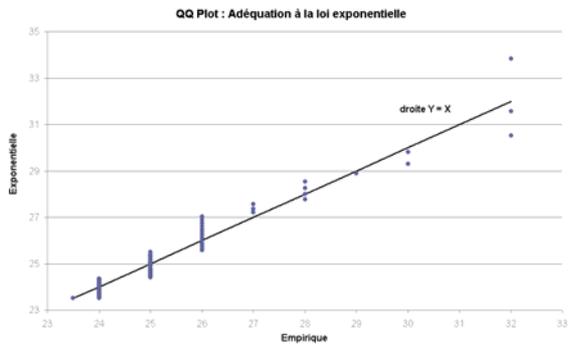
Ajustements des $X'_{3,t}$

La taille de l'échantillon des valeurs $X'_{3,t}$ est de 96.

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Exponentielle	$\hat{a} = 4,85E-01$	0,216	3,34
Pareto	$\hat{a} = 12,284$	0,228	3,59
Lognormale	$\hat{\mu} = 2,921$ $\hat{\sigma} = 0,176$	0,199	3,17
Weibull	$\hat{a} = 2,20$ $\hat{m} = 1,183$	0,187	3,18
Pareto Généralisée	$\hat{a} = 2,20$ $\hat{c} = -0,0680$	0,205	3,22

Les QQ Plots suivants permettent d'observer la qualité des ajustements :

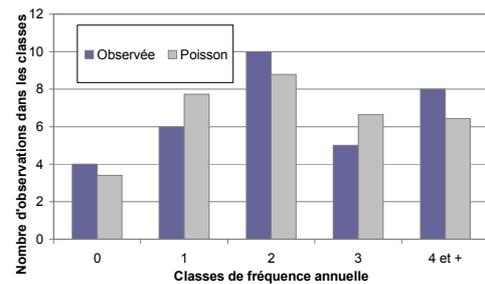


Ajustements de la fréquence N_3 :

Nous obtenons les ajustements suivants :

Loi étudiée	Paramètres	P-Value Test du Chi-deux
Poisson	$\hat{\lambda} = 2,27$	69,4 %
binomiale négative	inapplicable (variance empirique inférieure à la moyenne)	

Classes de fréquence annuelle	Observée	Poisson
0	4	3,40
1	6	7,73
2	10	8,78
3	5	6,65
4 et +	8	6,44
valeur du χ^2_{obs} :		1,45



Nous retiendrons la loi de Poisson.

- **Station de Orly**

Nous choisissons pour la station d'Orly un seuil de modélisation $z_4 = 25,5$ m/s.

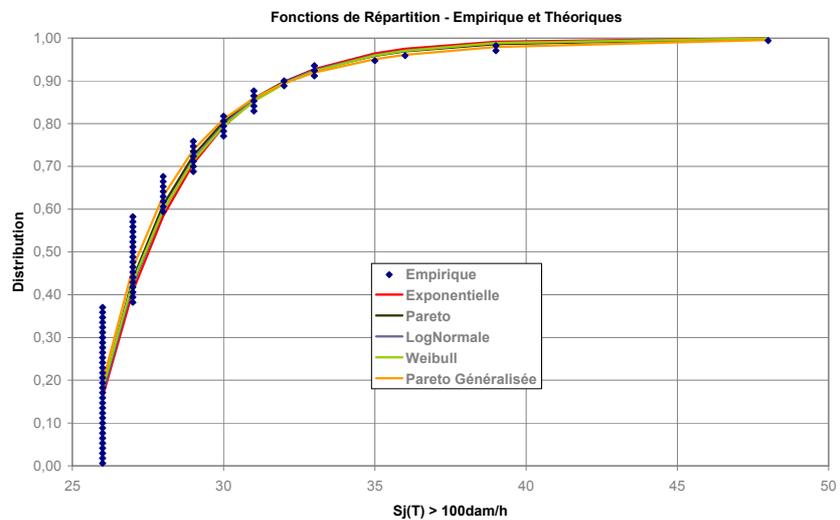
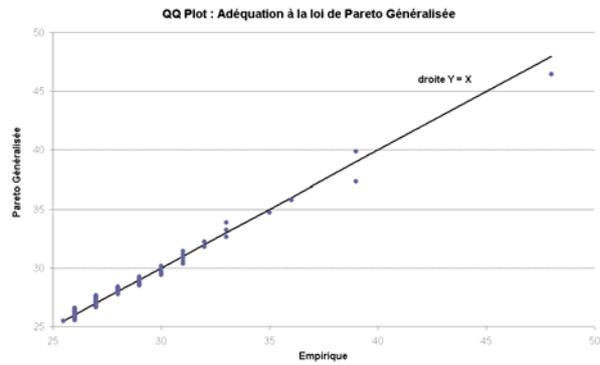
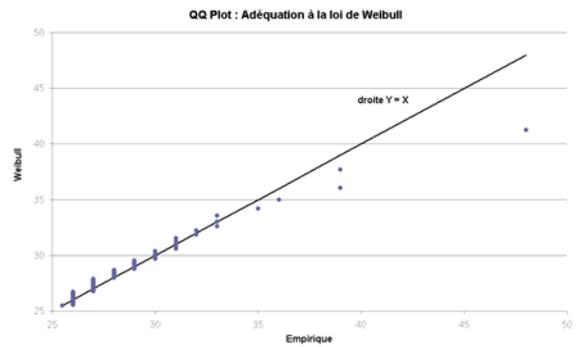
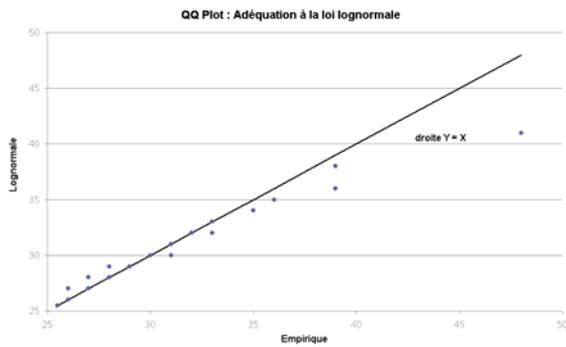
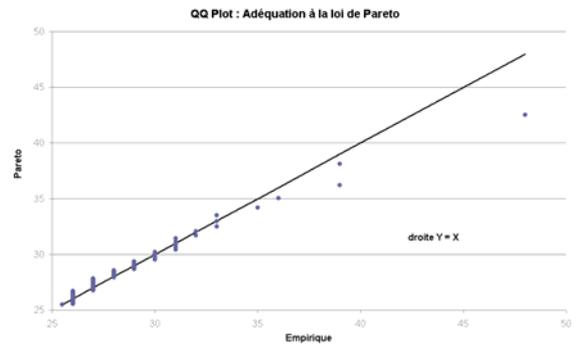
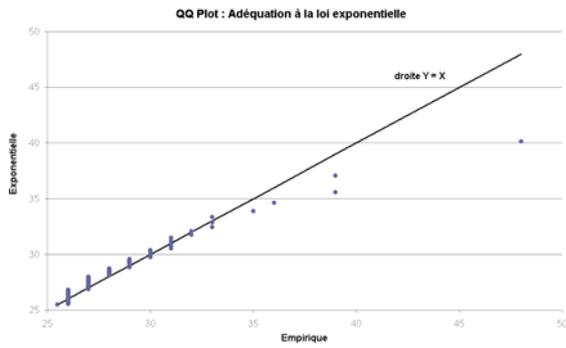
Ajustements des $X'_{4,t}$

La taille de l'échantillon des valeurs $X'_{4,t}$ est de 85.

Le tableau suivant synthétise les caractéristiques des ajustements effectués :

Loi étudiée	Paramètres	Kolmogorov-Smirnov	Anderson-Darling
Exponentielle	$\hat{a} = 3,51E-01$	0,216	3,88
Pareto	$\hat{a} = 10,048$	0,199	3,37
Lognormale	$\hat{\mu} = 1,302$ $\hat{\sigma} = 0,463$	0,207	3,58
Weibull	$\hat{a} = 2,76$ $\hat{m} = 0,941$	0,195	3,44
Pareto Généralisée	$\hat{a} = 2,256$ $\hat{c} = 0,2122$	0,195	3,20

Les QQ Plots suivants permettent d'observer la qualité des ajustements :

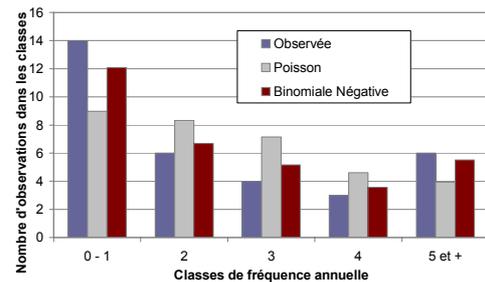


Ajustements de la fréquence N_4 :

Nous obtenons les ajustements suivants :

Loi étudiée	Paramètres	P-Value Test du Chi-deux
Poisson	$\hat{\lambda} = 2,58$	9,0 %
binomiale négative	$\hat{a} = 3,23$ $\hat{p} = 0,557$	68,0 %

Classes de fréquence annuelle	Observée	Poisson	bin neg
0 - 1	14	8,98	12,08
2	6	8,33	6,68
3	4	7,15	5,17
4	3	4,61	3,57
5 et +	6	3,93	5,51
valeur du χ^2_{obs} :		6,49	0,77



Nous retiendrons la loi binomiale négative (p-value du test du Chi-deux plus élevée).

- **Estimation de la prime de la couverture**

Nous modélisons donc en utilisant la méthode de Monte Carlo les fréquences annuelles N_i de jours où les vitesses de vent sont supérieures aux seuils z_i pour chaque station, en utilisant les lois de fréquence retenues.

En fonction du nombre de sinistres généré, nous simulons ensuite les valeurs des $X'_{i,t}$ de chaque station en utilisant les lois de Pareto Généralisée, qui nous sont globalement apparues comme les plus appropriées au vu des tests réalisés.

En effectuant pour chacune des 100 000 simulations réalisées le calcul $S_i(T) = \sum_{t=1}^{t=N_i} \max(0; X'_{i,t} - K_i)$, nous obtenons les résultats indiqués dans le tableau suivant :

	Bordeaux S₁(T)	Metz S₂(T)	Nice S₃(T)	Orly S₄(T)
Rappel moyenne historique	704	589	93	491
Rappel écart type historique	1 378	2 461	263	1 685
Moyenne (dam/h)	646	630	67	505
Ecart type (dam/h)	1243	1198	276	1526

Les estimations obtenues sont assez satisfaisantes. On note toutefois que la moyenne estimée de l'indice de Nice est inférieure à la moyenne historique, et que la volatilité historique de l'indice de Metz est nettement supérieure à celle que nous modélisons.

Afin d'estimer le prix de la couverture, nous effectuons le calcul de $C(T)$ pour chaque simulation :

$$C(T) = 1000 \times \max \left[\min \left(4000, \sum_{i=1}^{i=4} p_i \times S_i(T) - 1000 \right), 0 \right]$$

Nous obtenons ainsi les estimations suivantes du prix de la couverture, qui à nouveau apparaissent assez sensiblement en deçà des moyennes historiques :

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Modélisation lois de Pareto généralisée	106 430	403 008	187 032	267 633

L'approche de modélisation des vitesses de vent excédant un seuil permet donc d'appréhender plutôt correctement, et plus aisément qu'avec la modélisation au premier m/s, le comportement des indices de chaque station. Cependant, elle n'offre pas réellement dans notre exemple une meilleure estimation du prix de la couverture elle-même.

4.3.4 Conclusions

L'estimation du prix de la couverture indiciaire étudiée en exemple à partir de modélisations basées sur les vitesses de vent est problématique. En effet, les ajustements effectués, avec ou sans seuil, permettent de modéliser, de façon acceptable d'un point de vue statistique, le comportement des indices annuels par station $S_i(T)$. Ces ajustements offrent des estimations des moments d'ordre 1 et 2 en ligne avec les observations empiriques sur la période complète 1970-2002. Mais les modélisations de l'indice annuel de la couverture $S(T)$ et de l'indemnisation associée $C(T)$ semblent fortement sous-estimées compte tenu des valeurs historiques, qui sont pour nous des valeurs de référence. La modélisation des indices par station pris indépendamment les uns des autres ne semble donc pas adaptée à l'objectif de tarification des couvertures indiciaires, car elle n'appréhende pas correctement les liens qui doivent exister entre les indices individuels par station.

Deux remarques secondaires sont également à formuler concernant les approches qui ont été retenues en matière de modélisation des vitesses de vent. Tout d'abord, les analyses ont été effectuées comme si les variables aléatoires de vitesses de vent étaient réelles. Nous avons ainsi simulé des vitesses de vent à valeurs dans \mathbb{R} et non dans \mathbb{N} , en considérant que l'impact serait négligeable. Par ailleurs, il aurait été préférable lors des ajustements probabilistes de définir des bornes supérieures pour les valeurs de vitesses de vent. Toutefois, le choix de ces bornes aurait pu être discutable, et nous avons également jugé négligeable l'impact potentiel de limitation par le haut des vitesses de vent.

4.4 Autres approches envisageables

Afin de tarifier les couvertures indiciaires en général, et notre exemple de couverture indiciaire vent en particulier, il paraît intéressant d'étudier les approches traditionnellement adoptées dans le cadre de la tarification des options et autres instruments dérivés négociés sur les marchés financiers.

Le modèle de Black & Scholes a par exemple été utilisé pour la tarification des dérivés climatiques de température de type HDD-CDD. On pourrait imaginer l'appliquer à l'indice quotidien $S(t)$ afin de formuler le prix $C(t)$ d'une couverture indiciaire, semblable à un call de maturité T . Toutefois le modèle s'avère inapplicable. Tout d'abord, les variables climatiques ne sont pas des sous-jacents négociables et n'ont pas de prix en tant que telles. Enfin, il est conceptuellement impossible de constituer par couverture dynamique un portefeuille risque-neutre en combinant l'achat de l'option et du sous-jacent.

Les méthodes avancées de tarification des options exotiques, et la vision d'un indice climatique comme un processus stochastique, un processus de Levy par exemple, constituent cependant des approches intéressantes. Nous ne pourrions toutefois développer plus avant ces approches dans le cadre de ce mémoire.

4.5 Conclusions

Nous avons présenté au cours de cette partie plusieurs approches de tarification des couvertures indicielles. En exploitant les séries historiques de valeurs d'indices ou de valeurs de vitesses de vent, nous avons cherché à établir les ajustements probabilistes les plus adéquats en vue de construire des modèles d'estimation de la prime de la couverture indicielle vent étudiée en exemple. Le tableau suivant synthétise les résultats obtenus :

Approche	\bar{c} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Burning Cost 1970-2002	241 436	876 919	416 820	592 204
Modélisation de l'indice S(T)				
au premier dam/h : Loi gamma	198 129	679 733	334 076	470 022
Modélisation de l'indice S(T)				
avec un seuil de 100 dam/h : Loi de Pareto	217 762	811 985	377 159	539 556
Modélisation des vitesses de				
vent au premier m/s : Modèle final	103 081	317 307	166 542	230 004
Modélisation des vitesses de				
vent avec seuils : Lois de Pareto généralisée	106 430	403 008	187 032	267 633

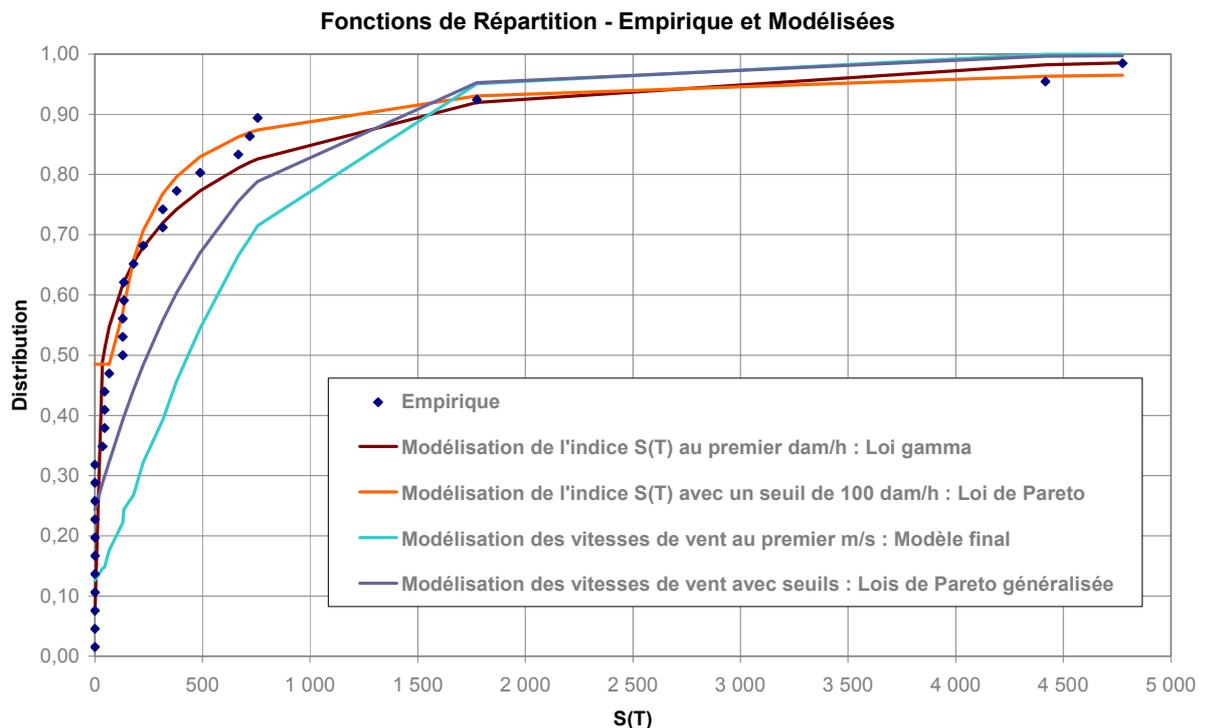
Il est à noter que les estimations des primes commerciales sont mentionnées à titre illustratif, et ne constituent pas en tant que telles le véritable enjeu des travaux de modélisation. En effet, de nombreuses méthodes alternatives de chargement et de calcul de la prime commerciale sont envisageables.

Les approches de modélisation de l'indice lui-même, de même que l'approche Burning Cost, permettent certes d'obtenir des estimations rapides et acceptables de

la couverture indicielle vent étudiée. Elles pourraient toutefois s'avérer inefficaces dans le cadre de la tarification de couvertures différentes, dans la mesure où elles reposent sur l'utilisation d'un échantillon de valeur d'indices de taille très réduite.

C'est pourquoi nous nous sommes attachés dans un deuxième temps à exploiter les séries historiques de données climatiques. Les résultats obtenus sont, comme nous l'avons vu précédemment, assez satisfaisants à l'échelle des indices individuels par station. Bien que l'approche ait l'avantage de pouvoir être généralisée à un ensemble bien plus vaste de couvertures indicielles climatiques, elle comporte malheureusement des lacunes lorsqu'il s'agit d'évaluer le prix de la couverture globale.

Le graphique suivant permet d'illustrer la qualité des approches en comparant la fonction de répartition empirique de $S(T)$ aux fonctions de répartition de $S(T)$ obtenues après simulations des différents ajustements :



Au cours de la cinquième partie, nous allons tenter d'améliorer les tarifications obtenues à partir des ajustements de vitesses de vent, en incorporant dans les modélisations des estimations des dépendances existant entre les quatre stations de la couverture indicielle vent.

5 ANALYSE DE LA DEPENDANCE SPATIALE & TARIFICATION AFFINEE DE LA COUVERTURE

Nous soulignerons tout d'abord l'existence de dépendances entre les valeurs des enregistrements effectués dans un réseau de stations météorologiques, et qui résultent de la nature même des phénomènes climatiques observés. Nous présenterons ensuite les mesures et outils mathématiques liés à la notion de dépendance sur lesquels nous nous appuyerons en vue d'affiner la tarification de la couverture indiciaire vent. Nous prendrons ainsi en compte la dépendance spatiale dans les modélisations de l'indice fondées sur les ajustements de vitesses de vent.

5.1 Mise en évidence du problème de la dépendance spatiale

5.1.1 Les caractéristiques de l'évènement climatique

Les phénomènes météorologiques susceptibles de faire l'objet d'une protection indiciaire concernent bien souvent des zones géographiques étendues. Une couverture indiciaire est généralement conçue à l'échelle d'un territoire national. A titre d'exemple, une vague de grand froid n'est pas localisée en un point géographique précis mais touche de vastes superficies. Lors de la survenance d'un tel phénomène, l'ensemble du réseau de stations choisies pour une éventuelle couverture indiciaire basée sur les observations des températures quotidiennes enregistre des valeurs extrêmes. Les séries de données climatologiques d'un réseau national de stations météorologiques présentent donc des corrélations certaines en ce qui concerne les températures. Selon les caractéristiques du phénomène climatique, on peut concevoir que cette dépendance entre les valeurs mesurées dans les stations, en particulier les valeurs extrêmes, soit plus ou moins aiguë.

Dans le cas d'une tempête moyenne ou forte, plusieurs stations (localisées dans une zone géographique précise⁶⁷) peuvent enregistrer conjointement des valeurs extrêmes de vitesses de vent sur un laps de temps assez réduit. Selon

⁶⁷ L'échelle caractéristique d'une tempête est de l'ordre de 1000 à 2000 km de diamètre, celle du cœur de la dépression étant de l'ordre de 100 à 200 km de diamètre.

Météo France⁶⁸, une tempête est considérée comme forte « lorsqu'au moins 20 % des stations ont relevé un vent maximal instantané quotidien supérieur à 100 km/h au cours de l'un des trois jours J, J+1 ou J+2 ». L'intégralité des stations retenues dans notre exemple ont ainsi enregistrées le 26/12/1999 des vitesses de vent supérieures à leur seuil respectif de modélisation z_i . Trois d'entre elles (i.e. 75 %) ont notamment enregistrées des vitesses de vent supérieures à 100 km/h.

5.1.2 La notion de dépendance

Avant de présenter plus avant les outils de mesure et de modélisation de la dépendance que nous allons appliquer, nous souhaitons introduire quelques commentaires concernant les notions de dépendance et de corrélation.

La définition de la dépendance donnée par le dictionnaire de l'Académie française est la suivante : « *Relation étroite et parfois réciproque, impliquant ou non une subordination, entre des réalités, des idées* ». Une approche plus mathématique de la notion de dépendance nécessite l'introduction des probabilités conditionnelles : la probabilité que B se réalise sachant que A s'est réalisé est :

$$P(A/B) = \frac{P(A \cap B)}{P(A)}$$

A et B sont dits indépendants si $P(A/B) = P(A)$

D'un point de vue littéraire la corrélation peut être définie comme le « *rapport existant entre deux phénomènes qui varient en fonction l'un de l'autre parce qu'il existe un lien de cause à effet entre eux, ou qu'ils comportent des causes communes* ». Par extension on emploie le terme corrélation pour indiquer « *le caractère de deux choses qui varient simultanément* ». Le terme corrélation est donc fréquemment utilisé pour faire référence à toute notion de dépendance. D'un point de vue mathématique, la corrélation n'est pourtant qu'une mesure parmi d'autres de la dépendance stochastique.

⁶⁸ voir « *Evolution du nombre de tempêtes en France sur la période 1950-1999* ». Christine Drevet. Météo France.

5.1.3 La dépendance entre les stations d'une couverture indicielle

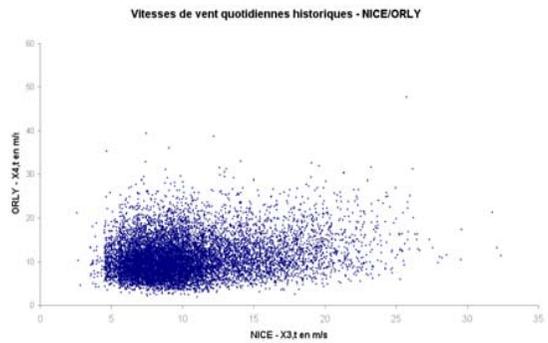
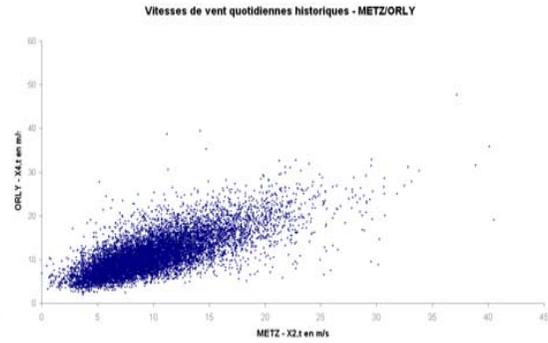
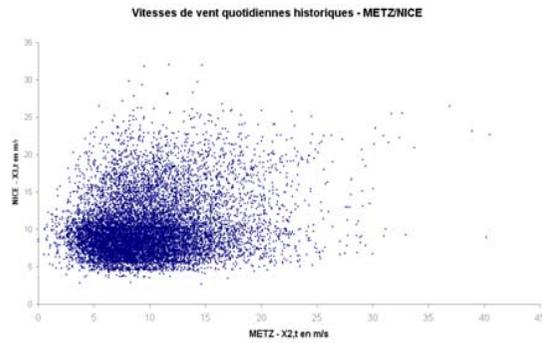
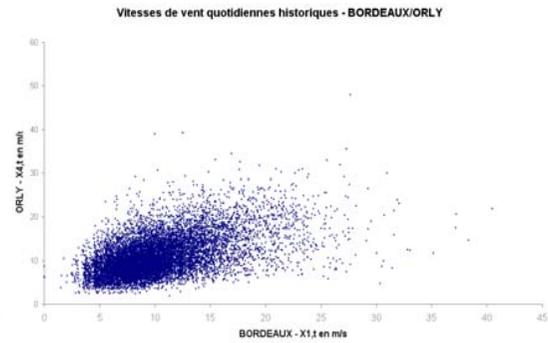
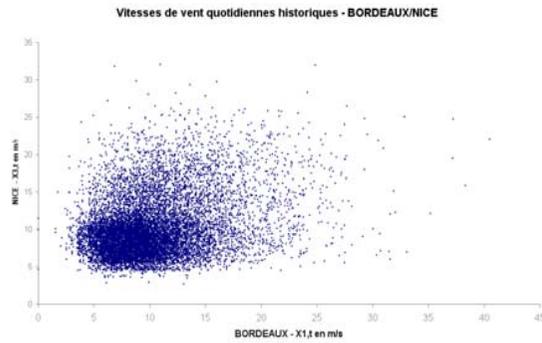
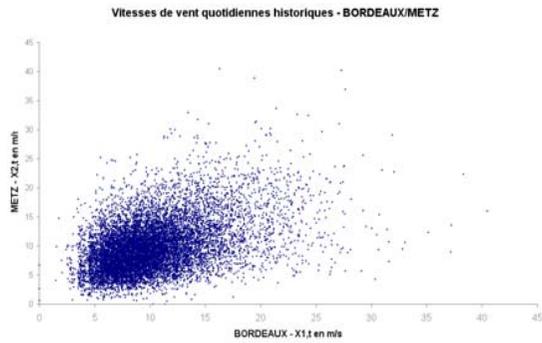
La situation de dépendance entre les observations réalisées dans un panel de stations influence directement le comportement de l'indice d'une couverture ; ceci d'autant plus que les protections indicielles sont généralement fondées sur un réseau relativement dense de stations. L'exemple d'une couverture avec un nombre très limité de stations, quatre, a été choisi afin de faciliter les analyses et n'est pas totalement représentatif des couvertures négociées. Dans la pratique, la proximité géographique des stations est bien plus marquée, afin de diminuer le risque de non corrélation entre valeur de l'indice et montant de sinistralité.

Cette proximité peut être synonyme de fortes corrélations entre les stations. Sur la période 1970-2002, la corrélation linéaire entre les vitesses de vent maximales instantanées quotidiennes enregistrées à Paris Montsouris et celles enregistrées à Orly est ainsi de près de 92 %.

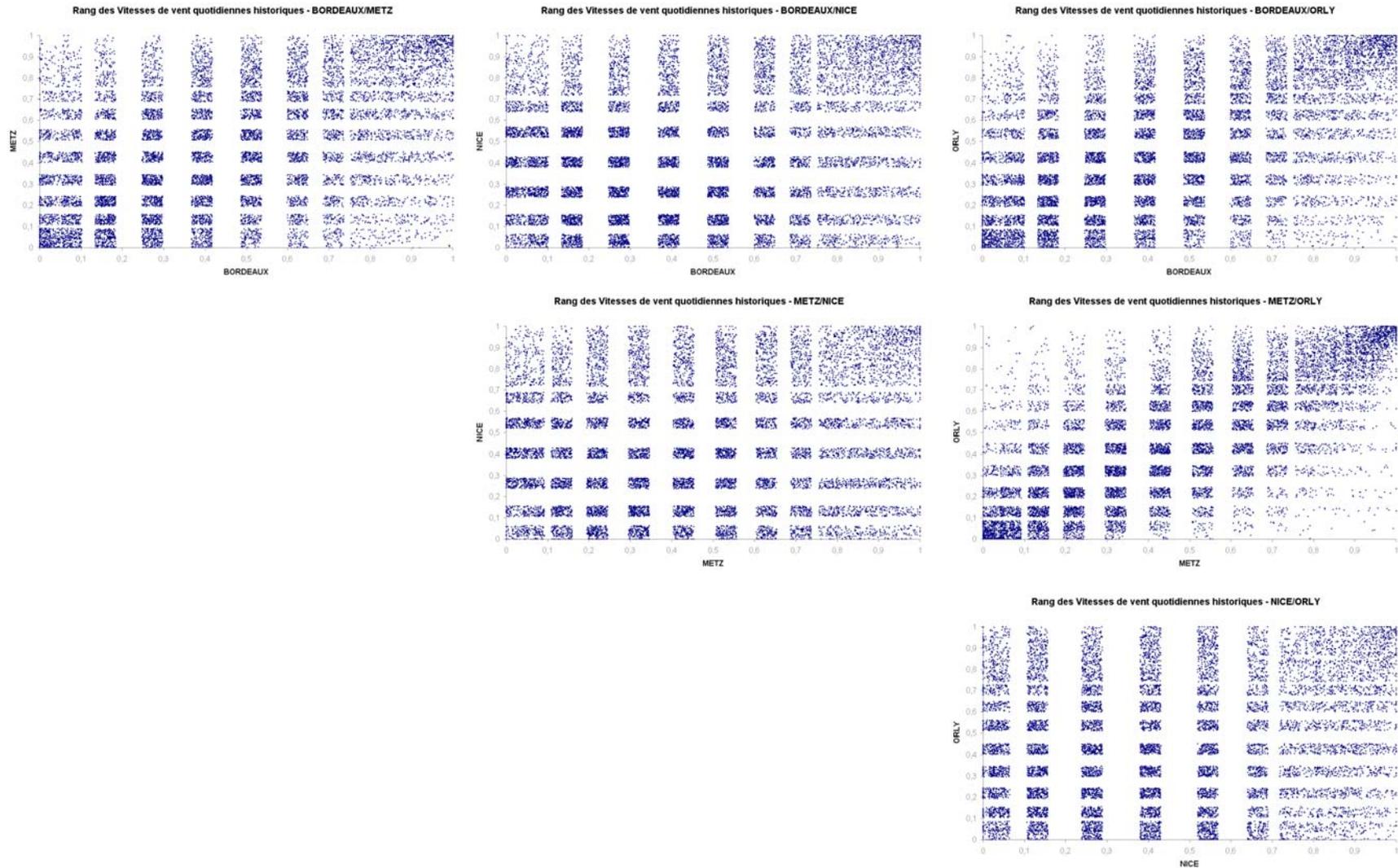
Sur la page suivante, nous avons mis en correspondance sur six graphiques les vitesses de vent historiques des différents couples de stations (i,j) retenues dans notre exemple de couverture indicielle vent. Nous y avons ajouté un bruit blanc afin de transformer les valeurs entières en valeurs réelles, et faciliter ainsi la lecture. On peut constater visuellement sur ces nuages de 12 053 points (nombre de jours entre le 01/01/1970 et le 31/12/2002) l'existence de corrélations plus ou moins fortes, l'exemple le plus frappant étant celui du couple (Metz - Orly).

Sur le même principe, nous présentons sur la page consécutive, les graphiques « quantile-quantile » associés à ces mêmes couples, auxquels nous avons ajouté un bruit blanc afin de faciliter l'interprétation.

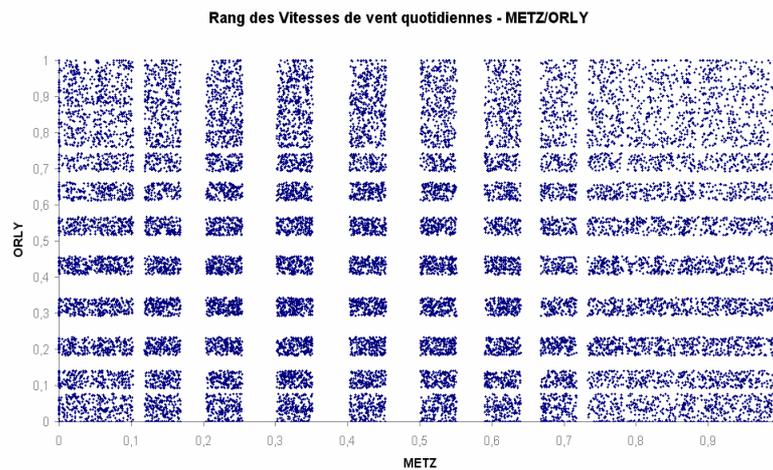
▪ Vitesses de vent historiques – Couples ($X_{i,t}$, $X_{j,t}$) :



- Q-Q plots historiques – Couples $(X_{i,t}, X_{j,t})$:



Les modélisations des vitesses de vent décrites dans la quatrième partie de ce mémoire sont fondées sur une approche individuelle de chaque station. Les éventuelles dépendances spatiales entre les vitesses de vent n'ont pas été simulées et ne sont donc pas prises en compte dans les estimations de la prime de la couverture. Le graphique « quantile-quantile » suivant concerne le couple (Metz-Orly), et est issu des simulations réalisées à partir du « modèle final » des ajustements des vitesses de vent au premier m/s :



On constate la dispersion plutôt uniforme des points, qui découle d'une simulation indépendante de chaque station, et qui contraste avec le nuage de points empirique présenté préalablement. Les représentations des autres couples ainsi modélisés dans la partie quatre sont similaires, et sont présentées en annexe 4.

Cette absence de prise en compte de la dépendance spatiale, et tout particulièrement celle des valeurs extrêmes, doit être l'origine des estimations basses du prix de la couverture que nous obtenons par les modélisations décrites dans le chapitre 4. Les tableaux suivants tendent à confirmer cette assertion.

Ce premier tableau indique la fréquence empirique (en ‰), sur l'intégralité de la période 1970-2002, du nombre de jours où les stations du couple (i,j) ont toutes deux enregistrées une vitesse de vent supérieure à leur seuil de modélisation respectif z_i :

fréquences empiriques en ‰	Metz	Nice	Orly
Bordeaux	0,747	0,498	0,830
Metz		0,415	2,489
Nice			0,332

Le tableau suivant indique ces mêmes fréquences, mais il est construit à partir de 1 million de simulations du « modèle final » des ajustements des vitesses de vent au premier m/s. On constate une très forte sous-estimation des fréquences empiriques :

fréquences modélisées en ‰	Metz	Nice	Orly
Bordeaux	0,076	0,052	0,048
Metz		0,081	0,059
Nice			0,040

La même approche appliquée aux nombres de jours où les stations des couples (i,j) ont simultanément enregistré des vitesses de vent supérieures à leur franchise K_i respectives nous donne les résultats suivants :

fréquences empiriques en ‰	Metz	Nice	Orly
Bordeaux	0,000	0,000	0,000
Metz		0,000	0,249
Nice			0,000

fréquences modélisées en ‰	Metz	Nice	Orly
Bordeaux	0,002	0,000	0,002
Metz		0,000	0,001
Nice			0,000

Ces différentes analyses nous confortent dans l'idée de chercher à affiner la tarification des couvertures indicielles en prenant en considération les situations de dépendance entre les stations.

5.2 Outils de mesure et de modélisation de la dépendance

Les notions et outils présentés dans cette partie ne sauraient rechercher l'exhaustivité, mais sont destinés à introduire les principaux éléments qui nous seront par la suite utiles dans l'étude et la modélisation de la dépendance spatiale de la couverture indicielle vent.

5.2.1 Préliminaire

La dépendance entre les variables aléatoires X_1, \dots, X_n est entièrement décrite par leur fonction de répartition conjointe :

$$F(x_1, \dots, x_n) = P[X_1 \leq x_1, \dots, X_n \leq x_n]$$

les variables aléatoires X_i , $i \in (1, \dots, n)$, ayant les fonctions de répartition suivantes, dites fonctions de répartition marginales :

$$F_i(x_i) = P[X_i \leq x_i]$$

5.2.2 Le coefficient de corrélation linéaire

Soient X et Y deux variables aléatoires. Le coefficient de corrélation linéaire, également appelé coefficient de corrélation de Pearson est :

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

où $cov(X,Y)=E(XY)-E(X)E(Y)$ est la covariance entre X et Y , et $Var(X)$ et $Var(Y)$ les variances respectives de X et Y . La définition du coefficient de corrélation linéaire est donc subordonnée à l'existence de variances finies de X et Y .

Le coefficient de corrélation linéaire a la propriété suivante :

$$-1 \leq r(X,Y) \leq 1$$

Deux variables indépendantes ont un coefficient de corrélation linéaire égal à zéro puisque $cov(X,Y) = 0$, mais la réciproque est fautive.

En cas de dépendance linéaire parfaite, c'est à dire $Y=\alpha X+\beta$, presque sûrement avec $\alpha \neq 0$, alors $r(X,Y)=\pm 1$ selon le signe de α .

Le coefficient de corrélation de Pearson est donc une mesure de dépendance linéaire, et il a pour propriété d'être invariable par transformation linéaire croissante :

$$r(\alpha X+\beta, \gamma Y+\delta) = \text{sign}(\alpha\gamma) \cdot r(X,Y)$$

où $\text{sign}(x)$ vaut 1 si $x>0$ et -1 sinon.

En revanche, si f et g sont deux fonctions non linéaires, alors :

$$r(f(X),g(Y)) \neq r(X,Y).$$

5.2.3 Le coefficient de corrélation de Spearman

Le coefficient de corrélation de Spearman est défini comme suit :

$$\rho(X,Y) = 3(P[(X-\tilde{X})(Y-Y') > 0] - P[(X-\tilde{X})(Y-Y') < 0])$$

où (X,Y) , (\tilde{X},\tilde{Y}) et (X',Y') sont des vecteurs de variables aléatoires identiques et indépendants.

Comme le coefficient de corrélation linéaire, le coefficient de corrélation de Spearman possède la propriété suivante :

$$-1 \leq \rho(X, Y) \leq 1$$

Le coefficient de corrélation de Spearman est une mesure de corrélation de rangs, et l'on peut montrer que :

$$\rho(X, Y) = r(F_X(X), F_Y(Y))$$

où F_X et F_Y sont respectivement les fonctions de répartition de X et de Y . Autrement dit, le coefficient de corrélation de Spearman mesure la corrélation linéaire de rangs.

Enfin, si f et g sont deux fonctions croissantes, linéaires ou non, alors :

$$\rho(f(X), g(Y)) = \rho(X, Y)$$

5.2.4 Simulation des corrélations

Soit ρ une matrice de corrélation de rangs, décrivant les coefficients de corrélation de Spearman des variables aléatoires (Y_1, \dots, Y_n) . Afin de simuler la corrélation entre les variables aléatoires (Y_1, \dots, Y_n) , nous utiliserons le logiciel de DFA ReMetrica. Le principe consiste à simuler dans un premier temps des variables aléatoires normales (X_1, \dots, X_n) avec la matrice de corrélation linéaire ρ . Dans un deuxième temps, on transforme les nombres générés afin d'obtenir les marginales : $(Y_1, \dots, Y_n) = (F_1^{-1}(\Phi(X_1)), \dots, F_n^{-1}(\Phi(X_n)))$. On peut montrer que la matrice de corrélation de rangs de (Y_1, \dots, Y_n) est identique à celle de (X_1, \dots, X_n) ⁶⁹ :

$$\rho(Y_i, Y_j) = \rho(X_i, X_j) = \frac{6}{\pi} \arcsin \frac{r(X_i, X_j)}{2} \approx r(X_i, X_j)$$

⁶⁹ Voir « *Correlation and dependency in risk management : Properties and Pitfalls* ». Paul Embrechts, Alexander McNeil, and Daniel Straumann.

5.2.5 Présentation des copules

Afin d'affiner la tarification de la couverture indicelle vent, l'utilisation des copules s'avère particulièrement intéressante, dans la mesure où celles-ci devraient permettre une meilleure représentation des dépendances statistiques existant entre les vitesses de vent des différentes stations.

Si les premiers travaux théoriques sur les copules remontent maintenant à plusieurs décennies (travaux de Sklar notamment, 1959), ce n'est que plus récemment (une dizaine d'années), que leur application pratique en finance et en actuariat se diffuse de façon remarquable. En effet, les copules offrent dans bon nombre de cas la possibilité de rendre compte, plus sincèrement qu'à l'aide de la mesure familière de corrélation, des structures de dépendance que l'on peut rencontrer dans la réalité. Elles nous seront utiles ici pour modéliser les dépendances propres aux vitesses de vent extrêmes et pourront s'appliquer dans notre cas multivarié (quatre stations).

- **Définition :**

Une copule est définie comme une fonction de répartition multivariée C définie sur $[0,1]^n$, dont les lois marginales sont uniformes sur $[0,1]$:

$$C(u_1, \dots, u_n) = P[U_1 \leq u_1, \dots, U_n \leq u_n]$$

La fonction C possède les propriétés suivantes :

- ❖ $C(u_1, \dots, u_n)$ est une fonction croissante de chaque composante u_i ,
- ❖ $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ pour tout $i \in (1, \dots, n)$

❖ Quels que soient (a_1, \dots, a_n) et $(b_1, \dots, b_n) \in [0, 1]^n$ avec $a_i \leq b_i$, on a :

$$\sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1 + \dots + i_n} C(u_{1i_1}, \dots, u_{ni_n}) \geq 0$$

où $u_{j1}=a_j$ et $u_{j2}=b_j$ pour tout $j \in (1, \dots, n)$

Par ailleurs, une copule C possède la propriété d'être invariante par transformations croissantes et continues des lois marginales. Si (X_1, \dots, X_n) a une copule C et que T_1, \dots, T_n sont des fonctions continues croissantes, alors $(T(X_1), \dots, T(X_n))$ a également C pour copule.

▪ **Densité des copules :**

Nous noterons c la densité d'une copule et qui s'exprime comme suit :

$$c(u_1, \dots, u_n) = \frac{\partial C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}$$

Si nous notons f_i la densité de la i -ème marginale, alors la densité f d'une distribution ayant F pour fonction de répartition peut s'écrire :

$$f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n)) \times \prod_{i=1}^n f_i(x_i)$$

▪ **Le théorème de Sklar :**

Soit H une fonction de répartition multivariée en n dimensions ayant des lois marginales F_1, \dots, F_n . Il existe une n -copule C telle que pour tout $x \in \mathbb{R}^n$:

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

Si F_1, \dots, F_n sont continues, alors C est unique ; sinon C est déterminée de façon unique en fonction de $E(F_1) \times \dots \times E(F_n)$ (où $E(F_i)$ est l'ensemble d'arrivée de la fonction F_i). Inversement, si C est une copule à n dimensions et que F_1, \dots, F_n sont des fonctions de répartition, alors la fonction H définie ci-dessus est une fonction de répartition à n dimensions ayant F_1, \dots, F_n pour marginales.

▪ **Corollaire du théorème de Sklar :**

Soient H une fonction de répartition multivariée en n dimensions ayant des lois marginales F_1, \dots, F_n et C une copule. Alors, pour tout vecteur $u \in [0, 1]^n$:

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

5.2.6 Distribution conditionnelle et simulation des copules

Soient $C_i(u_1, \dots, u_i) = C(u_1, \dots, u_i, 1, \dots, 1)$ avec $i \in (2, \dots, n-1)$ les distributions marginales en i dimensions de $C(u_1, \dots, u_n)$. Nous notons $C_1(u_1) = u_1$ et $C_n(u_1, \dots, u_n) = C(u_1, \dots, u_n)$. Si (U_1, \dots, U_n) sont liées par une copule C , alors la distribution conditionnelle de $(U_i | U_1, \dots, U_{i-1})$ peut être formulée à partir des dérivées et des densités des distributions marginales à i dimensions (si elles existent) :

$$C_i(u_i | u_1, \dots, u_{i-1}) = \frac{\frac{\partial^{i-1} C_i(u_1, \dots, u_i)}{\partial u_1 \dots \partial u_{i-1}}}{\frac{\partial^{i-1} C_{i-1}(u_1, \dots, u_{i-1})}{\partial u_1 \dots \partial u_{i-1}}}$$

On peut alors envisager d'appliquer l'algorithme suivant afin de simuler des variables (u_1, \dots, u_n) liées par une copule C^{70} :

- ▶ Simuler une valeur u_1 selon une loi uniforme $U(0, 1)$
- ▶ Simuler une valeur u_2 selon $C_2(u_2 | u_1)$
- ▶ ...
- ▶ Simuler une valeur u_n selon $C_n(u_n | u_1, \dots, u_{n-1})$

Pour simuler une valeur selon $C_i(u_i | u_1, \dots, u_{i-1})$, on peut simuler une valeur u selon une loi uniforme $U(0, 1)$, puis calculer si elle existe $C^{-1}(u | u_1, \dots, u_{i-1})$.

⁷⁰ Voir « Correlation and dependency in risk management : Properties and Pitfalls ». Paul Embrechts, Alexander McNeil, and Daniel Straumann.

Grâce à la propriété d'invariance de C , on peut ensuite utiliser les nombres générés (u_1, \dots, u_n) pour simuler des valeurs (y_1, \dots, y_n) liées par la même copule C .

5.2.7 Le coefficient de corrélation de Kendall

Le coefficient de corrélation de Kendall est défini comme suit :

$$\tau(X, Y) = P[(X - X')(Y - Y') > 0] - P[(X - X')(Y - Y') < 0]$$

où (X, Y) et (X', Y') sont des vecteurs de variables aléatoires identiques et indépendants.

Si (X, Y) a pour copule C , alors :

$$\tau(X, Y) = 4E[C(U, V)] - 1$$

5.3 Prise en compte de la dépendance dans la tarification : modélisation au premier m/s

Nous allons maintenant appliquer les outils de modélisation de la dépendance présentés dans la partie précédente afin d'affiner les estimations du prix de la couverture indiciaire vent étudiée en exemple. Dans un premier temps, nous utiliserons les ajustements des vitesses de vent au premier m/s des stations de Bordeaux, Metz, Nice et Orly, tels qu'ils ont été appliqués à l'issue de la partie 4.3.2 dans la modélisation « finale ».

5.3.1 Modélisation des corrélations

- **Matrices de corrélations**

Le tableau ci-dessous représente la matrice des coefficients de corrélation linéaire entre les observations des vitesses de vent des stations retenues :

$r(X_{i,t}, X_{j,t})$	Bordeaux $X_{1,t}$	Metz $X_{2,t}$	Nice $X_{3,t}$	Orly $X_{4,t}$
Bordeaux $X_{1,t}$	1	0,433	0,294	0,524
Metz $X_{2,t}$		1	0,200	0,739
Nice $X_{3,t}$			1	0,201
Orly $X_{4,t}$				1

On note que la plus forte corrélation linéaire observée est celle existant entre les stations de Metz et d'Orly. Ceci confirme les conclusions que nous avons pu tirer de l'observation graphique décrite en début de chapitre. Par ailleurs, on remarque que la vitesse de vent quotidienne enregistrée Nice est caractérisée par une faible dépendance linéaire vis-à-vis des autres stations.

Le tableau suivant représente la matrice des coefficients de corrélation de Spearman entre les stations retenues :

$\rho(X_{i,t}, X_{j,t})$	Bordeaux $X_{1,t}$	Metz $X_{2,t}$	Nice $X_{3,t}$	Orly $X_{4,t}$
Bordeaux $X_{1,t}$	1	0,398	0,210	0,496
Metz $X_{2,t}$		1	0,139	0,720
Nice $X_{3,t}$			1	0,148
Orly $X_{4,t}$				1

▪ **Résultats de la modélisation :**

Nous incluons la matrice de corrélation de Spearman au sein de la modélisation « finale » décrite dans le chapitre 4.3.2, et procédons à 50 000 simulations. Nous obtenons les nouvelles estimations suivantes du prix de la couverture :

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Rappel Modélisation sans prise en compte de la dépendance	103 081	317 307	166 542	230 004
Modélisation des corrélations	132 618	406 324	213 883	295 148

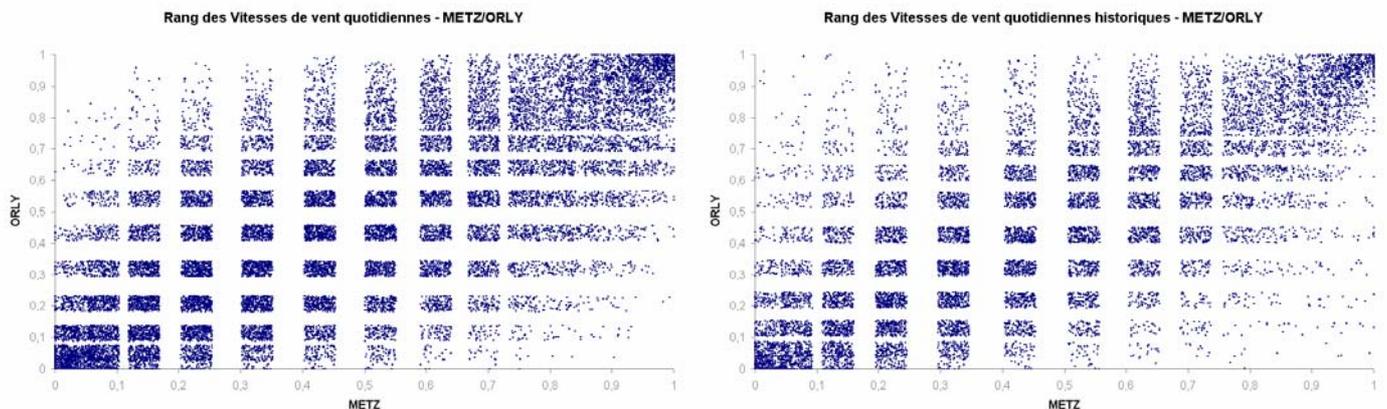
On constate donc une augmentation significative de la moyenne \bar{C} des indemnités annuelles $C(T)$ simulées, conduisant par conséquent à des estimations plus élevées du prix de la couverture que celles obtenues sans aucune modélisation des dépendances spatiales. Ces estimations demeurent toutefois nettement en deçà des moyennes historiques.

Les corrélations simulées permettent toutefois de mieux prendre en compte les relations entre les valeurs de vitesses de vent des quatre stations, comme

l'illustre la comparaison des graphiques quantile-quantile empiriques et simulés. Ces derniers sont exposés en annexe 4, nous présentons néanmoins ci-dessous ceux du couple Metz-Orly, pour lequel la dépendance historique est la plus importante. Il est à noter que le nuage historique comporte 12 053 points, alors que nous avons utilisé un échantillon de 20 000 simulations pour le nuage « simulation des corrélations ». Ceci rend malheureusement l'interprétation un peu moins immédiate.

QQ plot – simulation des corrélations

Rappel QQ plot empirique



Enfin, il convient de rester prudent dans les conclusions que l'on pourrait dresser sur la qualité de la modélisation effectuée en matière d'estimation de la prime de la couverture. Bien que la modélisation apparaisse relativement bonne dans l'ensemble, il est difficile de se rendre compte de sa capacité et retracer fidèlement le comportement des valeurs extrêmes.

Le tableau suivant permet de comparer les fréquences empiriques et simulées du nombre de jours où les stations de chaque couple (i,j) ont enregistré simultanément des vitesses de vent supérieures à leur seuil de modélisation respectif. Les fréquences empiriques sont indiquées en italique, par opposition aux fréquences obtenues après 1 million de simulations du modèle intégrant la matrice de corrélation de Spearman.

fréquences en ‰	Bordeaux	Metz	Nice	Orly
Bordeaux		0,676	0,176	0,728
Metz	0,747		0,227	2,487
Nice	0,498	0,415		0,151
Orly	0,830	2,489	0,332	

On procède de façon similaire en remplaçant les seuils z_i par les franchises K_i définies pour le déclenchement des indices par station.

fréquences en ‰	Bordeaux	Metz	Nice	Orly
Bordeaux		0,044	0,001	0,061
Metz	0,000		0,002	0,171
Nice	0,000	0,000		0,001
Orly	0,000	0,249	0,000	

Bien qu'imparfaites, les estimations obtenues par simulation nous semblent assez satisfaisantes, compte tenu de la rareté des réalisations historiques sur une période de 33 ans.

On peut toutefois considérer que les probabilités d'occurrences simultanées de valeurs extrêmes dans les différentes stations demeurent insuffisamment prises en compte dans la modélisation, ce qui pourrait expliquer l'inadéquation à nouveau constatée entre les charges historiques et estimées de la couverture.

Bien que difficilement exploitable dans notre exemple à quatre stations⁷¹, une comparaison des fréquences historiques et simulées pour les n-uplets de taille supérieure à deux semble abonder dans ce sens. Nous ne considérons que les fréquences des dépassements des seuils de modélisations z_i , aucune réalisation historique d'un dépassement des franchises pour un 3-uplet n'étant présente sur la période :

fréquences en ‰	empirique	corrélations
3-uplet Bordeaux-Metz-Orly ⁷²	0,581	0,327
4-uplet	0,083	0,016

⁷¹ L'approche est en revanche intéressante pour des couvertures indicielles avec un nombre plus élevé de stations.

⁷² Aucune réalisation historique pour les autres 3-uplets.

5.3.2 Modélisation par une Copule de Gumbel

Nous allons maintenant modéliser les dépendances entre les vitesses de vent des stations de la couverture indice vent, en utilisant une copule de Gumbel.

- **Présentation de la copule de Gumbel**

La copule de Gumbel est caractérisée par des concentrations importantes dans les queues de distribution. Etant asymétrique, la concentration est également plus importante à droite. Ceci s'avère intéressant pour modéliser les vitesses de vent dans le cadre de notre couverture indicielle, puisque ce sont avant tout les dépendances entre les fortes vitesses de vent qui influencent le comportement de l'indice S(T).

La copule de Gumbel bivariée est définie par :

$$C(u,v) = \exp\left(-\left((- \log(u))^a + (- \log(v))^a\right)^{1/a}\right) \text{ avec } a \geq 1$$

Sa fonction de densité est la suivante :

$$c(u,v) = C(u,v) u^{-1} v^{-1} \left((- \log(u))^a + (- \log(v))^a\right)^{-2+2/a} (\log(u) \cdot \log(v))^{a-1} \left(1 + (a-1) \left((- \log(u))^a + (- \log(v))^a\right)^{-1/a}\right)$$

Par ailleurs, il existe un lien direct entre le coefficient de corrélation de Kendall et le paramètre a de la copule de Gumbel :

$$\tau(a) = 1 - 1/a$$

▪ Approche de modélisation

La fonction $C_2(u|v)$ n'étant pas inversible, nous ne pouvons utiliser l'algorithme de simulation décrit précédemment. Nous utiliserons cependant une technique de simulation propre aux copulas archimédiennes⁷³.

Nous estimerons pour chaque couple de stations (i,j) avec $i \neq j$ le paramètre $a_{i,j}$ d'une copule de Gumbel bivariée. Nous appliquerons pour cela deux approches : d'une part l'utilisation des coefficients de corrélation de Kendall empiriques des couples de stations, d'autre part la méthode du maximum de vraisemblance.

Puis, en fonction des estimations des six paramètres $a_{i,j}$, nous modéliserons la copule de Gumbel de dimension 4 de la façon suivante :

Soit a_1 le paramètre $a_{i,j}$ ayant la valeur la plus élevée, qui correspond au couple de stations que nous notons (i_1, i_2) .

Soit a_2 le paramètre $a_{i,j}$ tel que $(i,j) \neq (i_1, i_2)$ et qui a la valeur la plus élevée. Nous noterons (i_1, i_3) le couple de stations auquel il correspond.

Soit enfin a_3 le paramètre $a_{i,j}$ tel que $(i,j) \neq ((i_1, i_2), (i_1, i_3), (i_2, i_3))$ et qui a la valeur la plus élevée. Le paramètre a_3 correspond donc au couple de la station que nous noterons i_4 et d'une des stations i_1, i_2 ou i_3 .

Nous modéliserons la dépendance du couple (X_{i_1}, X_{i_2}) en utilisant le paramètre a_1 .

Nous modéliserons la dépendance des couples (X_{i_1}, X_{i_3}) et (X_{i_2}, X_{i_3}) en utilisant le paramètre a_2 .

Enfin, nous modéliserons la dépendance des couples (X_{i_1}, X_{i_4}) , (X_{i_2}, X_{i_4}) et (X_{i_3}, X_{i_4}) en utilisant le paramètre a_3 .

Cette méthode nous force donc à accepter de modéliser la dépendance de certains couples de stations de façon identique. Par ailleurs, en retenant pour les paramètres a_2 et a_3 les valeurs les plus élevées possibles, nous avons fait le choix de modéliser les dépendances les plus faibles avec un risque de surestimation. Ce choix est subjectif, mais il est motivé par les résultats obtenus jusqu'à présent.

⁷³ Voir « *Modelling Dependence with Copulas and Applications to Risk Management* » Paul Embrechts, Filip Lindskog and Alexander McNeil.

- **Estimations des paramètres $a_{i,j}$ par les coefficients de Kendall :**

Le tableau ci-dessous représente la matrice des coefficients de corrélation de Kendall empiriques entre les stations retenues⁷⁴ :

$\tau(X_{i,t}, X_{j,t})$	Bordeaux $X_{1,t}$	Metz $X_{2,t}$	Nice $X_{3,t}$	Orly $X_{4,t}$
Bordeaux $X_{1,t}$	1	0,295	0,151	0,372
Metz $X_{2,t}$		1	0,100	0,570
Nice $X_{3,t}$			1	0,106
Orly $X_{4,t}$				1

Nous en déduisons directement les estimations $\hat{a}_{i,j}$ des paramètres $a_{i,j}$ de chaque copule de Gumbel $C(F^{-1}(X_{i,t}), F^{-1}(X_{j,t}))$:

Paramètres $\hat{a}_{i,j}$	Bordeaux $X_{1,t}$	Metz $X_{2,t}$	Nice $X_{3,t}$	Orly $X_{4,t}$
Bordeaux $X_{1,t}$	1	1,418	1,178	1,592
Metz $X_{2,t}$		1	1,111	2,326
Nice $X_{3,t}$			1	1,119
Orly $X_{4,t}$				1

Afin de mesurer la pertinence de la modélisation de la dépendance entre les variables $X_{i,t}$ et $X_{j,t}$ par une copule de Gumbel, nous effectuerons pour chaque couple de stations les calculs des fonctions $K(z)$ et de leur estimation non-paramétrique⁷⁵.

Pour une copule de Gumbel, la fonction $K(z)$ est définie par :

⁷⁴ Le calcul des coefficients empiriques de Kendall sur des échantillons de 12 053 valeurs pouvant être assez lourd pour les outils bureautiques standards, nous avons utilisé les bibliothèques de fonctions du logiciel R.

⁷⁵ Nous utiliserons ici les méthodes présentées dans le mémoire d'actuariat « *Tempêtes : Etude des dépendances entre les branches Auto et Incendie avec la théorie des copulas* », Olivier Belguise.

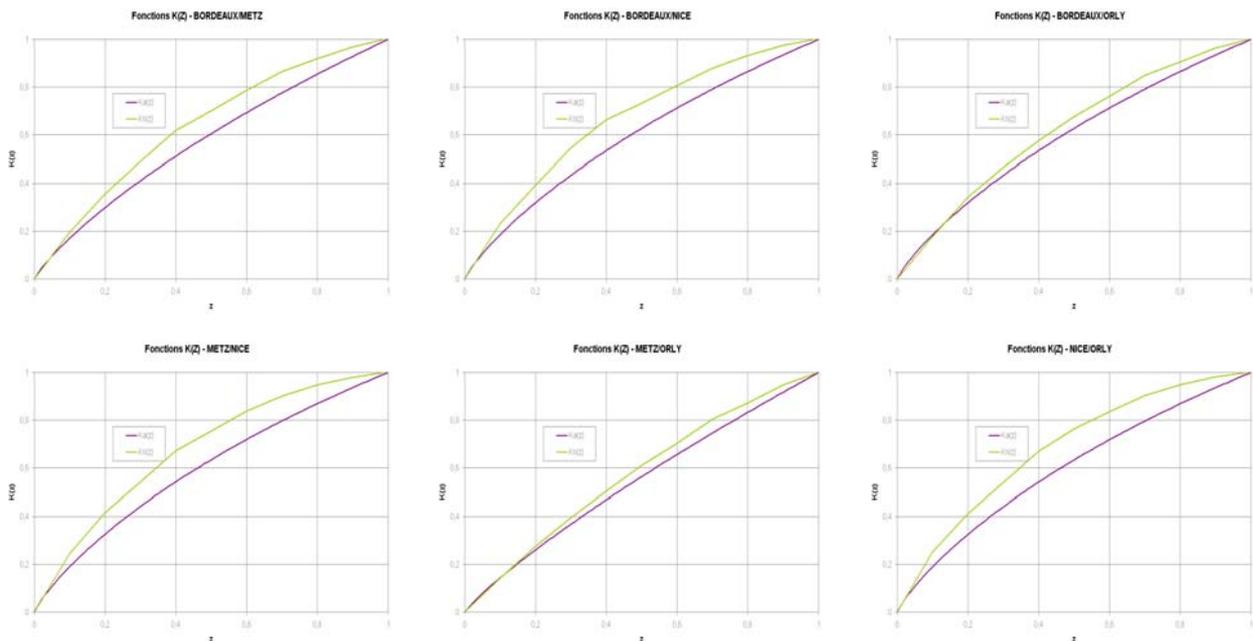
$$K_a(z) = z(1-\ln(z)/a)$$

Nous calculerons l'estimateur non paramétrique de $K(z)$ à partir des valeurs empiriques $X_{i,t}$ et $X_{j,t}$:

$$K_n(z) = \frac{\text{card}(\{z_i \text{ tq } z_i \leq z\})}{n}$$

$$\text{ou } z_i = \frac{\text{card}(\{(x_j, y_j) \text{ tq } x_j < x_i \text{ et } y_j < y_i\})}{n-1} \text{ pour } i \in (1, \dots, n)^{76}$$

La comparaison graphique de $K_a(z)$ et de $K_n(z)$ permet d'observer la qualité des ajustements :



On observe que les meilleurs ajustements sont ceux entre Metz et Orly, et entre Bordeaux et Orly. D'autres fonctions permettant d'étudier la qualité des ajustements, notamment sur les queues de distribution⁷⁷, pourraient s'avérer utiles dans l'analyse, mais ne seront malheureusement pas présentées ici.

⁷⁶ Le calcul du vecteur des 12 053 valeurs de z_i étant très consommateur de ressources informatiques, nous avons à nouveau utilisé la rapidité du logiciel R. La procédure écrite est présentée en annexe 5.

⁷⁷ Voir notamment les fonctions $L(z)$ et $R(z)$ présentées dans le mémoire d'actuariat « *Tempêtes : Etude des dépendances entre les branches Auto et Incendie avec la théorie des copulas* », Olivier Belguise.

- **Estimations des paramètres $a_{i,j}$ par le maximum de vraisemblance :**

En maximisant sur $\hat{a}_{i,j}$ la log-vraisemblance pour chaque couple de stations (i,j) avec $i \neq j$:

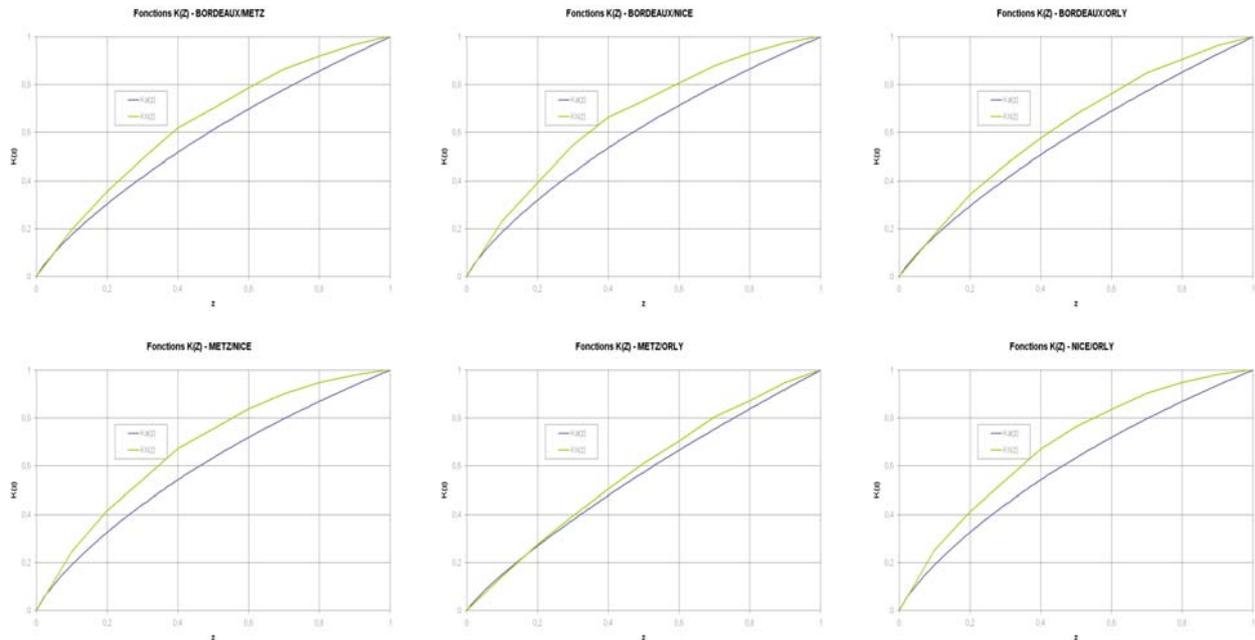
$$\sum_{t=1}^{12053} \log(c(F^{-1}(X_{i,t}), F^{-1}(X_{j,t})))$$

où $F^{-1}(X_{i,t})$ est la fonction de répartition inverse empirique de $X_{i,t}$ basée sur les 12 053 observations historiques, nous obtenons les estimations $\hat{a}_{i,j}$ suivantes :

Paramètres $\hat{a}_{i,j}$	Bordeaux $X_{1,t}$	Metz $X_{2,t}$	Nice $X_{3,t}$	Orly $X_{4,t}$
Bordeaux $X_{1,t}$	1	1,354	1,175	1,481
Metz $X_{2,t}$		1	1,114	2,025
Nice $X_{3,t}$			1	1,115
Orly $X_{4,t}$				1

On constate que les valeurs, bien que légèrement inférieures, sont comparables à celles obtenues par la méthode des coefficients de Kendall empiriques.

La comparaison graphique de $K_a(z)$ et de $K_n(z)$ permet d'observer la qualité des ajustements :



A nouveau, on observe que les meilleurs ajustements sont ceux entre Metz et Orly, et entre Bordeaux et Orly.

▪ **Modélisation et résultats :**

Nous décidons de retenir les estimations des paramètres $a_{i,j}$ obtenues à partir des coefficients de corrélation de Kendall empiriques, préférant le risque de surestimation à celui de sous-estimation des dépendances entre les stations.

Nous choisirons pour modéliser la copule de Gumbel les paramètres suivants :

$$a_1 = a_{2,4} = 2,326$$

$$a_2 = a_{1,4} = 1,592$$

$$a_3 = a_{1,3} = 1,178$$

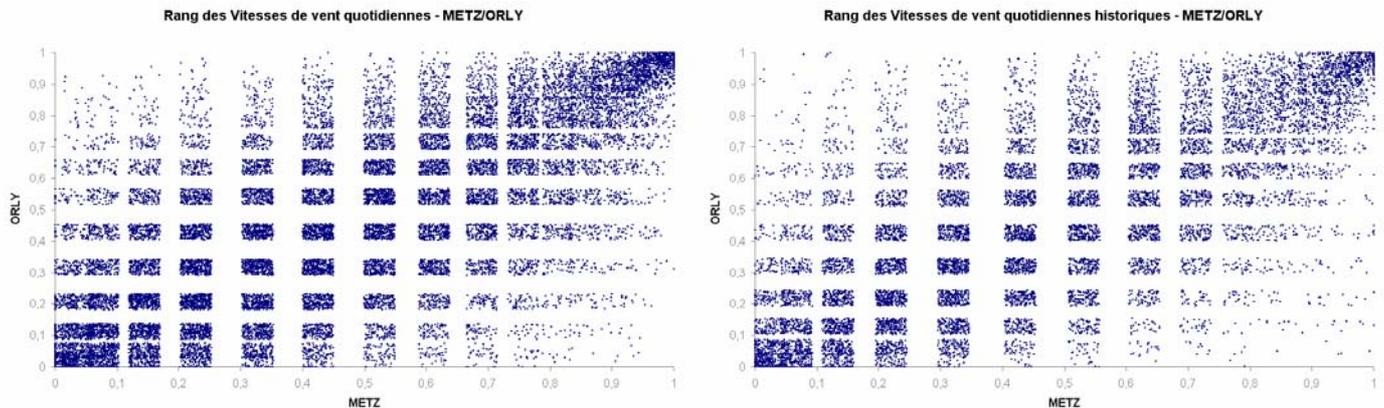
Au terme de 50 000 simulations, nous obtenons les résultats suivants :

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Modélisation sans prise en compte de la dépendance	103 081	317 307	166 542	230 004
Modélisation copule de Gumbel	206 373	615 838	329 541	452 708

On constate donc une augmentation significative de la moyenne \bar{C} des indemnités annuelles, des primes et également de la volatilité estimées. Les estimations demeurent certes inférieures aux moyennes historiques, mais nous paraissent plus crédibles que celles obtenues jusqu'à présent. Sans pour autant poursuivre aveuglément l'objectif du modèle reproduisant le passé à la perfection, force est d'admettre qu'une tarification, qui ne tient pas compte des évolutions potentielles des variables modélisées, ne doit pas aboutir à des résultats trop éloignés de la vérité historique. Elle serait autrement, à raison, contestable par les acteurs concernés.

Afin de mieux percevoir les mécanismes de dépendance modélisés, nous reprenons les analyses graphique et probabiliste introduites lors de la modélisation précédente. Les QQ plots ci-dessous offrent une comparaison empirique/modélisé des liens entre Metz et Orly⁷⁸. Les graphiques des autres stations sont exposés en annexe.

⁷⁸ Il est à noter que le nuage historique comporte 12 053 points, alors que nous avons utilisé un échantillon de 20 000 simulations pour le nuage « simulation copule de Gumbel ».



La modélisation apparaît relativement bonne dans l'ensemble.

Les tableaux suivants permettent de comparer les fréquences empiriques et simulées⁷⁹ des différents couples de stations. Les fréquences empiriques sont indiquées en italique.

Dépassements simultanés des seuil z_i :

fréquences en ‰	Bordeaux	Metz	Nice	Orly
Bordeaux		3,067	1,131	2,438
Metz	<i>0,747</i>		1,434	5,071
Nice	<i>0,498</i>	<i>0,415</i>		1,127
Orly	<i>0,830</i>	<i>2,489</i>	<i>0,332</i>	

Dépassements simultanés des franchises K_i :

fréquences en ‰	Bordeaux	Metz	Nice	Orly
Bordeaux		0,510	0,116	0,451
Metz	<i>0,000</i>		0,109	0,755
Nice	<i>0,000</i>	<i>0,000</i>		0,098
Orly	<i>0,000</i>	<i>0,249</i>	<i>0,000</i>	

⁷⁹ Fréquences obtenues après 1 million de simulations du modèle intégrant la copule de Gumbel.

Dépassements simultanés des seuils z_i pour les 3-uplets et le 4-uplet de stations :

fréquences en ‰	<i>empirique</i>	<i>copule de gumbel</i>
3-uplet Bordeaux-Metz-Orly ⁸⁰	0,581	2,265
4-uplet	0,083	0,738

La copule de Gumbel que nous avons retenue puis modélisée tend manifestement à surestimer les probabilités d'occurrences quotidiennes simultanées de vitesses de vent extrêmes. Ceci ne nous paraît pas choquant compte tenu des décisions prises lors de l'estimation des paramètres de la copule. Les montants estimés de la prime de la couverture ne sont quant à eux pas spécialement élevés, ce qui nous incite à penser que la modélisation peut encore être améliorée.

5.3.3 Modélisation par une Copule de Student

Nous utilisons à présent une copule de Student afin de modéliser les dépendances entre les stations de la couverture indice vent.

- **Preliminaire :**

Soit X un vecteur de variables aléatoires (X_1, \dots, X_n) en n dimensions. X suit une distribution de Student multivariée avec d degrés de liberté, un vecteur moyen μ et une matrice de dispersion positive définie Σ , si sa densité est donnée par :

$$f(x) = \frac{\Gamma(\frac{d+n}{2})}{\Gamma(\frac{d}{2}) \sqrt{(\pi d)^n |\Sigma|}} \left(1 + \frac{(x-\mu)' \Sigma^{-1} (x-\mu)}{d} \right)^{-\frac{d+n}{2}}$$

On dit que X suit une distribution $t_n(d, \mu, \Sigma)$.

⁸⁰ Aucune réalisation historique pour les autres 3-uplets.

- **Présentation de la copule de Student**

La copule de Student d'une distribution multivariée $t_n(d, \mu, \Sigma)$ est identique à celle d'une distribution $t_n(d, 0, P)$ où P est la matrice de corrélation correspondant à la matrice de dispersion Σ . La copule est donc unique et est définie par :

$$C_{d,P}^t(u) = \int_{-\infty}^{t_d^{-1}(u_1)} \dots \int_{-\infty}^{t_d^{-1}(u_n)} \frac{\Gamma(\frac{d+n}{2})}{\Gamma(\frac{d}{2}) \sqrt{(\pi d)^n |P|}} \left(1 + \frac{x' P^{-1} x}{d} \right)^{-\frac{d+n}{2}} dx$$

où t_d^{-1} est la fonction de répartition inverse d'une distribution de Student univariée standard.

La fonction de densité d'une copule de Student est la suivante :

$$c_{d,P}^t(u) = \frac{f_{d,P}(t_d^{-1}(u_1), \dots, t_d^{-1}(u_n))}{\prod_{i=1}^n f_d(t_d^{-1}(u_i))} \text{ avec } u \in (0,1)^n$$

où t_d est la fonction de répartition d'une loi de Student standard à d degrés de liberté, où $f_{d,P}$ est la fonction de densité d'une distribution de Student multivariée $t_n(d, 0, P)$, et où f_d est la fonction de densité standard d'une distribution de Student univariée à d degrés de liberté.

- **Approche de modélisation**

Afin de simuler une copule de Student, il suffit de générer un vecteur aléatoire multivarié X selon une distribution de Student multivariée $t_n(d, 0, P)^{81}$, puis de générer le vecteur $U = (t_d(X_1), \dots, t_d(X_d))'$.

⁸¹ Voir « *The t Copula and Related Copulas* » Stefano Demarta et Alexander J. McNeil.

Nous appliquerons une méthode d'estimation de la matrice P qui consiste à utiliser la matrice des coefficients de corrélation de Kendall empiriques $\tau(X_{i,t}, X_{j,t})$ et à construire la matrice P^* à partir des valeurs $\sin\left(\frac{\pi}{2}\tau(X_{i,t}, X_{j,t})\right)$.

Puis, nous estimerons d par la méthode du maximum de vraisemblance, en maximisant sur \hat{d} :

$$EMV = \sum_{i=1}^n \log(c_{\hat{d}, P^*}(\hat{U}_i))$$

où $\hat{U}_i = (F_1(X_{i,1}), \dots, F_d(X_{i,d}))'$ avec $F_d(X_{i,d})$ la valeur de la fonction de répartition empirique pour la i -ème observation, et n le nombre d'observations.

Une variante consisterait à maximiser la vraisemblance à la fois sur \hat{d} et sur P . Nous ne l'appliquerons cependant pas, compte tenu des difficultés de calcul rencontrées.

▪ **Modélisation et résultats :**

A partir de la matrice des coefficients de corrélation de Kendall empiriques, nous déduisons la matrice P^* ⁸²:

$\sin\left(\frac{\pi}{2}\tau(X_{i,t}, X_{j,t})\right)$	Bordeaux $X_{1,t}$	Metz $X_{2,t}$	Nice $X_{3,t}$	Orly $X_{4,t}$
Bordeaux $X_{1,t}$	1	0,447	0,235	0,552
Metz $X_{2,t}$	0,447	1	0,156	0,780
Nice $X_{3,t}$	0,235	0,156	1	0,166
Orly $X_{4,t}$	0,552	0,780	0,166	1

Nous obtenons par la méthode du maximum de vraisemblance l'estimation suivante du nombre de degrés de libertés de la copule de Student :

$$\hat{d} = 155$$

⁸² On peut constater que P^* est proche de la matrice des coefficients de corrélation linéaire. Celle-ci aurait ainsi pu constituer une bonne approximation.

Après 50 000 simulations, nous obtenons les résultats suivants :

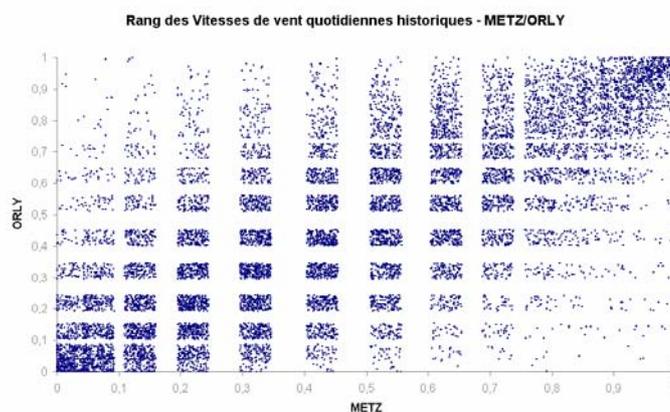
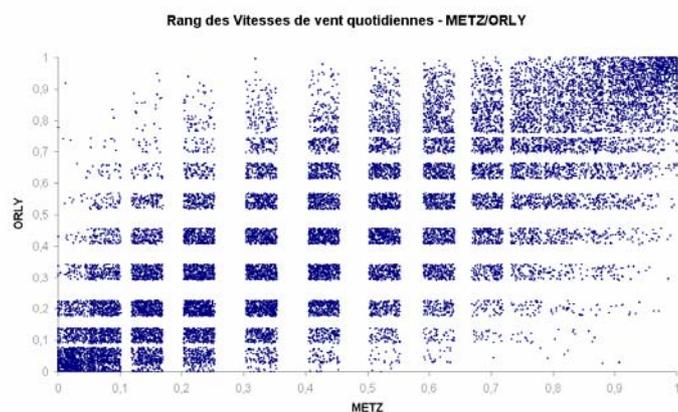
Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Modélisation sans prise				
en compte de la dépendance	103 081	317 307	166 542	230 004
Modélisation copule de Student				
	142 134	431 910	228 516	314 899

Les estimations sont légèrement supérieures à celles obtenues par simple modélisation des corrélations. Cependant, l'indemnisation moyenne \bar{C} estimée est de plus de 40% inférieure à la moyenne historique.

Nous présentons ci-après les analyses graphique et probabiliste introduites lors des modélisations précédentes. Les QQ plots ci-dessous offrent une comparaison empirique/modélisé des liens entre Metz et Orly⁸³.

QQ plot – simulation copule de Student

Rappel QQ plot empirique



⁸³ Il est à noter que le nuage historique comporte 12 053 points, alors que nous avons utilisé un échantillon de 20 000 simulations pour le nuage « simulation copule de Student ».

La modélisation apparaît relativement bonne dans l'ensemble.

Les tableaux suivants permettent de comparer les fréquences empiriques et simulées⁸⁴ des différents couples de stations. Les fréquences empiriques sont indiquées en italique.

Dépassements simultanés des seuil z_i :

fréquences en ‰	Bordeaux	Metz	Nice	Orly
Bordeaux		0,882	0,233	0,918
Metz	<i>0,747</i>		0,257	3,177
Nice	<i>0,498</i>	<i>0,415</i>		0,173
Orly	<i>0,830</i>	<i>2,489</i>	<i>0,332</i>	

Dépassements simultanés des franchises K_i :

fréquences en ‰	Bordeaux	Metz	Nice	Orly
Bordeaux		0,060	0,005	0,092
Metz	<i>0,000</i>		0,002	0,268
Nice	<i>0,000</i>	<i>0,000</i>		0,005
Orly	<i>0,000</i>	<i>0,249</i>	<i>0,000</i>	

Dépassements simultanés des seuil z_i pour les 3-uplets et le 4-uplet de stations :

fréquences en ‰	empirique	copule de student
3-uplet Bordeaux-Metz-Orly ⁸⁵	<i>0,581</i>	0,475
4-uplet	<i>0,083</i>	0,027

La modélisation par une copule de Student nous paraît assez satisfaisante en ce qu'elle restitue plutôt bien le comportement historique des observations quotidiennes jointes de valeurs extrêmes. Elle n'offre cependant pas une estimation de la prime de la couverture que nous jugeons acceptable.

⁸⁴ Fréquences obtenues après 1 million de simulations du modèle intégrant la copule de Student.

⁸⁵ Aucune réalisation historique pour les autres 3-uplets.

5.3.4 Conclusions

En intégrant la dépendance spatiale dans la modélisation des vitesses de vent au premier m/s, nous obtenons sans conteste des résultats bien plus satisfaisants pour notre couverture exemple.

La copule de Student, et dans une moindre mesure la matrice des corrélations de Spearman, permettent selon nous de modéliser assez correctement la dépendance entre les valeurs quotidiennes des stations, y compris celle des valeurs extrêmes. La qualité de modélisation de l'indice lui même est en revanche plus contestable. L'écart entre « les primes historiques » et les primes estimées demeure significatif.

La copule de Gumbel utilisée permet quant à elle une estimation plus crédible de la prime de la couverture (estimation plus proche de la moyenne historique, qui reste notre valeur de référence). Cependant nous avons remarqué que ce résultat repose sur un choix qui a exagéré les dépendances extrêmes. Ceci ne nous semble pas légitimement acceptable.

La famille des copules étant très large, il peut être intéressant d'utiliser des copules différentes. Une analyse plus approfondie de la qualité des ajustements effectués pour les copules de Gumbel et de Student est également envisageable. Toutefois, la réalisation de ces travaux nécessite déjà un temps certain pour notre exemple d'indice vent à quatre stations, en raison notamment des ressources informatiques nécessaires au traitement d'échantillons de données d'une telle taille. Dans le contexte plus général de couvertures où des dizaines de stations sont utilisées, l'approche en est encore plus ardue.

Ce problème n'est pas présent dans le cadre des modélisations des vitesses de vent supérieures à un seuil.

5.4 Prise en compte de la dépendance dans la tarification : modélisation avec seuils

Nous appliquons maintenant les outils de modélisation de la dépendance à partir des ajustements des vitesses de vent supérieures à un seuil.

Comme indiqué précédemment, la taille des échantillons de vitesses de vent à manipuler est bien plus faible, compte tenu des seuils que nous avons retenus :

Vitesses de vent quotidiennes supérieures aux seuils z_i	Bordeaux $X'_{1,t}$	Metz $X'_{2,t}$	Nice $X'_{3,t}$	Orly $X'_{4,t}$
Nombre d'observations	87	96	75	85

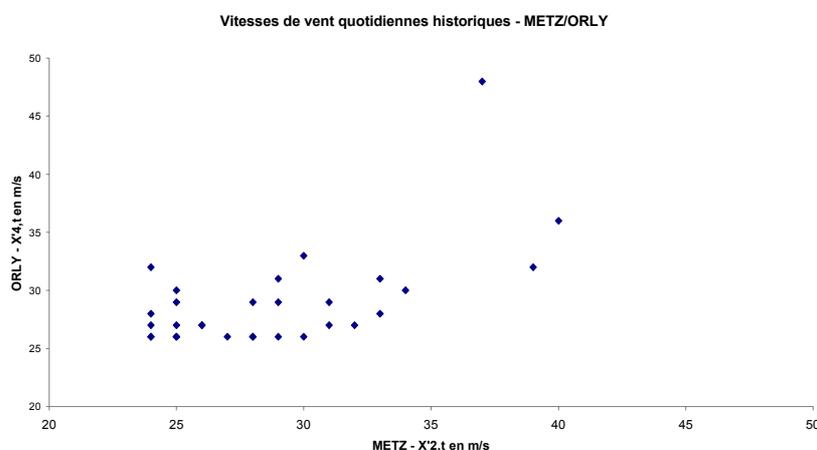
Un dénombrement rapide des dates communes à chaque couple de stations nous permet de dresser le tableau suivant. Il indique donc le nombre de jours sur la période 1970-2002 pour lesquels les vitesses de vent quotidiennes d'un couple de stations étaient simultanément supérieures aux seuils de modélisation (appliqués individuellement).

	Metz $X'_{2,t}$	Nice $X'_{3,t}$	Orly $X'_{4,t}$
Bordeaux $X'_{1,t}$	9	6	10
Metz $X'_{2,t}$		5	30
Nice $X'_{3,t}$			4

La relation la plus flagrante est celle existante entre Orly et Metz. Sur ces 30 observations communes de $(X'_{2,t}, X'_{4,t})$, on en dénombre 5 pour lesquels les vitesses de vent étaient supérieures ou égales aux franchises (K_2, K_4) . Il est à noter que nous ne dénombrons aucune observation de ce type pour les autres couples.

En conséquence, et compte tenu également du peu d'observations pour les autres couples, nous choisissons de ne prendre en compte dans les modélisations que la dépendance entre Metz et Orly.

Nous avons représenté les 30 observations communes du couple Metz-Orly sur le graphique suivant⁸⁶ :



Au préalable, nous présenterons néanmoins brièvement une approche de prise en compte de la dépendance entre les stations au travers de l'étude des corrélations entre les distributions du nombre annuel de valeurs supérieures aux seuils.

5.4.1 Modélisation de la corrélation des fréquences

Nous disposons pour chaque station i d'un historique de trente-trois valeurs de fréquences annuelles (nombre de jours sur l'année où la vitesse de vent est supérieure au seuil de modélisation), sur lequel nous avons ajusté les lois de Poisson et binomiale négative.

Nous modéliserons les corrélations entre les fréquences N_i et N_j à partir de la matrice des coefficients de corrélation de Spearman :

$\tau(N_i, N_j)$	Metz N_2	Nice N_3	Orly N_4
Bordeaux N_1	0,443	0,020	0,190
Metz N_2		-0,021	0,613
Nice N_3			0,269

⁸⁶ Les valeurs étant entières, certains points sont confondus.

A l'issue des analyses décrites dans le chapitre quatre, nous avons retenu des lois de Poisson pour N_2 et N_3 , et des lois binomiales négatives pour N_1 et N_4 . N'ayant pu simulé correctement les corrélations entre ces lois différentes, nous remplacerons les lois binomiales négatives par les lois de Poisson pour modéliser N_1 et N_4 . Le tableau suivant permet de constater qu'il n'y a pas de changement significatif dans les estimations obtenues. Les résultats de la modélisation retenue dans le chapitre quatre sont indiqués en italique.

	Bordeaux S₁(T)	Metz S₂(T)	Nice S₃(T)	Orly S₄(T)
<i>moyenne (dam/h)</i>	<i>646</i>	<i>630</i>	<i>67</i>	<i>505</i>
<i>écart type (dam/h)</i>	<i>1243</i>	<i>1198</i>	<i>276</i>	<i>1526</i>
nouvelle moyenne (dam/h)	648	630	67	506
nouvel écart type (dam/h)	1232	1197	276	1502

Après 100 000 simulations des modèles avec et sans prise en compte des corrélations de fréquences, nous obtenons les résultats suivants :

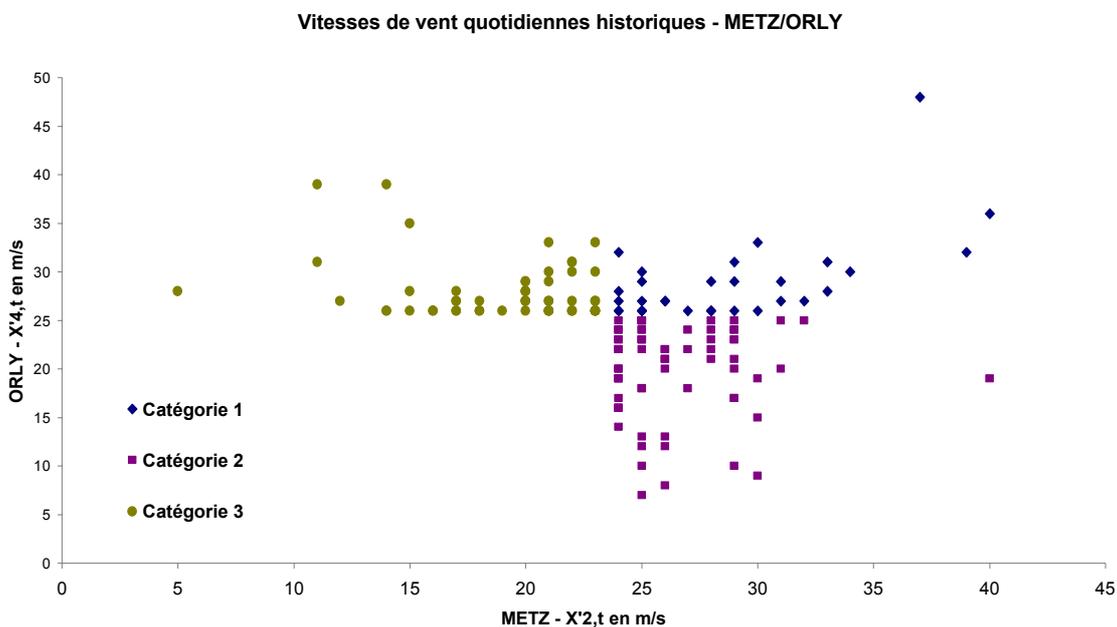
Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
Modélisation sans prise en compte de la dépendance	99 146	382 918	175 730	252 313
Modélisation des corrélations de fréquence	104 028	392 280	182 484	260 940

L'impact est manifestement marginal. Nous n'irons pas plus loin dans cette voie, car l'utilisation de copules ne paraît pas adaptée, dans la mesure où les dépendances entre les fréquences ne semblent pas très marquées à ce niveau de seuil.

5.4.2 Modélisation de la dépendance Metz-Orly à l'aide de copules

■ Préliminaire :

Si la modélisation des vitesses de vent avec introduction de seuils résout le problème de traitement des données, elle a l'inconvénient de multiplier les phases d'ajustement probabilistes. En effet, le graphique suivant permet de distinguer trois catégories différentes de situations :



La modélisation de la dépendance entre Metz et Orly nécessite donc un travail en trois étapes :

- La modélisation des « événements » appartenant à la catégorie 1, c'est à dire les observations de dépassements des seuils dans les deux stations.
- La modélisation des « événements » appartenant à la catégorie 2, c'est à dire les observations, où la vitesse de vent à la date t est supérieure au seuil à Metz mais pas à Orly. Nous ne modéliserons que les valeurs de la station de Metz.

- La modélisation des « événements » appartenant à la catégorie 3, c'est à dire les observations, où la vitesse de vent à la date t est supérieure au seuil à Orly mais pas à Metz. Nous ne modéliserons que les valeurs de la station d'Orly.

On se rend compte que l'approche peut rapidement rencontrer des limites, puisque le nombre d'étapes croît plus rapidement que le nombre de stations : l'application de cette approche à n stations, pour lesquelles on souhaiterait envisager tous les cas de figure envisageables en matière de dépendance, nécessite un nombre d'étapes égal à $\sum_{i=1}^n C_n^i$.

- **Modélisation des vitesses de vent de la catégorie 1 :**

Nous travaillons donc sur deux échantillons de 30 observations de vitesses de vent.

Après avoir appliqué les principes d'ajustement probabiliste tels que nous les avons présentés au cours de ce mémoire, voici les principales caractéristiques des modélisations qui seront effectuées pour la catégorie 1 :

Loi de fréquence commune :

loi binomiale négative de paramètres $a = 1,5$ et $p = 0,623$

Loi de sévérité marginale pour les vitesses de vent d'Orly :

loi de pareto généralisée de paramètres $a = 2,92$ et $c = 0,173$, et de seuil $z = 25,5$ m/s

Loi de sévérité marginale pour les vitesses de vent de Metz :

loi de pareto de paramètre $a = 5,169$ et de seuil $z = 23,5$ m/s

Nous appliquerons dans un premier temps ces paramètres sans modélisation supplémentaire de la dépendance. Par la suite nous utiliserons les copules de Gumbel puis de Student.

Nous avons estimé par la méthode du maximum de vraisemblance à 7,434 le paramètre a de la Copule de Gumbel bivariée Metz-Orly.

Pour la copule de Student, nous utiliserons la matrice $P^* = (0,53)$ et un nombre de degrés de liberté $d = 6$.

- **Modélisation des vitesses de vent de la catégorie 2 :**

Nous travaillons sur un échantillon de 66 observations pour lesquelles la vitesse de vent enregistrée à Metz était supérieure à 23,5 m/s, et où celle enregistrée à Orly était inférieure à 25,5 m/s.

Loi de fréquence :

loi binomiale négative de paramètres $a = 8$ et $p = 0,8$

Loi de sévérité :

loi de pareto généralisée de paramètres $a = 3,24$ et $c = -0,0712$, et de seuil $z = 23,5$ m/s

- **Modélisation des vitesses de vent de la catégorie 3 :**

Nous travaillons sur un échantillon de 55 observations pour lesquelles la vitesse de vent enregistrée à Orly était supérieure à 25,5 m/s, et où celle enregistrée à Metz était inférieure à 23,5 m/s.

Loi de fréquence :

loi binomiale négative de paramètres $a = 4,44$ et $p = 0,727$

Loi de sévérité :

loi de pareto généralisée de paramètres $a = 1,97$ et $c = 0,214$, et de seuil $z = 25,5$ m/s

- **Résultats des modélisations :**

Nous présentons tout d'abord les résultats obtenus station par station au terme de 100 000 simulations, et sans utilisation des copules :

	Bordeaux S ₁ (T)	Metz S ₂ (T)	Nice S ₃ (T)	Orly S ₄ (T)
Rappel : valeurs historiques				
Moyenne (dam/h)	704	589	93	491
Ecart type (dam/h)	1 378	2 461	263	1 685
Rappel : modélisation finale des vitesses de vent avec seuil (partie 4.3.3) :				
Moyenne (dam/h)	646	630	67	505
Ecart type (dam/h)	1243	1198	276	1526
Modélisation distincte des catégories 1, 2 et 3 pour Metz-Orly:				
Moyenne (dam/h)	646	880	67	506
Ecart type (dam/h)	1243	2255	276	1477

Les estimations pour Bordeaux et Nice sont bien évidemment inchangées. Les nouveaux ajustements réalisés ne modifient pas significativement la modélisation d'Orly, mais conduisent globalement à une meilleure restitution de la volatilité historique pour la station de Metz.

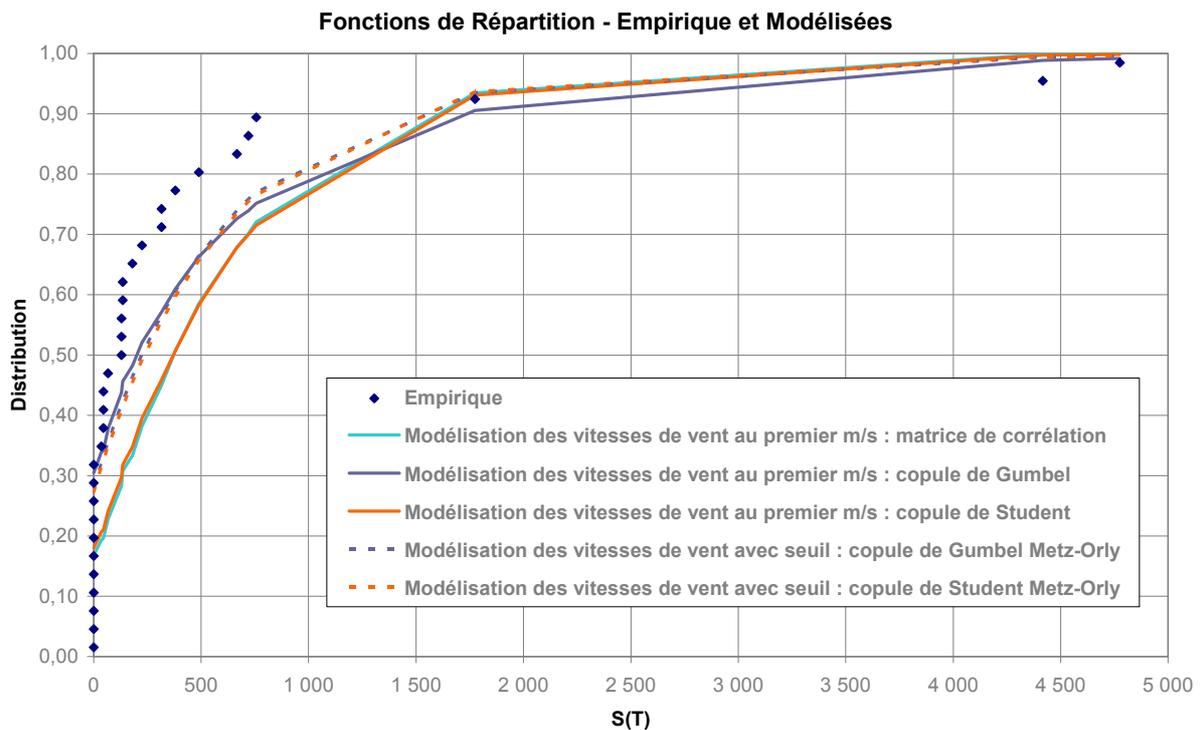
Nous obtenons les estimations suivantes du prix de la couverture, qui sont assez sensiblement inférieures aux moyennes historiques :

Estimation	\bar{C} (en EUR)	s (en EUR)	$\Pi_{20\%}$ (en EUR)	$\Pi_{40\%}$ (en EUR)
Rappel Burning Cost	241 436	876 919	416 820	592 204
nouvelle modélisation sans copule	138 625	465 590	231 743	324 861
nouvelle modélisation copule de Gumbel	146 554	500 303	246 615	346 675
nouvelle modélisation copule de Student	141 513	480 734	237 660	333 806

5.4.3 Conclusions

La prise en compte de la dépendance spatiale entre Metz et Orly dans la modélisation des vitesses de vent supérieures à un seuil permet globalement d'obtenir des résultats plus satisfaisants pour notre couverture exemple. La méthode peut cependant s'avérer fastidieuse dans le cadre de la tarification d'une couverture qui comporte un nombre élevé de stations présentant des liens de dépendance.

Toutefois, comme pour les ajustements de vent au premier m/s, nous n'obtenons pas une modélisation correcte de l'indice $S(T)$. L'observation des fonctions de répartition empirique et simulées nous le confirme.



Par conséquent, nous estimons que les ajustements effectués sur les vitesses de vent n'offrent pas à l'heure actuelle une tarification suffisamment satisfaisante de la couverture.

5.5 Introduction au problème de dépendance temporelle

Les méthodes employées sur les valeurs de vitesses de vent permettent pourtant de modéliser correctement d'une part le comportement des indices individuels par station, et d'autre part les dépendances spatiales entre les observations quotidiennes des différentes stations.

L'écart entre estimations empiriques et simulées peut en grande partie s'expliquer par le poids des années 1990 et 1999 dans l'historique des indemnisations de la couverture. En 1990, la station de Metz a enregistré six de ses sept valeurs les plus élevées sur la période 1970-2002. En 1999, Metz et Orly ont enregistré une valeur extrême le 26/12/1999 lors de la survenance de Lothar, Bordeaux le 27/12/1999 et Nice le 28/12/1999 lors de la survenance de Martin. Il semble ainsi que, pour une même station, l'occurrence et la valeur à la date t d'une vitesse de vent extrême ne soit pas indépendante des observations précédentes.

Ces situations extrêmes ont dans les modélisations que nous avons effectuées, une probabilité de survenance plus faible sur une échelle de temps d'une trentaine d'années.

Ainsi, dans les modélisations des vitesses de vent avec seuil, nous pensons qu'il est nécessaire, en vue d'obtenir une meilleure tarification de la couverture, d'étudier la dépendance par station entre la fréquence et la valeur des vitesses de vent supérieures aux seuils.

De même dans les modélisations des vitesses de vent au premier m/s, une prise en compte de la dépendance temporelle paraît légitime sur les variables de vitesses de vent $(X_{i,1}, \dots, X_{i,t}, \dots, X_{i,T})$.

Pour ce faire, nous conseillons l'étude d'une approche économétrique sur données temporelles (utilisant par exemple les vitesses de vent des jours précédents).

CONCLUSION

Après avoir présenté les principaux enjeux liés à l'utilisation et au fonctionnement d'une couverture indicielle Cat, nous nous sommes attachés à explorer les voies envisageables en matière de tarification. Nous avons pour cela travaillé sur l'exemple très simplifié d'une couverture indicielle basée sur les vitesses de vent maximales instantanées enregistrées quotidiennement dans les stations de Bordeaux, Metz, Nice et Orly.

L'étude des valeurs historiques de l'indice nous a permis tout d'abord d'estimer le prix de la couverture par l'intermédiaire du Burning Cost. Nous avons également vu qu'il était possible d'ajuster des lois de probabilité à ces valeurs historiques. Dans notre exemple, la modélisation par une loi gamma s'est avérée la plus satisfaisante. Afin de mieux prendre en compte le comportement de la queue de distribution, nous avons étudié l'impact sur la tarification de l'introduction d'un seuil de modélisation de l'indice à 100 dam/h. Plusieurs lois de probabilité, notamment celles recommandées pour l'étude des valeurs extrêmes, ont permis d'obtenir une tarification de la couverture dont on pourrait se satisfaire. Ces premières méthodes, simples et rapides, peuvent être appliquées à la grande majorité des couvertures que l'on peut imaginer en termes de variable climatique sous-jacente et de mécanisme d'indemnisation (niveau des franchises et des plafonds, période de couverture, etc.). Elles seraient toutefois difficilement applicables pour des protections très hautes (franchises élevées), ou de période de couverture supérieure à l'année. De plus le choix d'un seuil de modélisation est toujours problématique. Enfin, elles présentent l'inconvénient de fournir des cotations qui paraissent très sensibles à un historique de taille relativement réduite de valeurs d'indice, qui comporte une très grande volatilité.

Nous avons donc cherché dans un deuxième temps à exploiter les valeurs historiques de la variable climatique elle-même. L'obtention d'ajustements fiables pour chaque station permettrait d'envisager la tarification de n'importe quelle couverture indicielle (à sous-jacent climatique identique) sur des bases communes. Bien que la taille des échantillons historiques ait entraîné quelques difficultés dans la manipulation informatique des données, nous avons réussi à obtenir pour chaque station des ajustements assez satisfaisants des séries de vitesses de vent

quotidiennes, avec et sans introduction de seuils. Il est à noter que les ajustements au-delà des seuils individuels de chaque station nous ont paru plus robustes. Nous avons ainsi pu simuler le fonctionnement de la couverture indicielle, et obtenir des résultats station par station que nous estimons tout à fait acceptables. En revanche, les estimations du prix de la couverture ont été systématiquement inférieures aux tarifications sur base d'indice, et de façon significative.

En ayant mis en évidence le problème de la dépendance entre les stations et présenté des outils nous permettant de la modéliser, nous avons pu construire des modèles offrant une tarification affinée de la couverture. L'utilisation d'une simple matrice de corrélations, d'une copule de Gumbel et surtout d'une copule de Student ont permis de reproduire assez fidèlement les liens de dépendance constatés historiquement sur l'intégralité des vitesses de vent quotidiennes des stations. Cependant, bien que meilleures, les estimations du prix de la couverture nous paraissent encore trop faibles compte tenu de l'expérience passée. Un travail plus approfondi dans le choix des copules (utilisation de copules extrêmes notamment) et dans l'analyse de la qualité des ajustements pourrait être utile. Il nécessite toutefois des moyens informatiques très performants, et pourrait ne pas conduire à une amélioration significative de la tarification. Enfin, la prise en compte de la dépendance spatiale dans la modélisation avec introduction de seuils ne nous a pas permis d'obtenir des résultats plus satisfaisants. En outre, son application à un nombre élevé de stations est, comme nous l'avons fait remarquer, problématique.

Afin d'améliorer la tarification des couvertures indicielles vent, nous pensons qu'il est souhaitable d'étudier, par une approche économétrique, les dépendances temporelles entre les valeurs quotidiennes de vitesses de vent.

Une vision plus pragmatique peut néanmoins consister à retenir une tarification obtenue par surestimation des dépendances spatiales, de façon similaire à ce qui a été réalisé pour la copule de Gumbel multivariée.

Enfin, il conviendra de vérifier la pertinence des approches pour d'autres types de variables climatiques, dont les comportements pourraient être différents en matière de valeurs extrêmes et de dépendance spatiale.

BIBLIOGRAPHIE

▪ PARTIE 1 :

BELOW R. et GUHA-SAPIR D. [2002] : « *The Quality and Accuracy of Disaster Data. A comparative analyses of three global datasets* ». WHO Centre for Research on the Epidemiology of Disasters. University of Louvain School of Medicine.

BOBTCHEFF C. [2003] : « *La couverture des risques liés aux événements naturels : Cotation des traités en excédent de sinistre par événement* ». Mémoire d'actuariat ISFA.

DESCHEEMAERE D. et PERRON B. [1996] : « *Les événements naturels en France : étude du risque tempête pour une compagnie d'assurances* ». Mémoire d'actuariat CEA.

GAUDRON C. [1999] : « *Les protections en excédent de sinistre en réassurance des événements naturels en France* ». Mémoire d'actuariat ESSEC-ISUP.

MOISSELIN J., SCHNEIDER M., CANELLAS C. et MESTRE O. [2002] : « *Les changements climatiques en France au XX^e siècle. Etude des longues séries homogénéisées de données de température et de précipitations* ». Météo France. La Météorologie n°38, août 2002.

SIGMA [2004] : « *Catastrophes naturelles et techniques en 2003 : Un grand nombre de victimes, mais des dommages assurés modérés* ». Sigma n°1/2004, Swiss Re.

TOPICS [2004] : « *Annual Review : Natural Catastrophes 2003* ». Munich Re Group.

YEATMAN J. [1998] : « *Manuel International de l'Assurance* ». Ecole Nationale d'Assurances de Paris. Ed Economica.

▪ PARTIE 2 :

BANQUE MONDIALE [2000] : « *The Management of Catastrophic Risks using Pooled Insurance Structures and Alternative Financing & Risk Transfer Mechanisms. The Insurance Market and the case of the Caribbean Region* ». Discussion Draft January 27, 2000.

BENFIELD PARIS [2002,2003,2004] : « *Etude des Renouvellements des protections tempête du marché français* ».

CCR [2003] : « *Les catastrophes naturelles en France* ». CCR.

CHARPENTIER A. [2002] : « *Titrisation des risques catastrophes : les Cats-Bonds* ». FFSA / Direction Etudes, Statistiques et Systèmes d'Information.

COSSU P. et MANSION Y. [1996] : « *Assurance et Finance : Vers un pont à double voie ?* ». Risques n°28, Octobre-Décembre 1996.

COSTESEQUE L., FITE C. et FOUCAULT C. [1998] : « *Catastrophes Naturelles et Assurance* ». Mémoire de maîtrise de Sciences Economiques. Université des Sciences Sociales de Toulouse 1.

CUMMINS D.J. et GEMAN H. [1992] : « *An Asian Option Approach to the Valuation of Insurance Futures Contracts* ». Working Paper.

DREVETON C. [2002]: « *Evolution du nombre de tempêtes en France sur la période 1950-1999* ». Météo France.

FFSA [2001] : « *Catastrophes naturelles. Les différents systèmes d'indemnisation en Europe* ». Les nouvelles de l'assurance. Février 2001.

RISK PERCEPTION [2003] : « *Catastrophes naturelles et réassurance* ». Swiss Re.

SCOR TECH [1996] : « *L'assurance des catastrophes naturelles* ». Scor.

SIGMA [1996]: « *Insurance derivatives and securitization: New hedging perspectives for the US catastrophe insurance market?* ». Sigma n°5/1996, Swiss Re.

SIGMA [2001] : « *Les marchés de capitaux, source d'innovation pour le secteur de l'assurance* ». Sigma n°3/2001, Swiss Re.

SIGMA [2003] : « *Transfert alternatif des risques (ART) : état des lieux* ». Sigma n°1/2003, Swiss Re.

WRMA [2003] : « *The Weather Risk Management Industry. Survey Findings for April 2002 to March 2003* ». PriceWaterhouseCoopers.

▪ **PARTIE 3 :**

ARGUS DE L'ASSURANCE [2002] : « *Les dérivés climatiques : un marché dans le vent* ». Article paru dans le n°6779 du 29 Mars 2002.

DE PAIVA E PONA R. [2000] : « *Les Dérivés Climatiques : des Parapluies Financiers ?* ». Mémoire de Magistère Juriste d'Affaires. Université Paris 2.

METEO FRANCE [1996] : « *Normales Climatiques – Période 1961-1990 – Tome 1 – Stations de métropole* ». Service Central d'Exploitation de la Météorologie. Météo France

▪ **PARTIE 4 :**

COLES S.G. [2001] : « *An Introduction to Statistical Modelling of Extreme Values* ». Ed Springer Verlag.

D'AGOSTINO R.B. et STEPHENS M. [1986] : « *Goodness-of-fit techniques* ». Ed Marcel-Dekker, New York.

DAYKIN C.D., PENTIKAINEN T. et PESONEN M. [1994] : « *Practical Risk Theory for Actuaries* ». Ed Chapman & Hall, London.

DISCHEL B. [1998] : « *Black-Scholes won't do* ». Article paru dans Weather Risk Special Report, octobre 1998.

DORNIER F. et QUERUEL M. [2000] : « *Caution to the wind* », Article paru dans Weather Risk Special Report, août 2000.

EMBRECHTS P., KLUPPELBERG C. et MIKOSH T. [1997] : « *Modelling Extremal Events for Insurance and Finance* ». Ed Springer Verlag, 4th edition.

GUMBEL E.J. [1958] : « *Statistics of Extremes* ». Ed Columbia University Press.

HOGG R. et KLUGMAN S. [1984] : « *Loss Distributions* ». Ed Wiley.

MCNEIL A.J. [1997] : « *Estimating the tails of loss severity distributions using extreme value theory* ». Astin Bulletin, Vol 27.

PARTRAT C. [2003] : « *Valeurs extrêmes et périodes de retour* », issu de « *La Réassurance : Approche technique* ». Ed Economica

SMITH R.L. [2002] : « *Statistics of extremes, with applications in environmental science, insurance and finance* ». Department of Statistics, University of North Carolina. <http://www.stat.unc.edu/postscript/rs/semstatrls.ps>.

SMITS A. [2001] : « *Estimation of extreme return levels of wind speed : an analysis of storm maxima* ». KNMI.
<http://www.knmi.nl/samenw/hydra/documents/storms/storms.pdf>

STEPHENSON A. [2003] : « *A User's Guide to the evd Package (Version 2.0)* ». Mathematics and Statistics Dept – Lancaster University, UK.
<http://www.maths.lancs.ac.uk/~stephen/>

TANCREDI A., ANDERSON C. et O'HAGAN A. [2002] : « *Accounting for Threshold Uncertainty in Extreme Value Estimation* ». Dipartimento di Scienze Statistiche, Università di Padova, Italy and Department of Probability and Statistics, University of Sheffield, England.

▪ PARTIE 5 :

BELGUISE O. [2001] : « *Tempêtes : Etude des dépendances entre les branches Auto et Incendie avec la théorie des copulas* ». Mémoire d'actuariat ULP Strasbourg.

CHARPENTIER A. [2003] : « *Tail distribution and dependence measures* ». Colloque ASTIN, Berlin, 27 août 2003.

CHERUBINI U. et LUCIANO E. [2000] : « *Multivariate option pricing with copulas* ». Paper provided by ICER – International Centre for Economic Research in its series ICER Working Papers – Applied Mathematics Series with number 05-2002.

DEMARTA S. et MCNEIL A.J. [2004] : « *The t Copula and Related Copulas* ». Department of Mathematics, Federal Institute of Technology, ETH Zurich.

DURRLEMAN V., NIKEGHBALI A. et RONCALLI T. [2000] : « *Which copula is the right one ?* ». Groupe de Recherche Opérationnelle, Crédit Lyonnais.

EMBRECHTS P., MCNEIL A.J. et STRAUMANN D. [1998] : « *Correlation and Dependency in Risk Management : Properties and Pitfalls* ».

EMBRECHTS P., LINDSKOG F. et MCNEIL A.J. [2001] : « *Modelling Dependence with Copulas and Applications to Risk Management* ». Department of Mathematics, Federal Institute of Technology, ETH Zurich.

HASLETT J. et RAFTERY A.E. [1989] : « *Space-time modelling with long-memory dependence : Assessing Ireland's wind power resource* ». Royal Statistical Society.

NELSEN R. [1999] : « *An Introduction to copulas* ». Ed Springer Verlag.

VENTER G. [2000] : « *Tails of Copulas* ». Guy Carpenter InStrat.

VENTER G. [2003] : « *Fit to a t – Estimation, Application and Limitations of the t-copula. Topic 1 : Risk Evaluation* ». Guy Carpenter InStrat. Astin, Berlin 2003.

ANNEXES

Annexe 1 : Tableaux statistiques et rose des vents – Station de Metz

Les tableaux suivants sont extraits de l'ouvrage de Météo France (1996), « Normales Climatiques – Période 1961-1990 – Tome 1 – Stations de métropole ».



METZ

Latitude : 49°05' N Longitude : 6°08' E Altitude : 190 m Département : 57

Période des calculs : 1961-1990

Températures

Température minimale (en °C)

	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Aou	Sep	Oct	Nov	Déc	Année
Q80	+1,1	+1,6	+3,3	+5,0	+9,0	+12,5	+13,7	+13,3	+10,8	+7,8	+3,9	+1,8	+6,3
MOY	-1,1	-0,5	+1,7	+4,3	+8,0	+11,2	+12,9	+12,7	+9,9	+6,5	+2,4	-0,1	+5,7
Q20	-4,0	-2,1	+0,0	+3,3	+7,2	+10,2	+11,9	+12,0	+8,7	+5,0	+1,1	-1,4	+5,2
TNABS DATE	-20,1 02/01/71	-17,8 05/02/63	-12,2 06/03/71	-5,1 13/04/66	-1,8 06/05/79	+1,9 02/06/75	+4,3 22/07/80	+4,7 24/08/80	-1,1 29/09/72	-3,8 25/10/65	-9,4 26/11/89	-17,0 03/12/73	-20,1 02/01/71

Température moyenne (en °C)

Q80	+3,4	+4,8	+7,5	+10,2	+14,5	+17,6	+19,8	+19,0	+16,0	+11,6	+6,6	+3,9	+10,6
MOY	+1,5	+2,8	+5,8	+9,1	+13,2	+16,4	+18,4	+18,0	+15,0	+10,6	+5,3	+2,4	+9,9
Q20	-0,8	+1,1	+4,4	+8,3	+12,0	+15,4	+16,8	+16,7	+13,8	+9,5	+4,1	+1,1	+9,3

Température maximale (en °C)

Q80	+5,6	+8,3	+12,1	+15,6	+20,5	+23,1	+25,9	+25,1	+21,3	+16,1	+9,5	+6,9	+14,9
MOY	+4,0	+6,2	+9,9	+13,9	+18,3	+21,6	+23,8	+23,4	+20,2	+14,7	+8,3	+4,9	+14,1
Q20	+2,5	+4,3	+8,0	+12,6	+16,5	+20,1	+22,0	+21,5	+18,7	+13,7	+6,8	+3,4	+13,3
TXABS DATE	+15,7 15/01/75	+20,4 24/02/90	+24,3 30/03/89	+28,2 22/04/68	+29,7 08/05/76	+34,5 30/06/76	+36,6 18/07/84	+35,4 03/08/86	+34,0 03/09/82	+28,8 04/10/66	+19,7 05/11/63	+17,5 16/12/89	+36,6 18/07/84

Nombre de jours de gel

Q80	22,6	20,0	17,0	5,0	1,0	0,0	0,0	0,0	0,0	3,0	14,0	19,8	67,6
MOY	16,5	14,4	10,6	3,6	0,3	0,0	0,0	0,0	0,0	1,8	9,0	15,2	71,6
Q20	11,0	8,0	5,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	4,0	10,2	62,4

Nombre de jours avec :

TX >= 30°C	0,0	0,0	0,0	0,0	0,0	0,8	3,2	2,2	0,5	0,0	0,0	0,0	6,7
TX >= 25°C	0,0	0,0	0,0	0,2	2,9	7,7	12,4	10,7	3,9	0,3	0,0	0,0	38,0
TX <= 0°C	5,5	2,7	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,6	4,4	13,5
TM >= 18°C	0,0	0,0	0,0	0,1	2,6	9,9	15,8	15,7	5,2	0,5	0,0	0,0	49,8
TN <= -5°C	6,3	4,2	1,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,9	4,8	17,7

Nombre d'années sans gel : 0

Date de la première gelée sous abri

Date de la dernière gelée sous abri

La plus précoce : 29/09/72

La plus tardive : 19/05/68

Médiane : 27/10

Médiane : 20/04

La plus tardive : 27/11/77

La plus précoce : 25/03/80

METZ

Latitude : 49°05' N Longitude : 6°08' E Altitude : 190 m Département : 57

Période des calculs : 1961-1990

Précipitations

Cumul mensuel des précipitations (en mm)

	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Aou	Sep	Oct	Nov	Déc	Année
Q80	91,0	80,8	80,2	88,8	85,3	108,5	85,6	93,3	90,6	95,3	82,6	119,2	876,3
MOY	63,5	57,7	63,1	53,5	68,9	72,0	61,5	62,5	59,7	63,5	66,6	73,0	765,6
Q20	34,6	16,3	37,4	25,9	45,4	37,4	43,5	35,9	30,6	16,0	48,6	37,2	635,2

Hauteur maximale des précipitations en 24 h (en mm)

	43,6	32,7	32,4	32,9	41,1	49,5	45,4	38,5	48,0	56,5	34,3	28,6	56,5
DATE	12/01/62	14/02/90	24/03/66	24/04/68	11/05/70	07/06/63	24/07/72	30/08/90	21/09/67	15/10/61	18/11/71	31/12/78	15/10/81

Durée mensuelle des précipitations (en heures)

MOY	123,3	100,0	99,7	80,7	67,5	56,3	40,2	44,4	48,3	72,3	102,4	116,6	947,7
-----	-------	-------	------	------	------	------	------	------	------	------	-------	-------	-------

Nombre de jours avec précipitations (RR) :

RR>=1mm	12,7	9,5	11,3	10,1	11,7	10,0	8,7	9,5	8,5	9,7	10,7	11,0	123,6
RR>=5mm	3,7	4,0	4,3	3,8	5,0	5,0	4,1	4,2	4,0	4,2	4,3	5,3	51,8
RR>=10mm	1,5	1,6	1,5	1,1	1,9	2,1	2,1	1,8	1,8	1,8	1,8	2,2	21,2

Insolation

Durée totale d'insolation (en heures)

	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Aou	Sep	Oct	Nov	Déc	Année
Q80	56,7	102,2	158,4	217,0	229,7	242,4	295,0	248,1	190,3	126,4	66,1	54,3	1817,8
MOY	43,2	61,1	122,4	163,1	202,4	213,1	238,5	209,1	162,6	104,0	55,6	43,0	1638,1
Q20	27,9	57,5	84,1	122,7	163,3	177,6	201,0	178,4	126,5	78,4	41,4	29,7	1513,3

Nombre de jours avec fraction d'insolation (sigma) :

sigma=0	15,2	6,1	5,3	3,7	2,3	1,3	0,8	1,3	2,0	6,1	11,7	16,2	74,0
sigma>=0,8	0,9	2,8	3,2	4,4	4,8	4,0	5,6	5,1	4,1	2,7	1,6	1,5	40,7

METZ

Latitude : 49°05' N Longitude : 6°08' E Altitude : 190 m Département : 57

Période des calculs : 1961-1990

Humidité

Humidité relative minimale

	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Aou	Sep	Oct	Nov	Déc	Année
Q80	79	71	62	57	54	55	56	56	62	72	78	80	63
MOY	77	66	56	50	50	51	49	51	57	67	74	77	60
Q20	73	60	49	44	46	45	43	44	52	63	71	74	58

Humidité relative moyenne

Q80	89	86	81	79	78	78	79	81	85	89	90	90	82
MOY	87	82	78	73	74	74	73	76	81	87	87	88	80
Q20	85	79	73	68	69	69	69	70	78	84	86	86	78

Phénomènes

Nombre de jours avec :

	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Aou	Sep	Oct	Nov	Déc	Année
Brouillard	5,5	3,2	2,2	2,1	2,5	1,8	1,4	3,3	7,7	9,9	6,5	5,2	51,3
Orage	0,2	0,2	0,5	1,1	3,8	4,8	4,1	4,0	1,8	0,4	0,2	0,2	21,3
Grêle	0,1	0,3	0,5	0,4	0,3	0,4	0,1	0,0	0,1	0,0	0,2	0,2	2,8
Neige	7,8	6,3	4,6	1,9	0,1	0,0	0,0	0,0	0,0	0,1	3,0	5,7	29,5

Vent

Nombre de jours avec vent maximal instantané (FXI) :

	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Aou	Sep	Oct	Nov	Déc	Année
FXI >= 16m/s	5,7	4,6	6,5	3,3	2,7	2,2	1,7	1,3	2,2	3,9	2,8	4,3	40,4
FXI >= 28m/s	0,3	1,0	0,4	0,0	0,2	0,1	0,1	0,0	0,1	0,0	0,4	0,1	2,2

Maximum absolu du vent maximal instantané (en m/s)

VIT	30	40	40	23	41	30	29	25	29	26	29	28	41
DIR	280	240	280	260	250	200	160	240	280	260	260	260	250
DATE	22/01/88	03/02/90	01/03/90	07/04/85	11/05/89	02/06/82	30/07/82	03/08/84	15/09/85	20/10/86	24/11/84	19/12/88	11/05/89

METZ



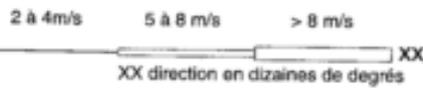
Latitude : 49°05' N Longitude : 6°08' E Altitude : 190 m Département : 57

Période des calculs : 1981-1990

Fréquences moyennes annuelles des directions du vent en % par groupes de vitesses

Valeurs trihoraires de 00 à 21 heures UTC

Groupes de vitesses



Pourcentage par direction

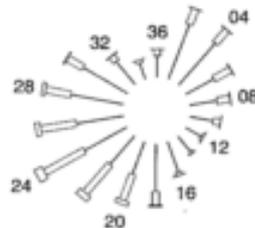


Tableau récapitulatif (en pourcentage)

dir (en °)	2-4 m/s	5-8 m/s	>8 m/s	tot
020	4,3	1,2	+	5,6
040	4,1	1,3	+	5,4
060	2,3	1,4	+	3,7
080	2,0	1,0	+	3,0
100	1,7	0,5	+	2,3
120	1,3	0,2	+	1,5
140	1,5	0,1	0,0	1,6
160	2,3	0,2	+	2,6
180	3,5	1,1	0,1	4,7
200	2,5	2,1	0,3	4,9
220	2,8	2,8	0,5	6,0
240	3,6	3,4	0,6	7,6
260	3,5	2,7	0,4	6,5
280	3,9	1,9	0,3	6,1
300	3,8	1,3	0,1	5,2
320	2,2	0,5	+	2,8
340	1,2	0,2	+	1,5
360	1,6	0,5	+	2,1
tot	48,1	22,2	2,5	72,8

Le signe + indique une fréquence non nulle mais inférieure à 0,05%

Fréquence des vents inférieurs à 2 m/s = 27,2%

Nombre de cas observés = 29046

Nombre de cas manquants = 170

Annexe 2 : Fiches descriptives des séries utilisées



© CatiXL 2003

GENERAL STATION INFORMATION SERIE 19 - BORDEAUX-MERIGNAC (m/s)

STATION

Usual Name		Alternate Name	
BORDEAUX-MERIGNAC		AEROPORT DE BORDEAUX	
Country		Latitude	Longitude
France		44,832 °N	0,692 °W
Local ID	WMO ID	Altitude	
33281001	7510	47 m above sea level	

ELEMENT

Element	Measure Details	Precision
Maximum Daily Wind Speed	00-24 UT	Nearest Integer.
Distribution Threshold	Recording Units	Start Date
Above 18,0	m/s	01janv1970

DATA COMPLETENESS

Bordeaux has less than 0.2% missing values. Data from St Gervais and Cognac was used after correlation analysis.

The final serie has no missing values.

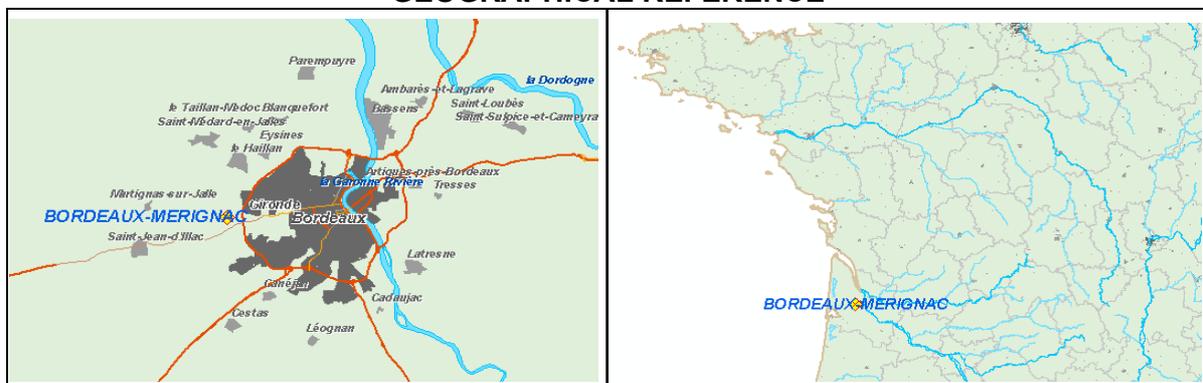
AVAILABLE METADATA

Due to important heterogeneities in most French wind records prior to 1970, all time series have been truncated to Jan 1st, 1970. Contact CatiXL for more information.

The station first opened on 19200101. It was relocated on 20001026.

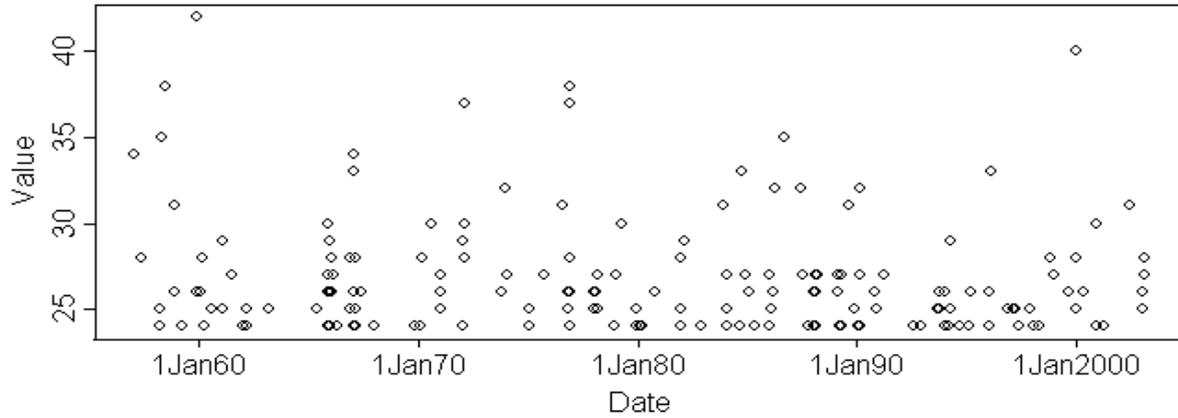
The first Wind recordings go back to 19490101.

GEOGRAPHICAL REFERENCE



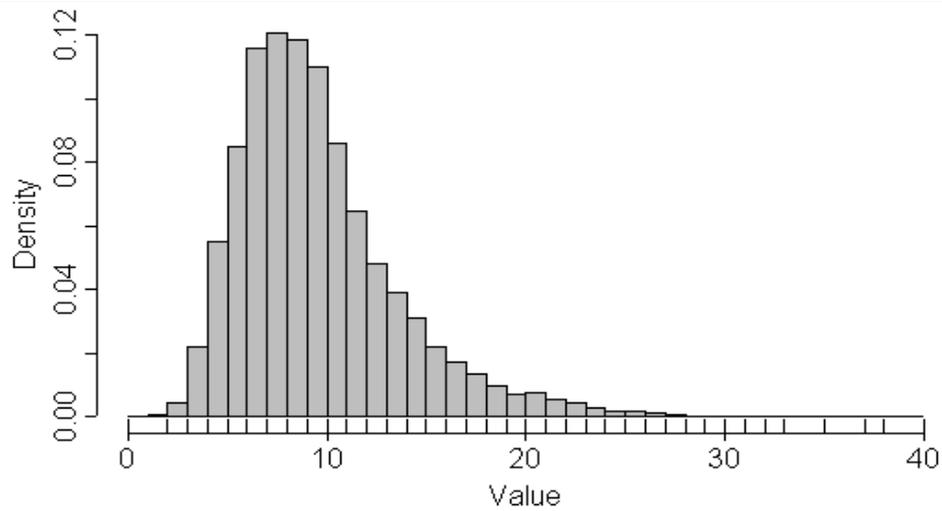
STATION STATISTICS
SERIE 19 - BORDEAUX-MERIGNAC (m/s)

EXTREME HISTORICAL VALUES



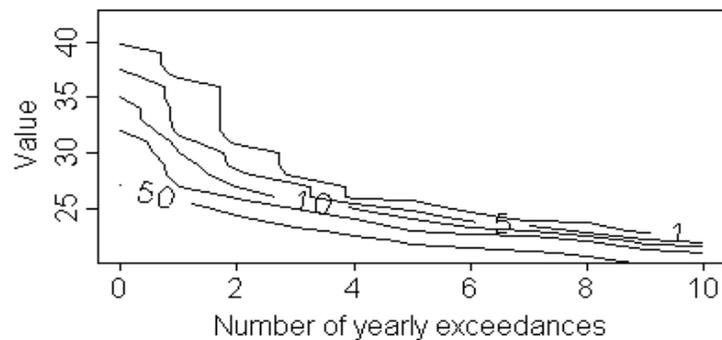
DATA DISTRIBUTION (1970 - 2000)

Quantile	Value
95,0%	18,00
97,5%	21,00
98,0%	22,00
98,5%	23,00
99,0%	24,00
99,5%	26,00
100,0%	40,00



PROBABILITY OF EXCEEDANCE (1970 - 2000)

	5%	10%	20%
2 days	28,50	27,00	25,86
4 days	25,50	25,00	24,00
6 days	23,83	23,33	22,75
8 days	22,90	22,60	22,00

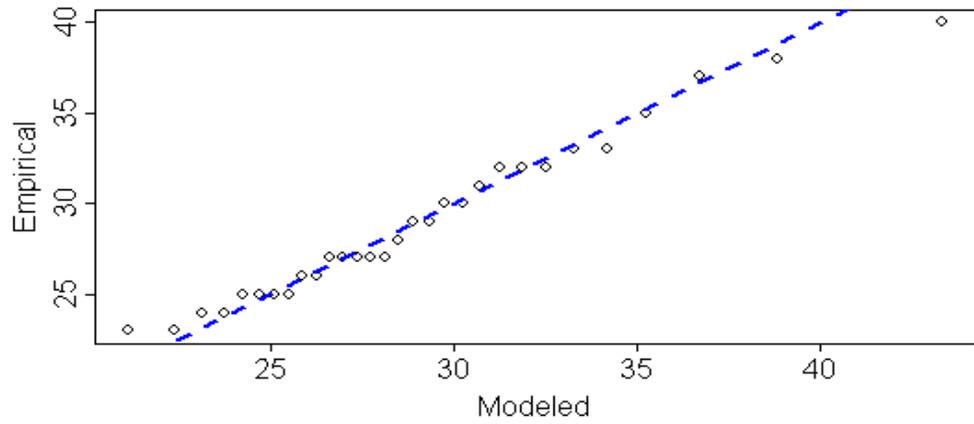


RETURN PERIOD ESTIMATION
SERIE 19 - BORDEAUX-MERIGNAC (m/s)

GUMBEL RETURN PERIOD (1970 - 2000)

T (Years)	Value
2	28,31
5	32,89
10	35,92
15	37,63

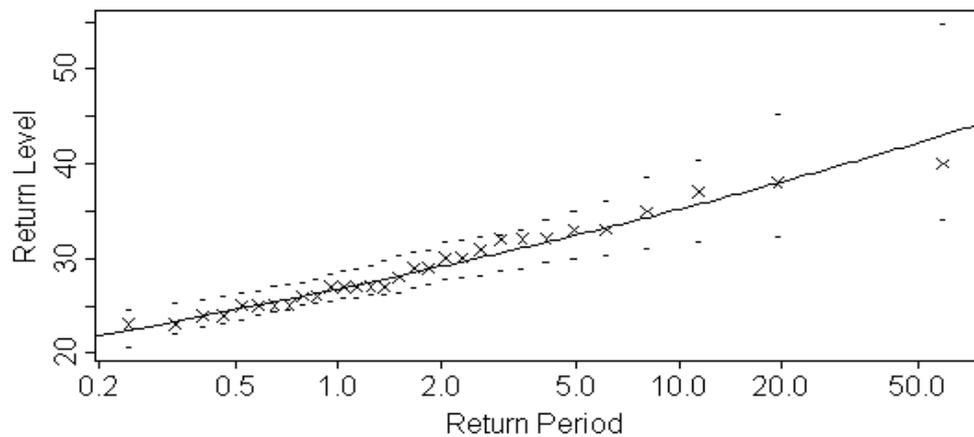
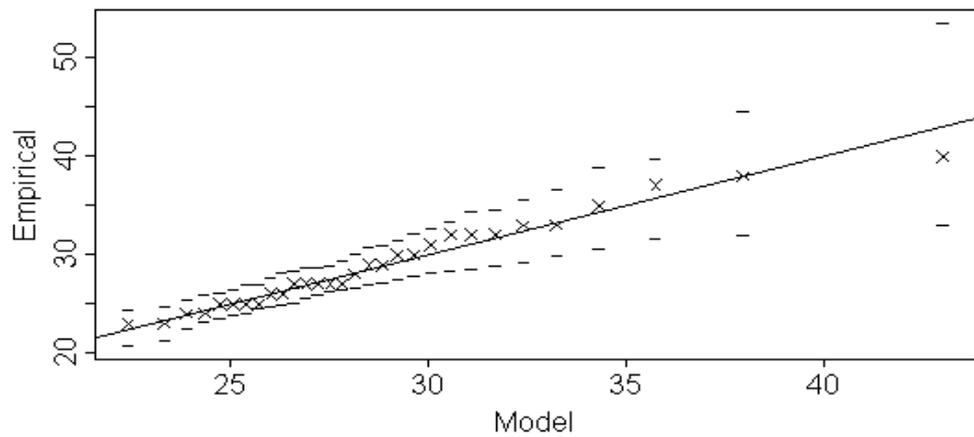
Parameters	
Location	26,825
Scale	4,041



GEV RETURN PERIOD (1970 - 2000)

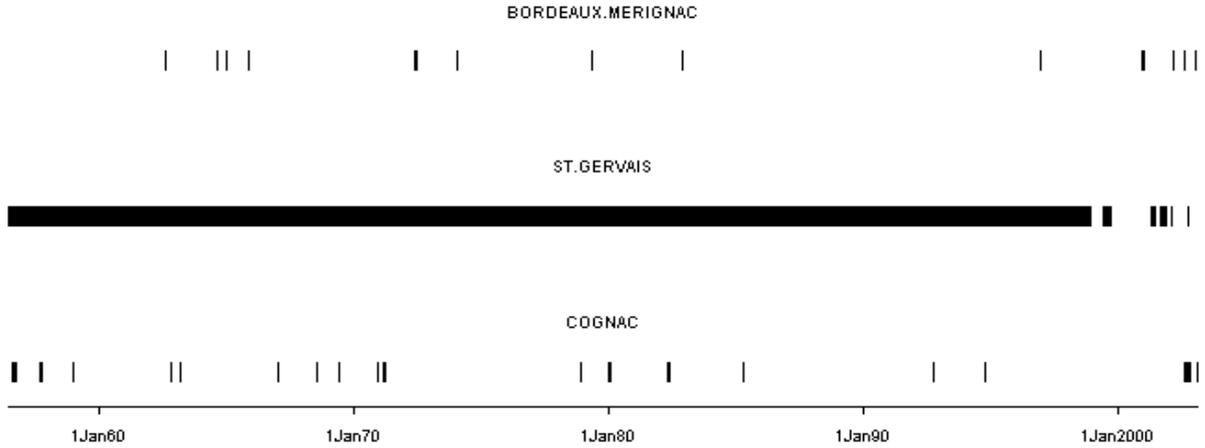
T (Years)	Value
2	28,02
5	32,06
10	34,98
15	36,72

Parameters	
Location	26,793
Scale	3,280
Shape	0,091



STATION FALLBACK INFORMATION
SERIE 19 - BORDEAUX-MERIGNAC (m/s)

AVAILABLE SURROUNDING DATA



PRINCIPAL FALLBACK
ST-GERVAIS

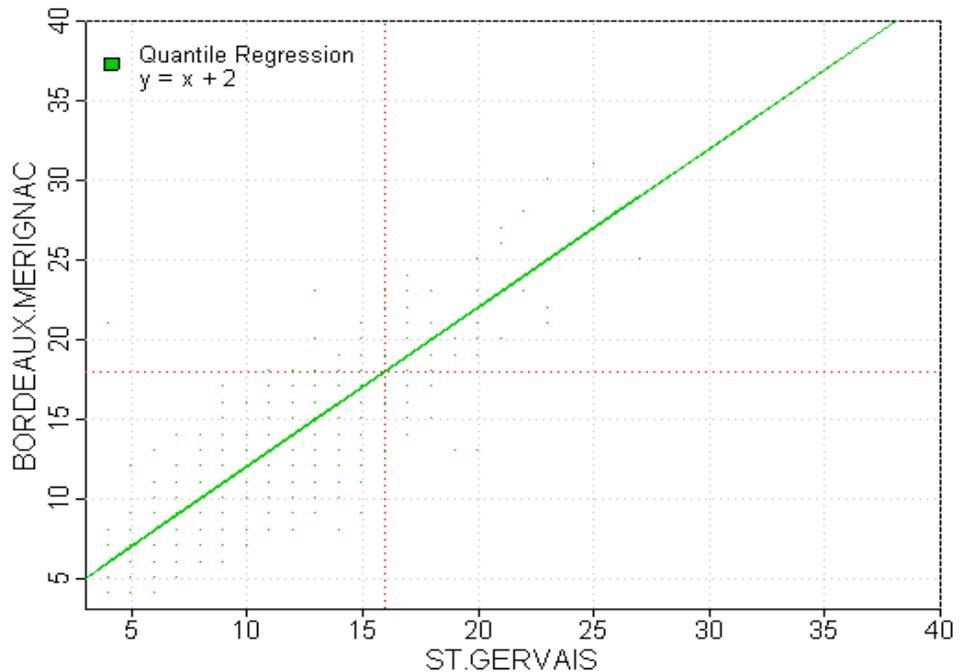
Station Name	
Local ID	WMO
33415001	

Correlation Dates	
01janv1999	15mars2003
R ²	87,90%

Regression	
a	b
1,000	2,000
95% IQR	4,000
95% MAE	2,000

95% IQR
 The 95% IQR corresponds to the difference between the 3rd and the 1st quartiles of the error distribution for the 5% most extreme readings.

95% MAE
 The 95% MAE corresponds to the mean of the absolute error for the 5% most extreme readings.



GENERAL STATION INFORMATION
SERIE 64 - METZ-FRESCATY (m/s)

STATION

Usual Name		Alternate Name	
METZ-FRESCATY		METZ - FRESCATY	
Country		Latitude	Longitude
France		49,082 °N	6,115 °E
Local ID	WMO ID	Altitude	
57039001	7090	190 m above sea level	

ELEMENT

Element	Measure Details	Precision
Maximum Daily Wind Speed	00-24 UT	Nearest Integer.
Distribution Threshold		Start Date
Above	18,0	01janv1970

DATA COMPLETENESS

The station Metz-Frescaty has less than 0.25% missing values. Data from Aéroport MNL was used after correlation analysis.

The final serie has missing values on 1Aug56, 24-25Jan57, 17Dec57, 10Oct60, 16-17Oct60 and 20Oct60. However, nearby stations reveal no extreme events around these periods.

AVAILABLE METADATA

Due to important heterogeneities in most French wind records prior to 1970, all time series have been truncated to Jan 1st, 1970. Contact CatiXL for more information.

The station first opened on 19291201. It was relocated on 19560531.

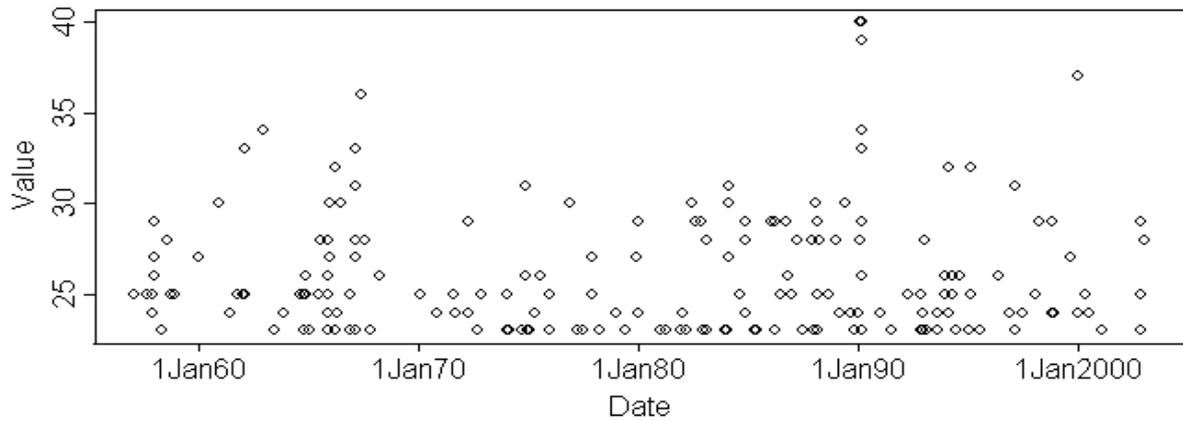
The first Wind recordings go back to 19490101.

GEOGRAPHICAL REFERENCE



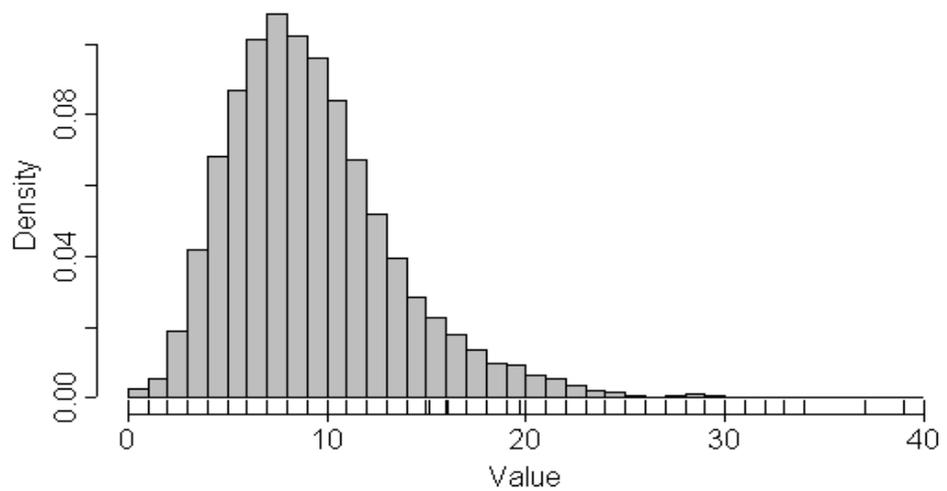
STATION STATISTICS
SERIE 64 - METZ-FRESCATY (m/s)

EXTREME HISTORICAL VALUES



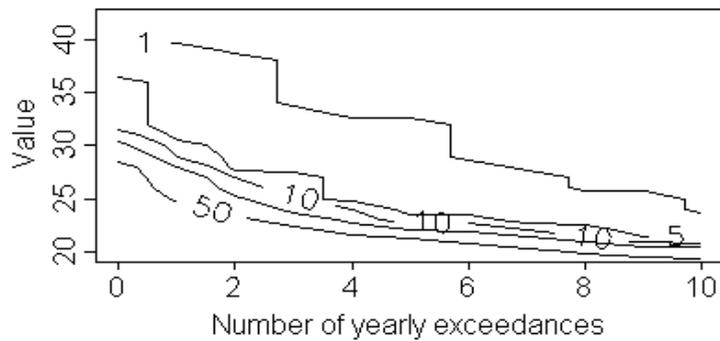
DATA DISTRIBUTION (1970 - 2000)

Quantile	Value
95,0%	18,00
97,5%	20,00
98,0%	21,00
98,5%	22,00
99,0%	23,00
99,5%	25,00
100,0%	40,00



PROBABILITY OF EXCEEDANCE (1970 - 2000)

	5%	10%	20%
2 days	27,75	27,00	25,33
4 days	24,75	24,00	22,71
6 days	23,50	22,67	21,88
8 days	22,50	21,75	21,00

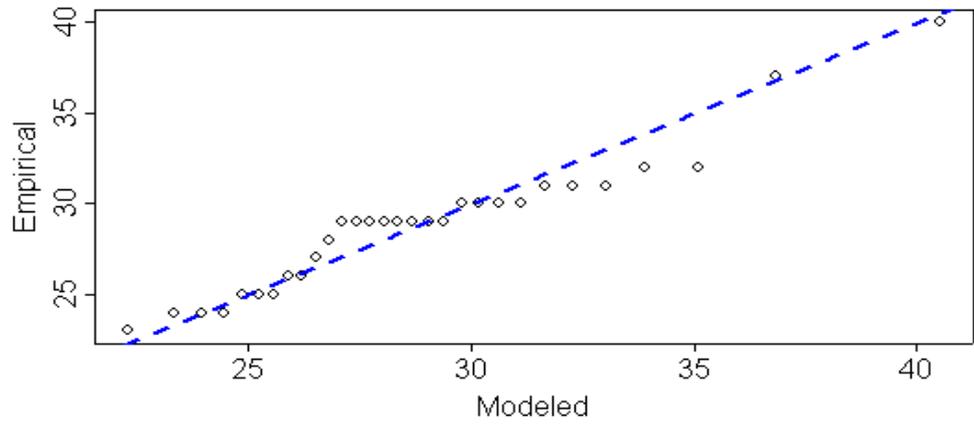


RETURN PERIOD ESTIMATION
SERIE 64 - METZ-FRESCATY (m/s)

GUMBEL RETURN PERIOD (1970 - 2000)

T (Years)	Value
2	28,20
5	31,96
10	34,45
15	35,86

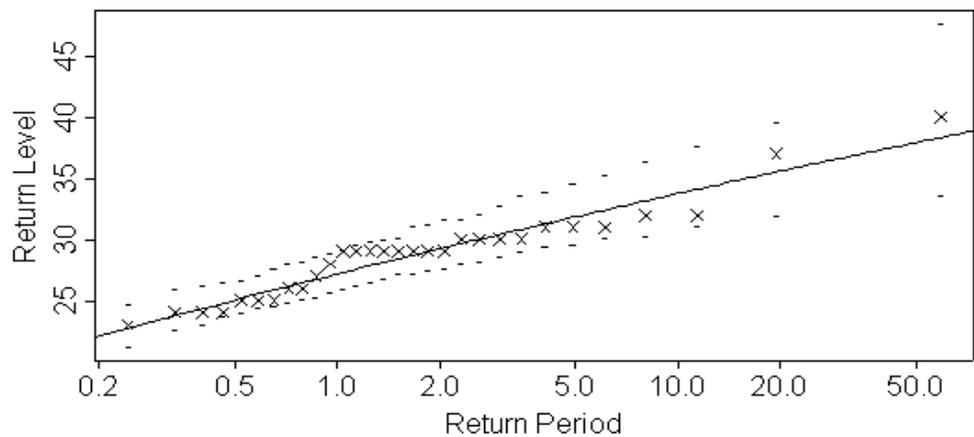
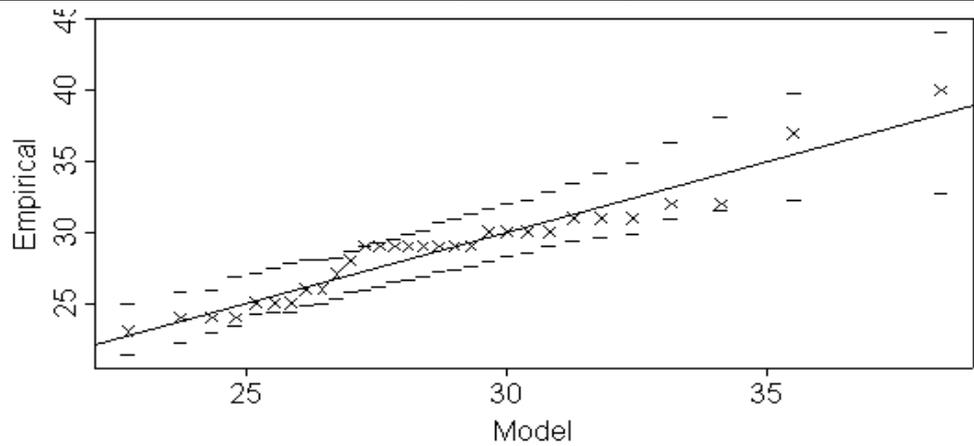
Parameters	
Location	26,980
Scale	3,319



GEV RETURN PERIOD (1970 - 2000)

T (Years)	Value
2	28,27
5	31,55
10	33,62
15	34,75

Parameters	
Location	27,170
Scale	3,031
Shape	-0,051



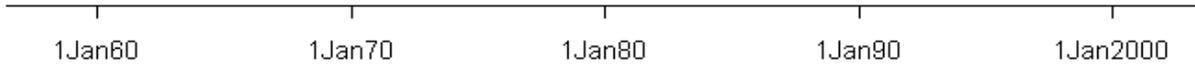
STATION FALLBACK INFORMATION
SERIE 64 - METZ-FRESCATY (m/s)

AVAILABLE SURROUNDING DATA

METZ.FRESCATY



AEROPORT.M.N.L.



PRINCIPAL FALLBACK
AEROPORT M.N.L.

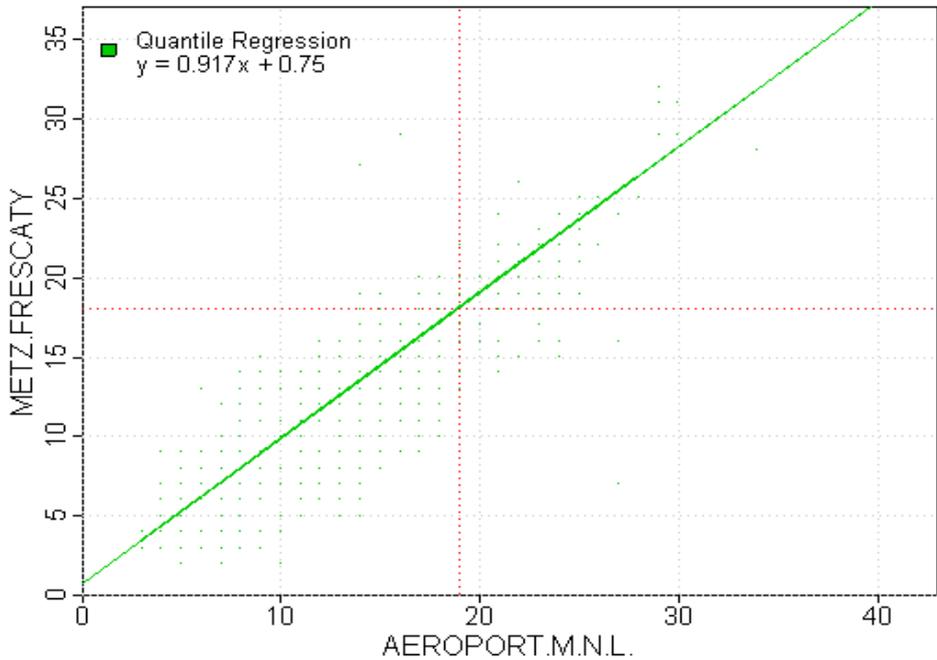
Station Name	
Local ID	WMO
57251001	

Correlation Dates	
01janv1995	30avr2003
R ²	88,70%

Regression	
a	b
0,917	0,750
95% IQR	2,040
95% MAE	1,395

95% IQR
 The 95% IQR corresponds to the difference between the 3rd and the 1st quartiles of the error distribution for the 5% most extreme readings.

95% MAE
 The 95% MAE corresponds to the mean of the absolute error for the 5% most extreme readings.



GENERAL STATION INFORMATION
SERIE 28 - NICE (m/s)

STATION

Usual Name		Alternate Name	
NICE		AEROPORT	
Country		Latitude	Longitude
France		43,650 °N	7,200 °E
Local ID	WMO ID	Altitude	
6088001	7690	4 m above sea level	

ELEMENT

Element	Measure Details	Precision
Maximum Daily Wind Speed	00-24 UT	Nearest Integer.
Distribution Threshold		Start Date
Above	19,0	01janv1970

DATA COMPLETENESS

The station Nice has less than 0.2% missing values. Data from Cap Ferrat was used after correlation analysis.

The final serie has missing values on 20Jul94, 29Jun-4Jul95 and 1-2Nov95; however, surrounding stations reveal no extreme events around these periods.

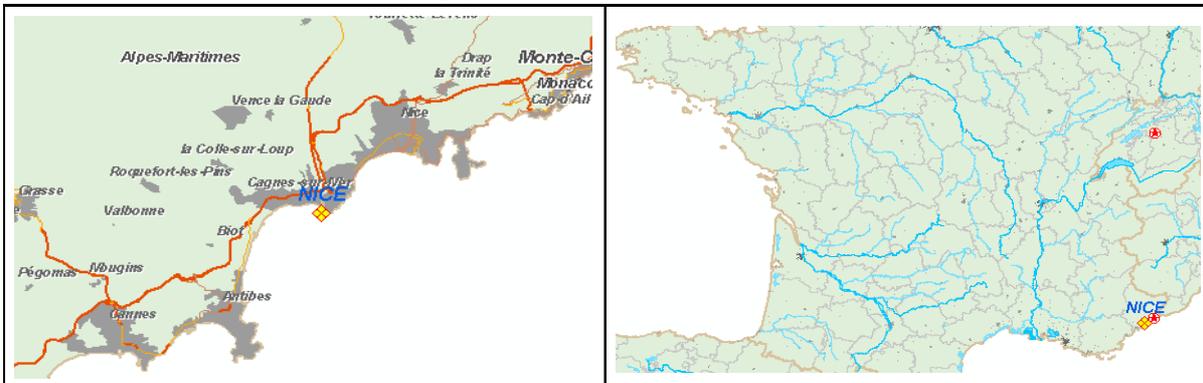
AVAILABLE METADATA

Due to important heterogeneities in most French wind records prior to 1970, all time series have been truncated to Jan 1st, 1970. Contact CatiXL for more information.

The station first opened on 19420701. It was relocated on 19421127; 19440104; 19440910; 19441104; 19460404; 19570425.

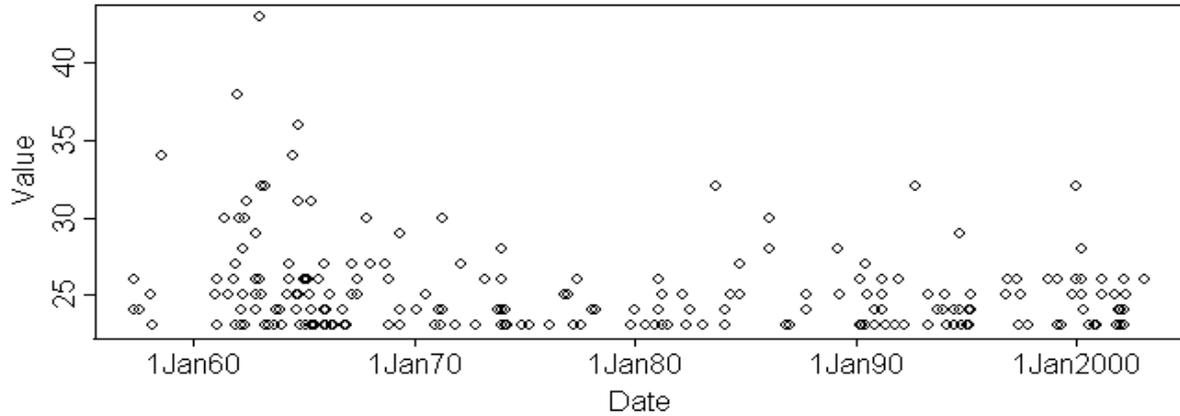
The first Wind recordings go back to 19490101.

GEOGRAPHICAL REFERENCE



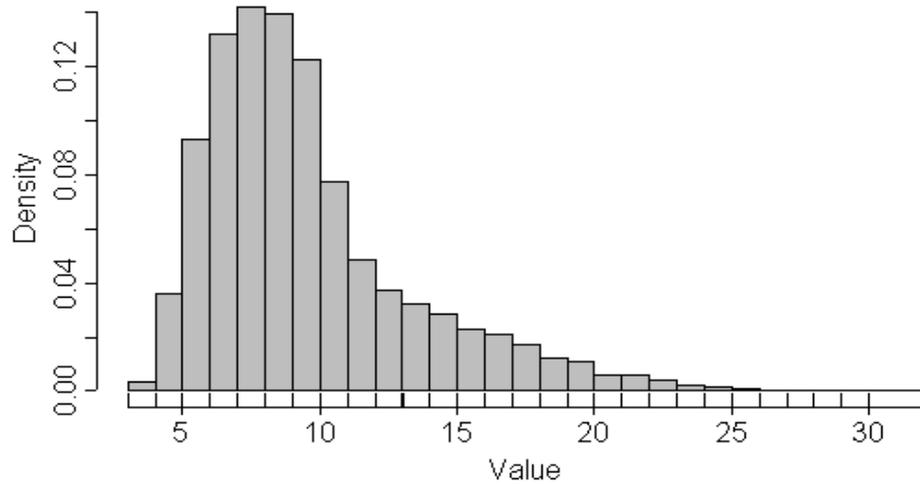
STATION STATISTICS
SERIE 28 - NICE (m/s)

EXTREME HISTORICAL VALUES



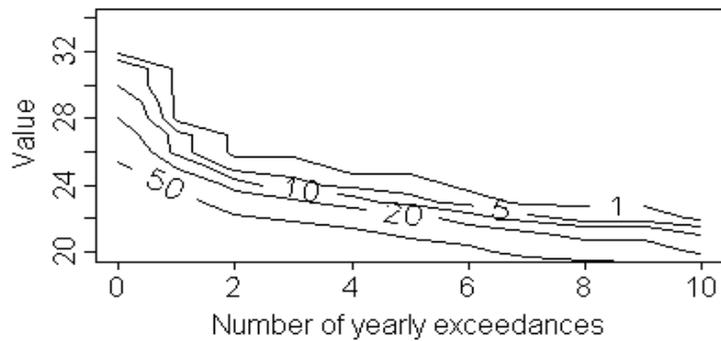
DATA DISTRIBUTION (1970 - 2000)

Quantile	Value
95,0%	18,00
97,5%	20,00
98,0%	21,00
98,5%	22,00
99,0%	22,00
99,5%	24,00
100,0%	32,00



PROBABILITY OF EXCEEDANCE (1970 - 2000)

	5%	10%	20%
2 days	24,83	24,33	23,67
4 days	23,83	23,33	22,67
6 days	22,83	22,33	21,67
8 days	21,88	21,50	20,75

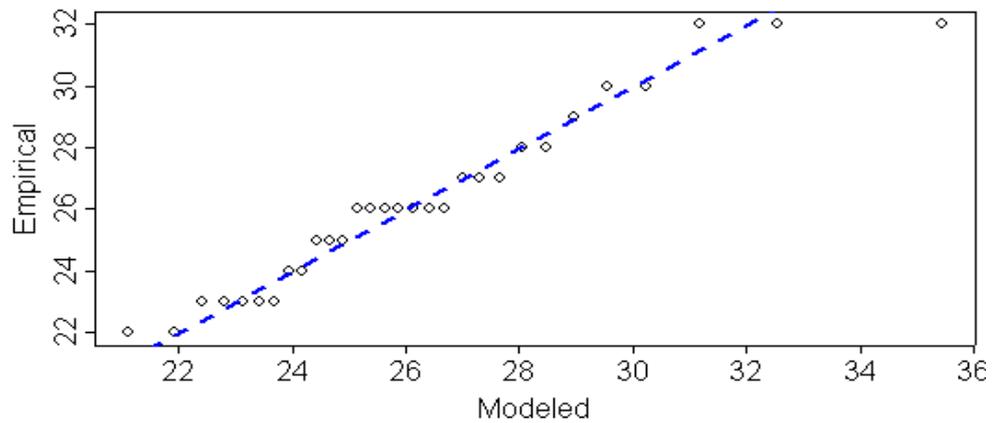


RETURN PERIOD ESTIMATION
SERIE 28 - NICE (m/s)

GUMBEL RETURN PERIOD (1970 - 2000)

T (Years)	Value
2	25,75
5	28,71
10	30,67
15	31,77

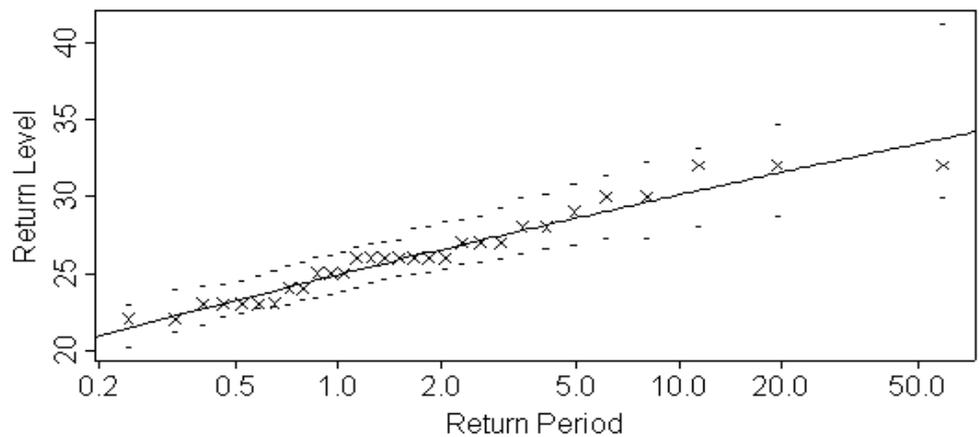
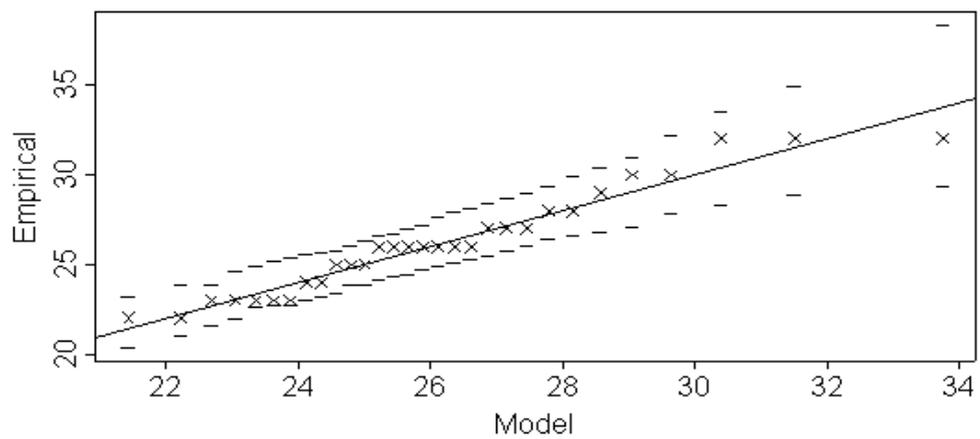
Parameters	
Location	24,795
Scale	2,609



GEV RETURN PERIOD (1970 - 2000)

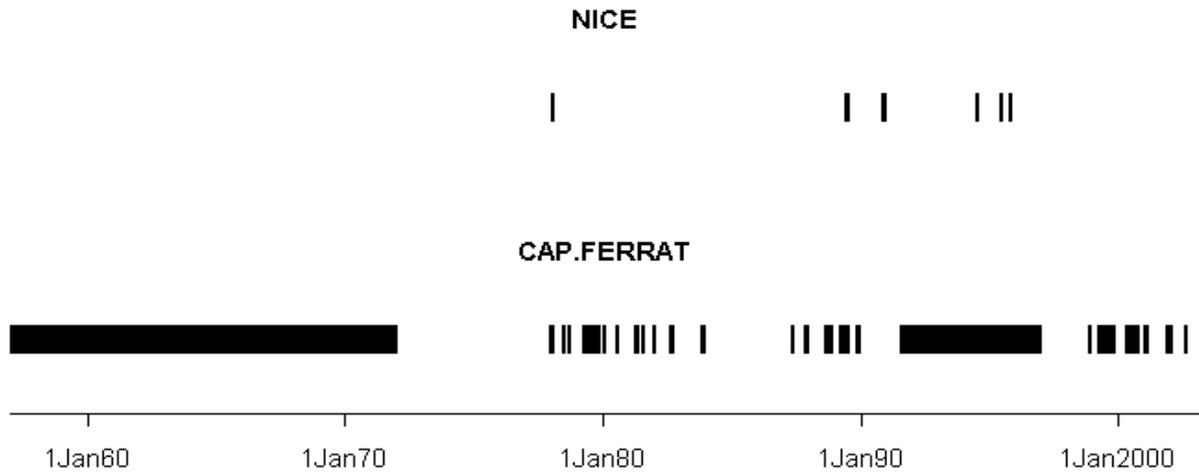
T (Years)	Value
2	25,78
5	28,36
10	29,99
15	30,89

Parameters	
Location	24,910
Scale	2,382
Shape	-0,048



STATION FALLBACK INFORMATION
SERIE 28 - NICE (m/s)

AVAILABLE SURROUNDING DATA



PRINCIPAL FALLBACK
CAP FERRAT

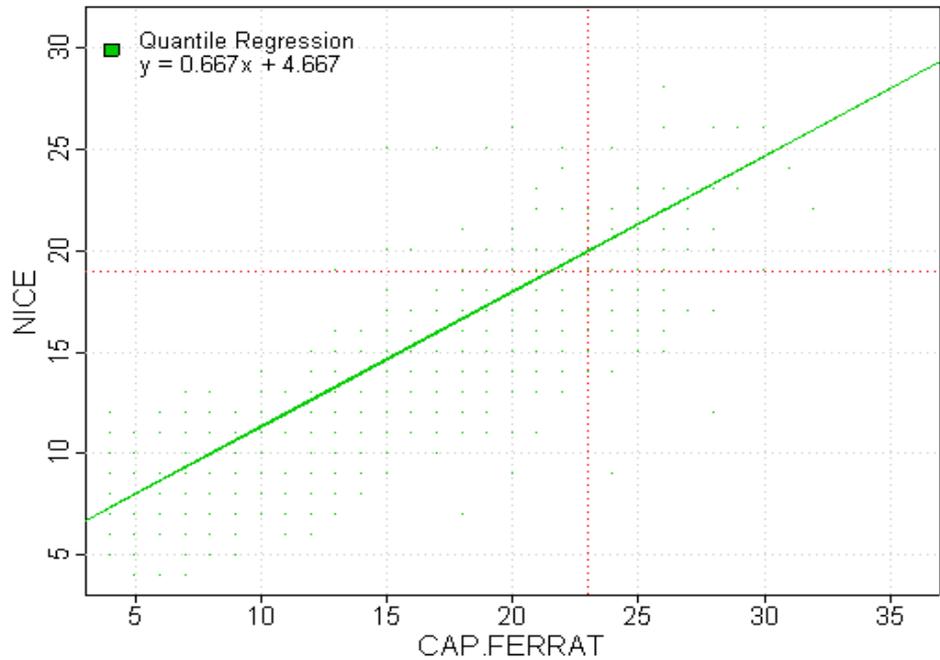
Station Name	
Local ID	WMO
06121001	07695

Correlation Dates	
01janv1998	15mars2003
R ²	86,60%

Regression	
a	b
0,667	4,667
95% IQR	2,745
95% MAE	1,806

95% IQR
 The 95% IQR corresponds to the difference between the 3rd and the 1st quartiles of the error distribution for the 5% most extreme readings.

95% MAE
 The 95% MAE corresponds to the mean of the absolute error for the 5% most extreme readings.



GENERAL STATION INFORMATION
SERIE 18 - ORLY (m/s)

STATION

Usual Name		Alternate Name	
ORLY		ORLY AEROPORT	
Country		Latitude	Longitude
France		48,717 °N	2,383 °E
Local ID	WMO ID	Altitude	
91027002	7149	89 m above sea level	

ELEMENT

Element	Measure Details	Precision
Maximum Daily Wind Speed	00-24 UT	Nearest Integer.
Distribution Threshold	Recording Units	Start Date
Above 20,0	m/s	01janv1970

DATA COMPLETENESS

The station Orly has less than 0.15% missing values. Data from Toussus-Le-Noble was used after correlation analysis.

The final serie has missing values on 20Nov49, 15Mar50, 10-17Feb54, 2Dec59, 20Jul61, 10Nov70. However surrounding stations reveal no extreme events around these periods.

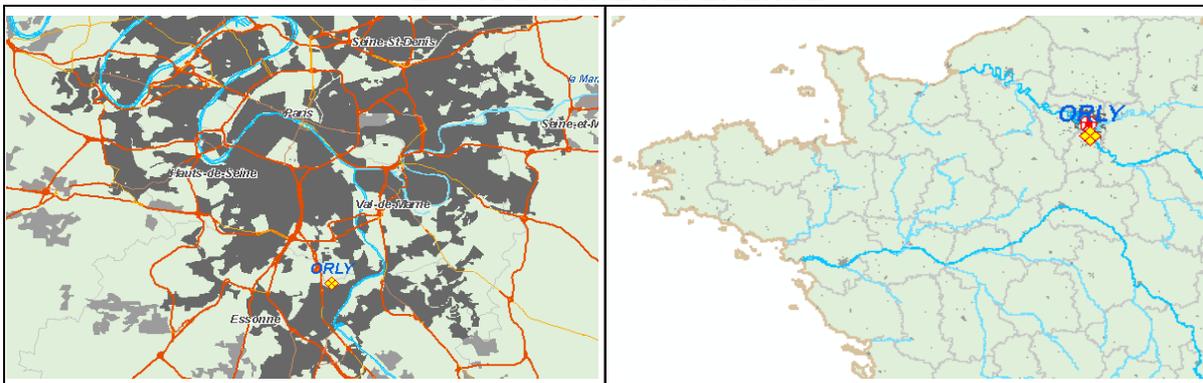
AVAILABLE METADATA

Due to important heterogeneities in most French wind records prior to 1970, all time series have been truncated to Jan 1st, 1970. Contact CatiXL for more information.

The station first opened on 19210301. It was relocated on 19930630.

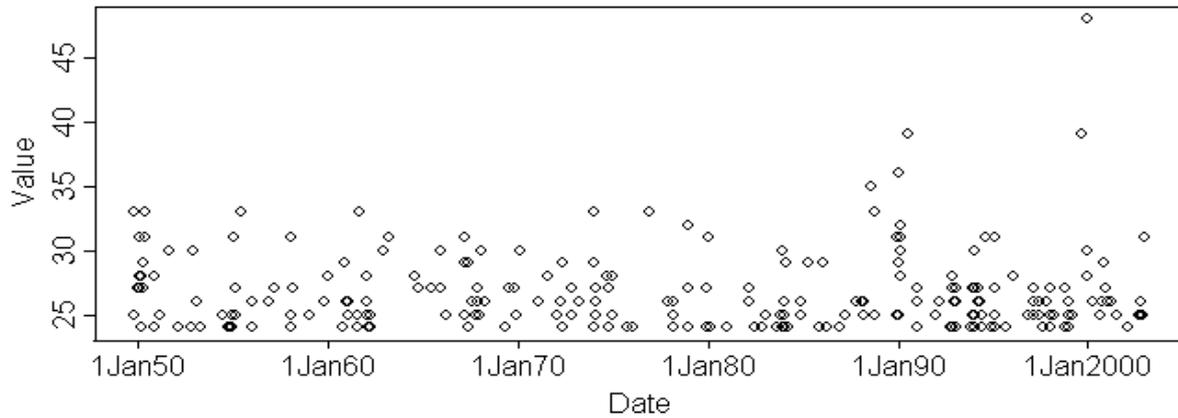
The first Wind recordings go back to 19490102.

GEOGRAPHICAL REFERENCE



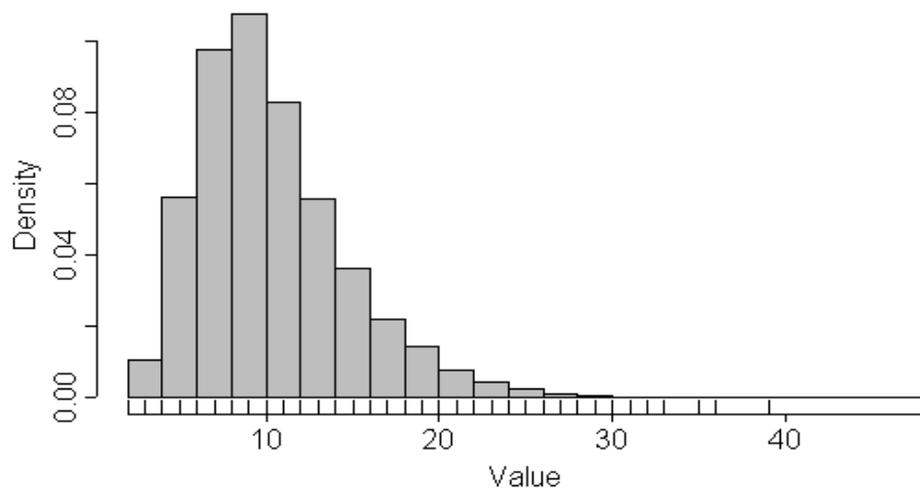
STATION STATISTICS
SERIE 18 - ORLY (m/s)

EXTREME HISTORICAL VALUES



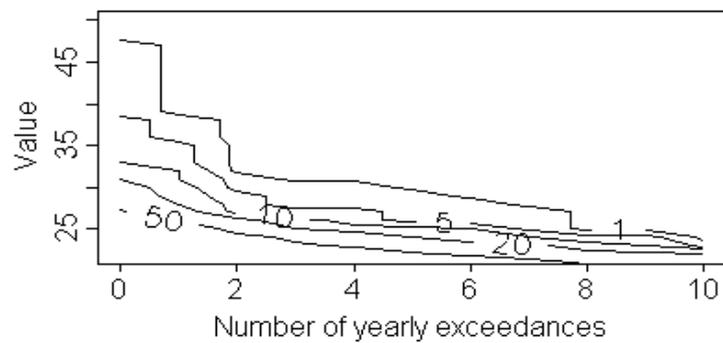
DATA DISTRIBUTION (1970 - 2000)

Quantile	Value
95,0%	19,00
97,5%	22,00
98,0%	22,00
98,5%	23,00
99,0%	24,00
99,5%	26,00
100,0%	48,00



PROBABILITY OF EXCEEDANCE (1970 - 2000)

	5%	10%	20%
2 days	29,50	26,80	26,20
4 days	27,50	25,50	24,60
6 days	25,75	25,00	23,50
8 days	24,25	23,50	22,50

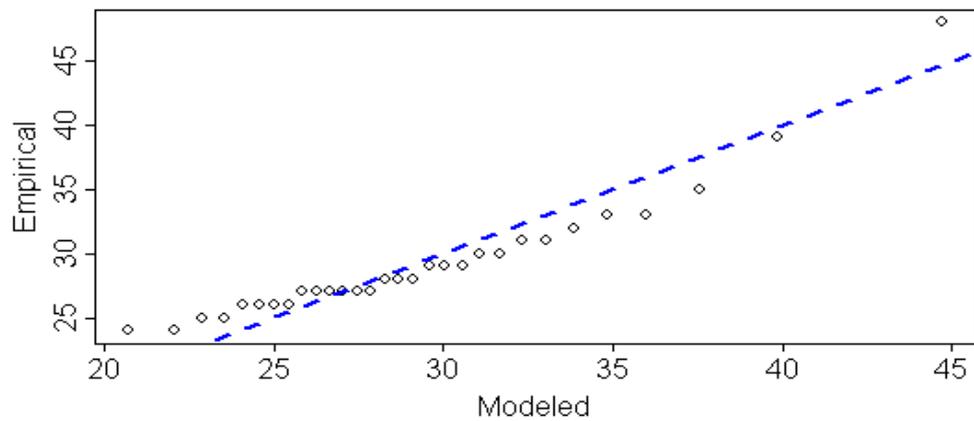


RETURN PERIOD ESTIMATION
SERIE 18 - ORLY (m/s)

GUMBEL RETURN PERIOD (1970 - 2000)

T (Years)	Value
2	28,48
5	33,43
10	36,71
15	38,56

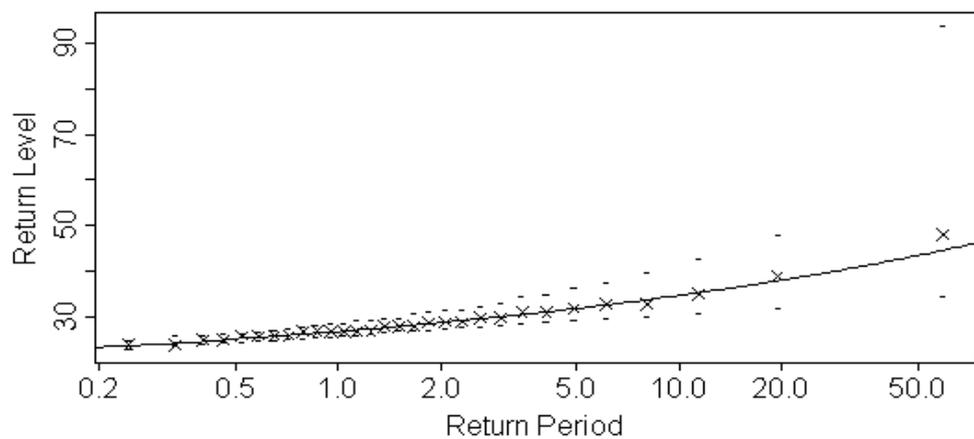
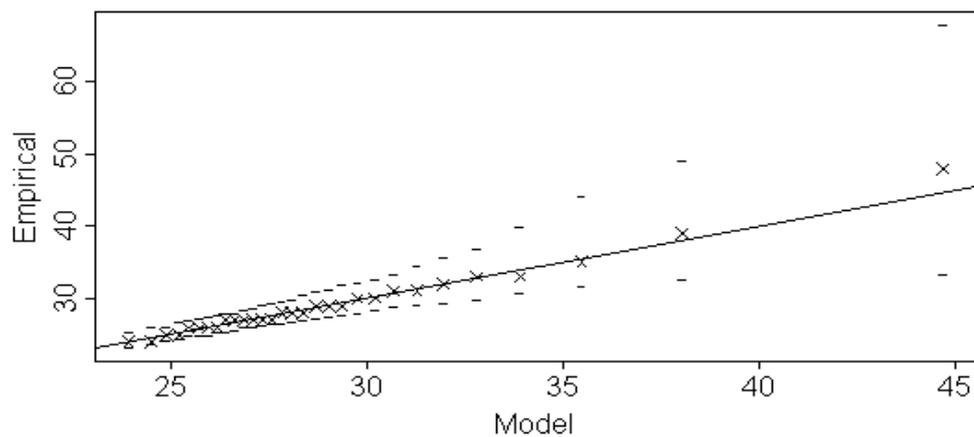
Parameters	
Location	26,882
Scale	4,367



GEV RETURN PERIOD (1970 - 2000)

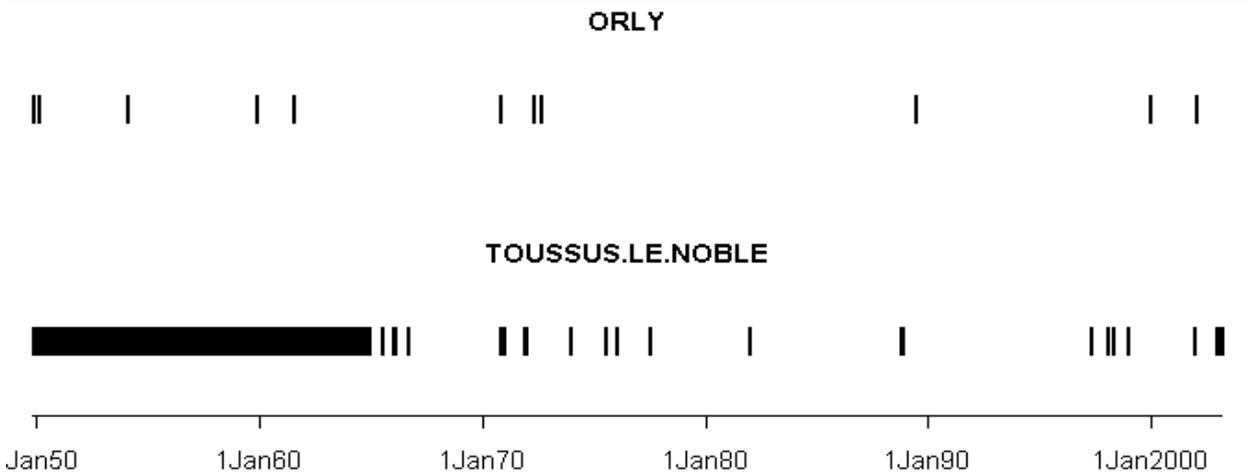
T (Years)	Value
2	27,97
5	31,61
10	34,62
15	36,57

Parameters	
Location	26,991
Scale	2,561
Shape	0,238



STATION FALLBACK INFORMATION
SERIE 18 - ORLY (m/s)

AVAILABLE SURROUNDING DATA



PRINCIPAL FALLBACK
TOUSSUS LE NOBLE

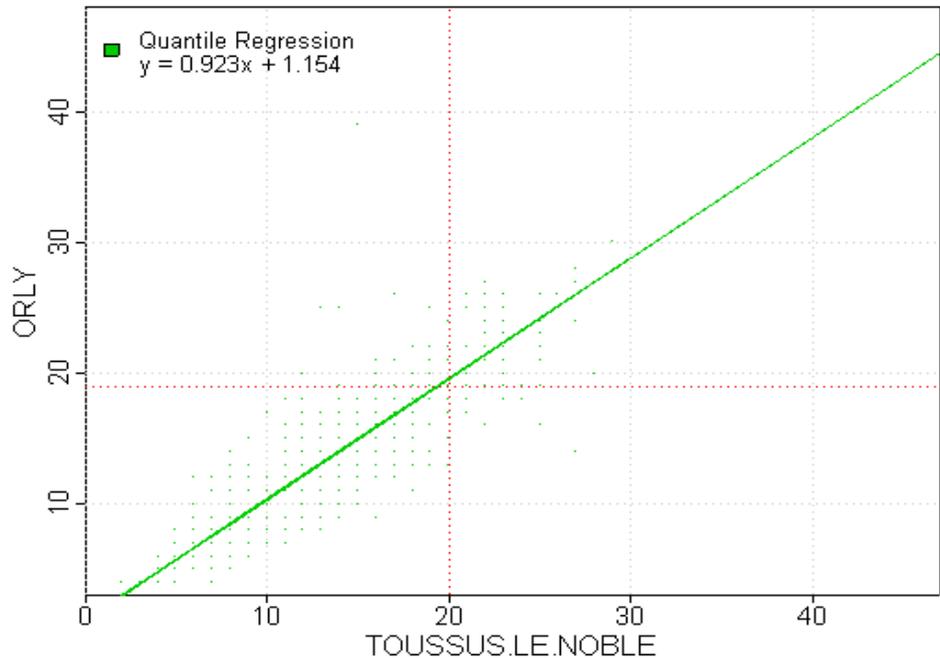
Station Name	
Local ID	WMO
78620001	

Correlation Dates	
01janv1998	31déc2002
R ²	90,90%

Regression	
a	b
0,923	1,154
95% IQR	2,150
95% MAE	1,703

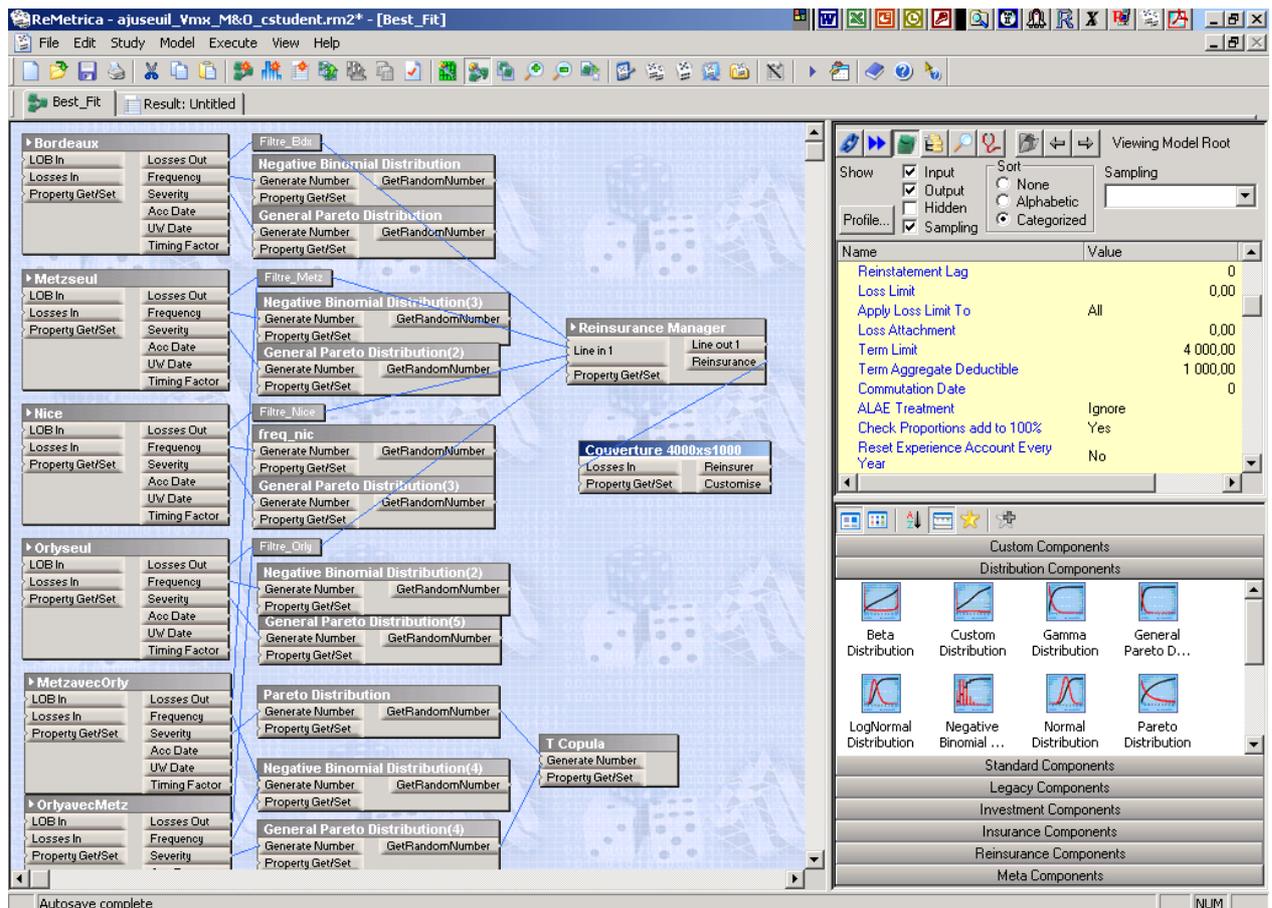
95% IQR
 The 95% IQR corresponds to the difference between the 3rd and the 1st quartiles of the error distribution for the 5% most extreme readings.

95% MAE
 The 95% MAE corresponds to the mean of the absolute error for the 5% most extreme readings.



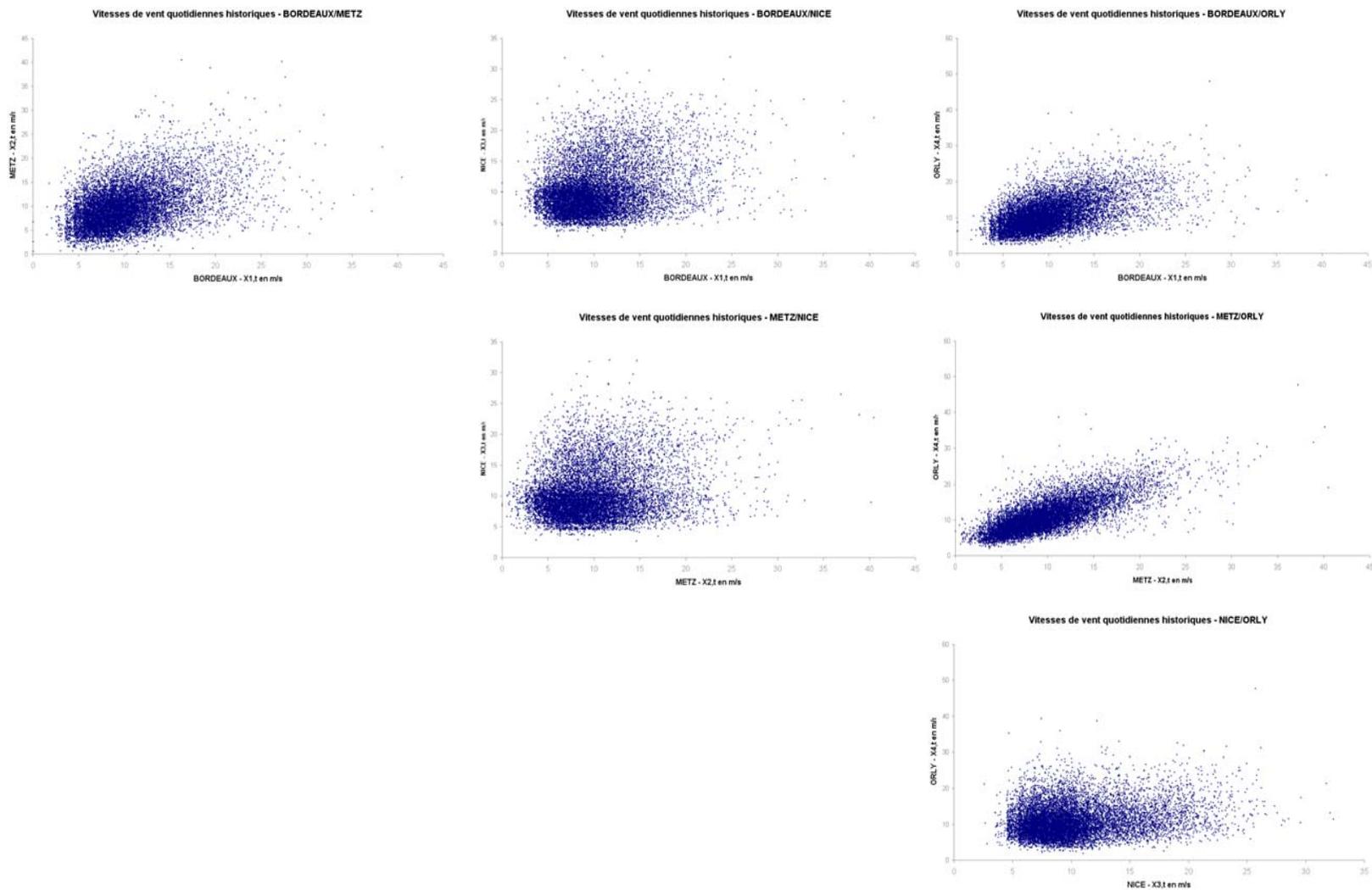
standards de distribution de probabilité (lois gamma, lognormale, poisson, etc.) et de copules (copule de Gumbel, copule de Student). Certains ont été programmés sous Visual Basic spécialement pour l'occasion (loi GEV notamment).

L'exemple d'un modèle utilisé pour le mémoire est présenté ci-dessous :

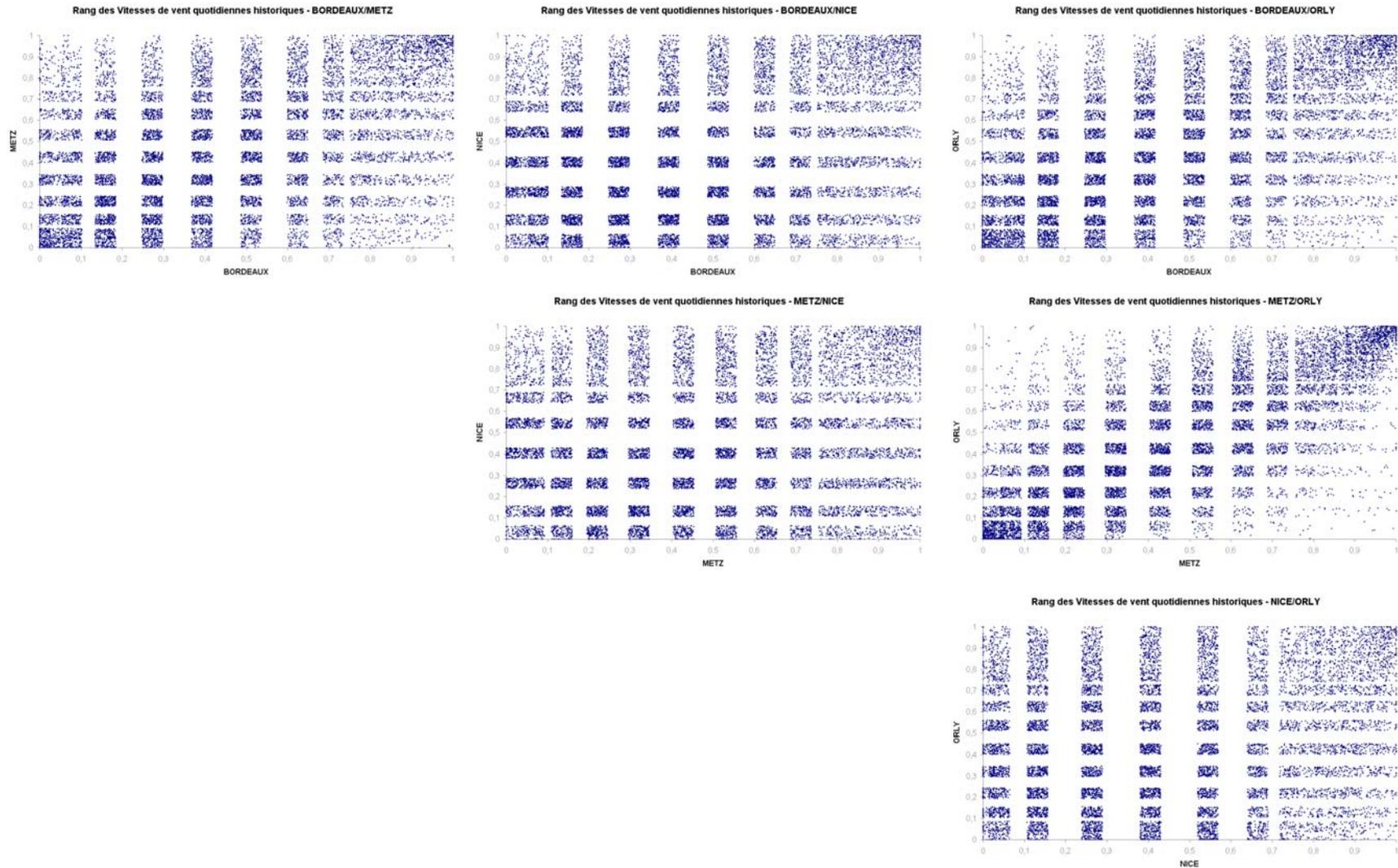


Annexe 4 : Nuages des vitesses de vent et QQ Plots

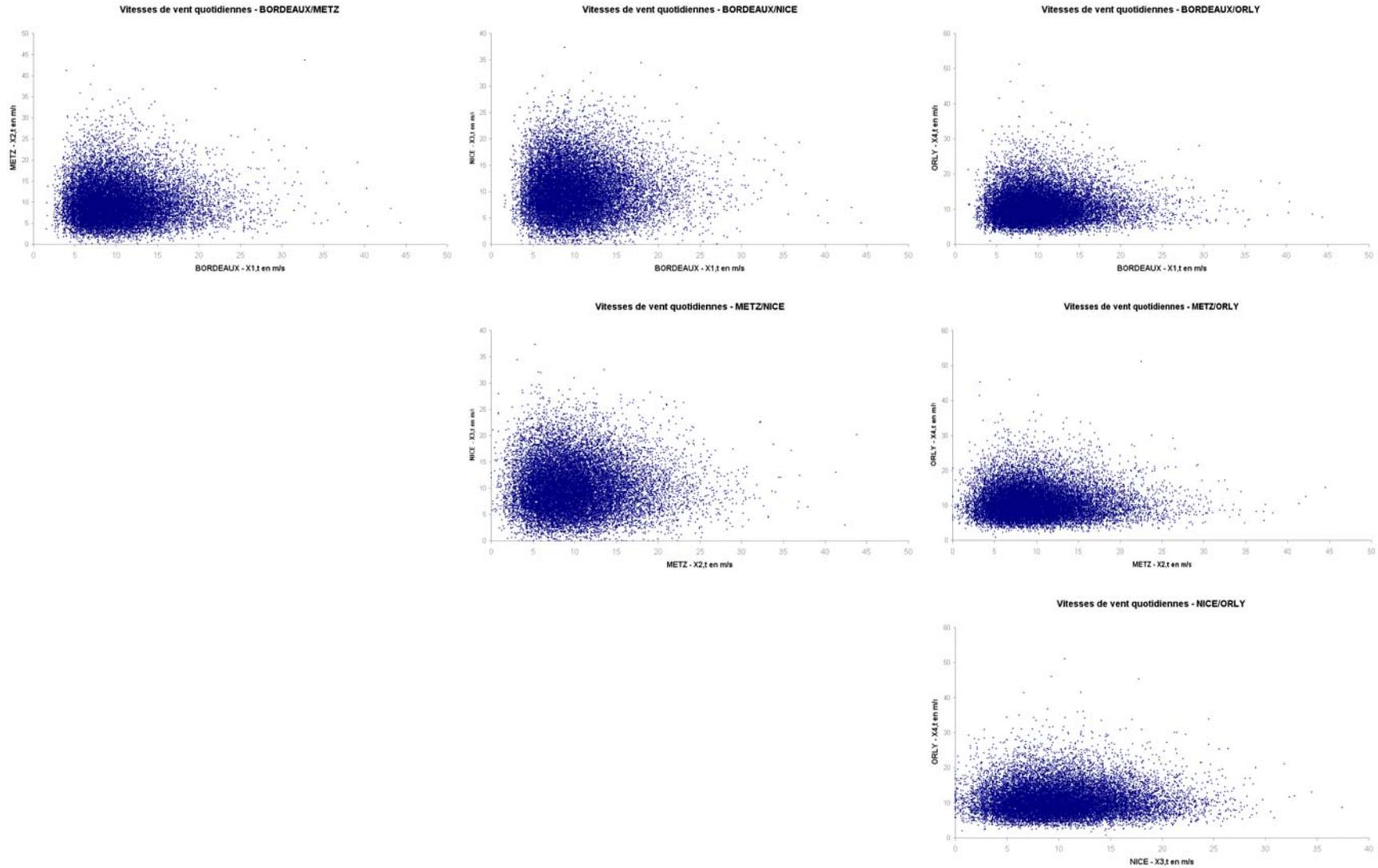
- Vitesses de vent historiques – Couples $(X_{i,t}, X_{j,t})$:



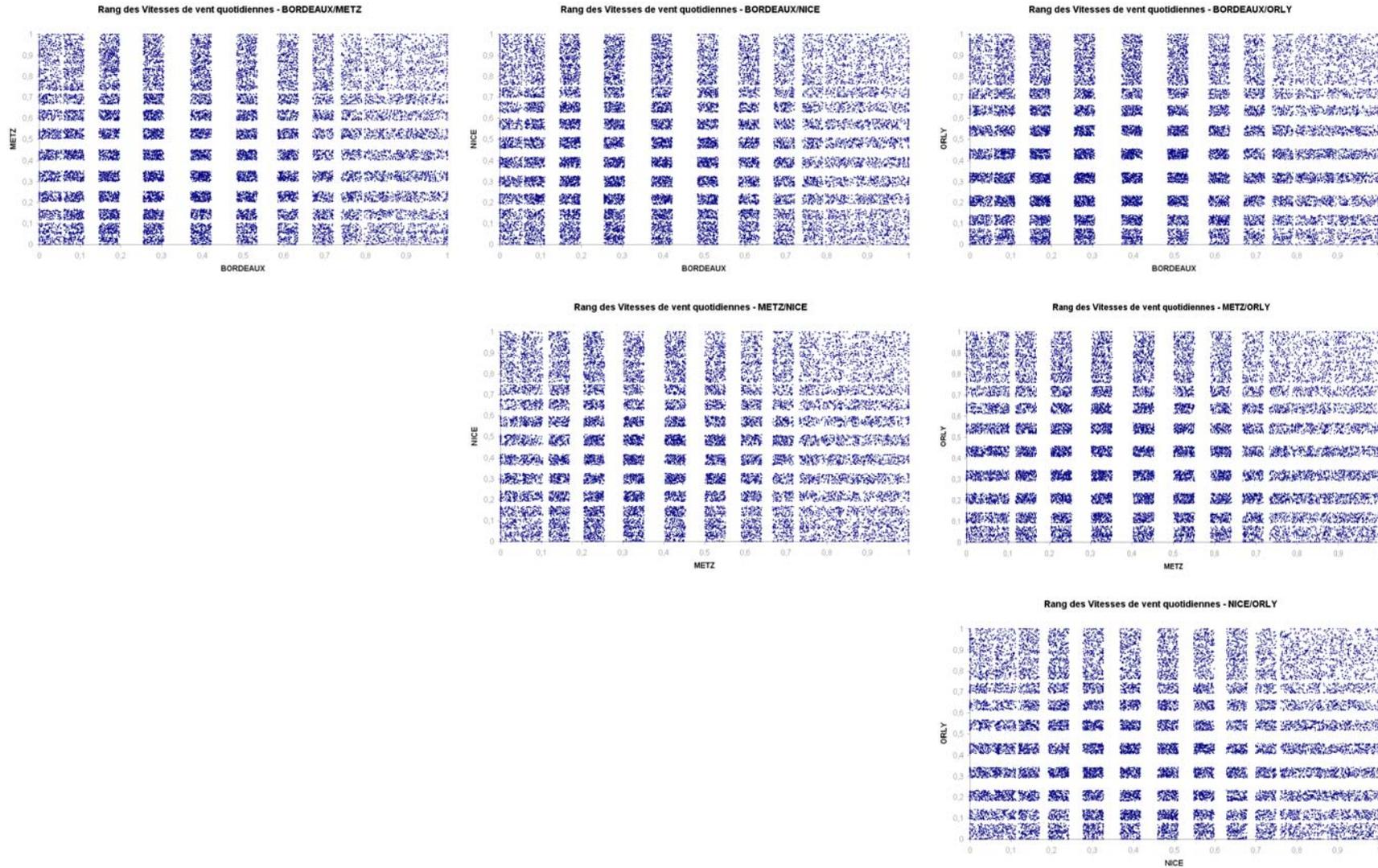
- Q-Q plots historiques – Couples $(X_{i,t}, X_{j,t})$:



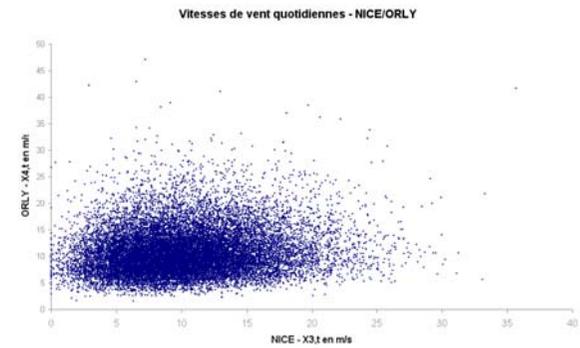
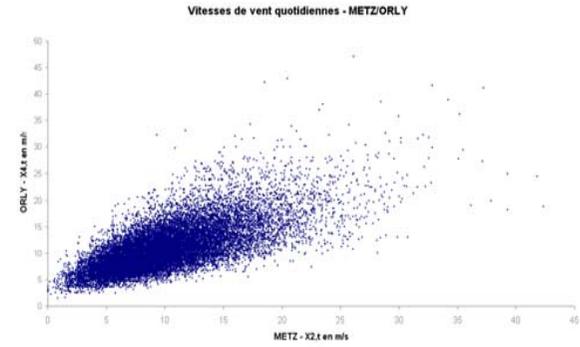
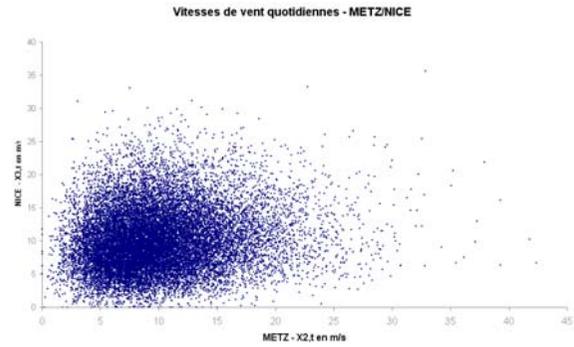
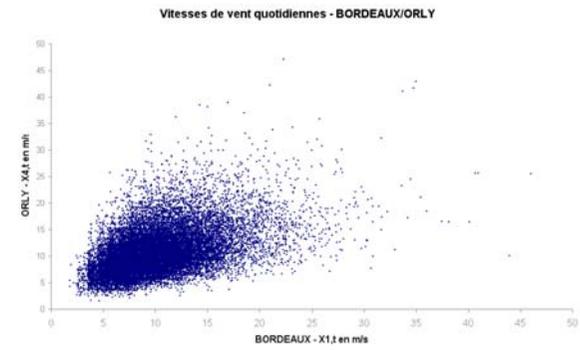
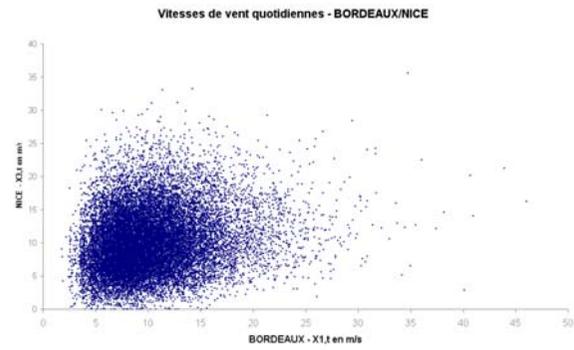
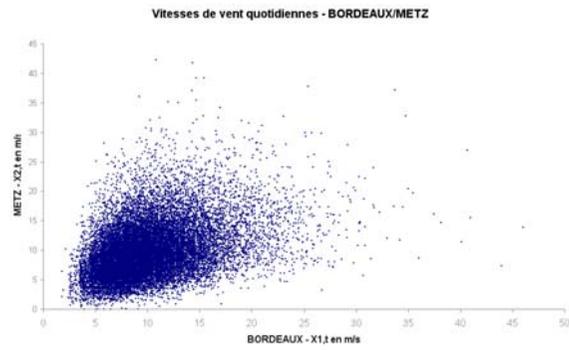
- Vitesses de vent des Couples ($X_{i,t}$, $X_{j,t}$) : « modèle final » sans prise en compte de la dépendance :



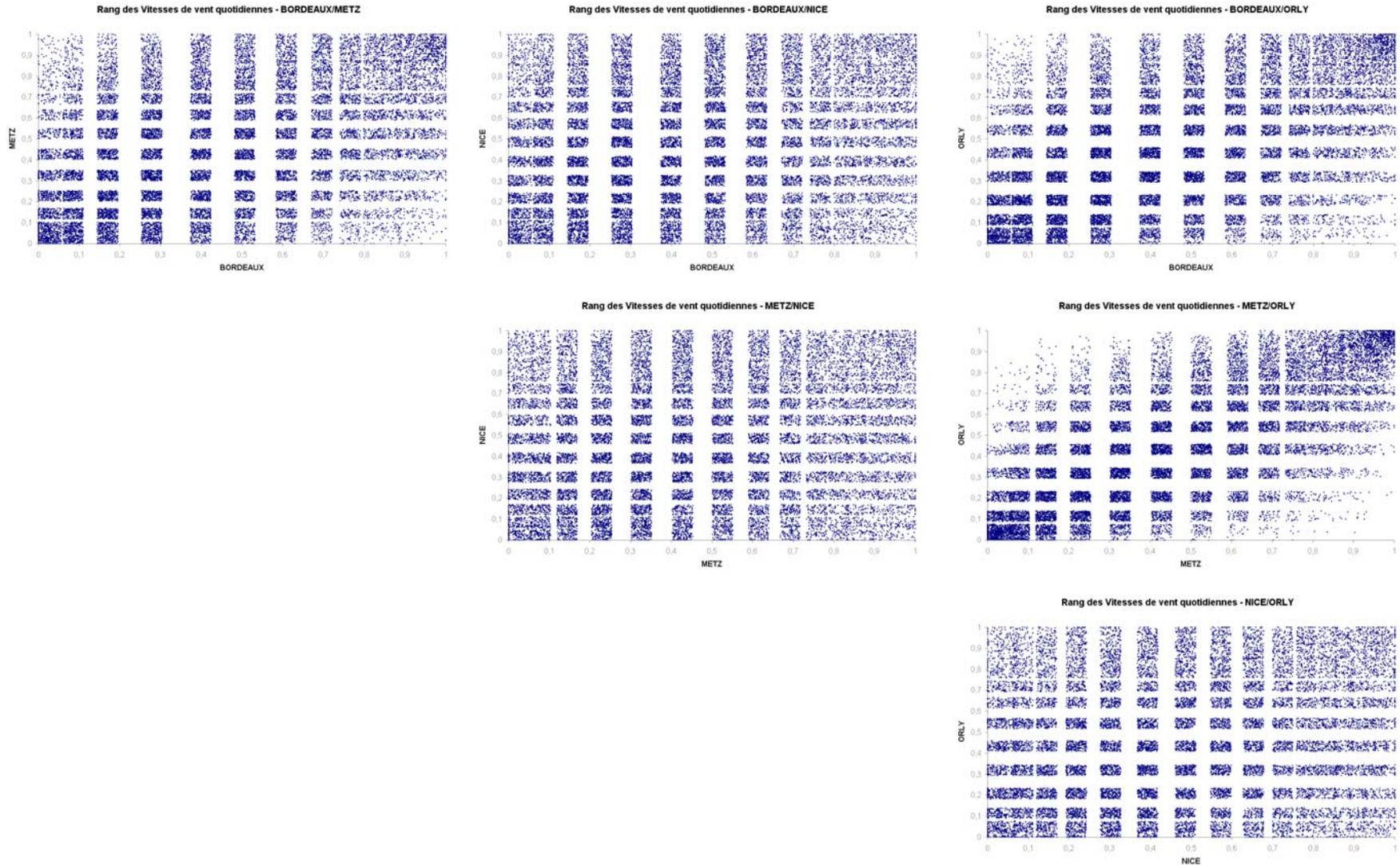
- Q-Q plots des Couples $(X_{i,t}, X_{j,t})$: « modèle final » sans prise en compte de la dépendance :



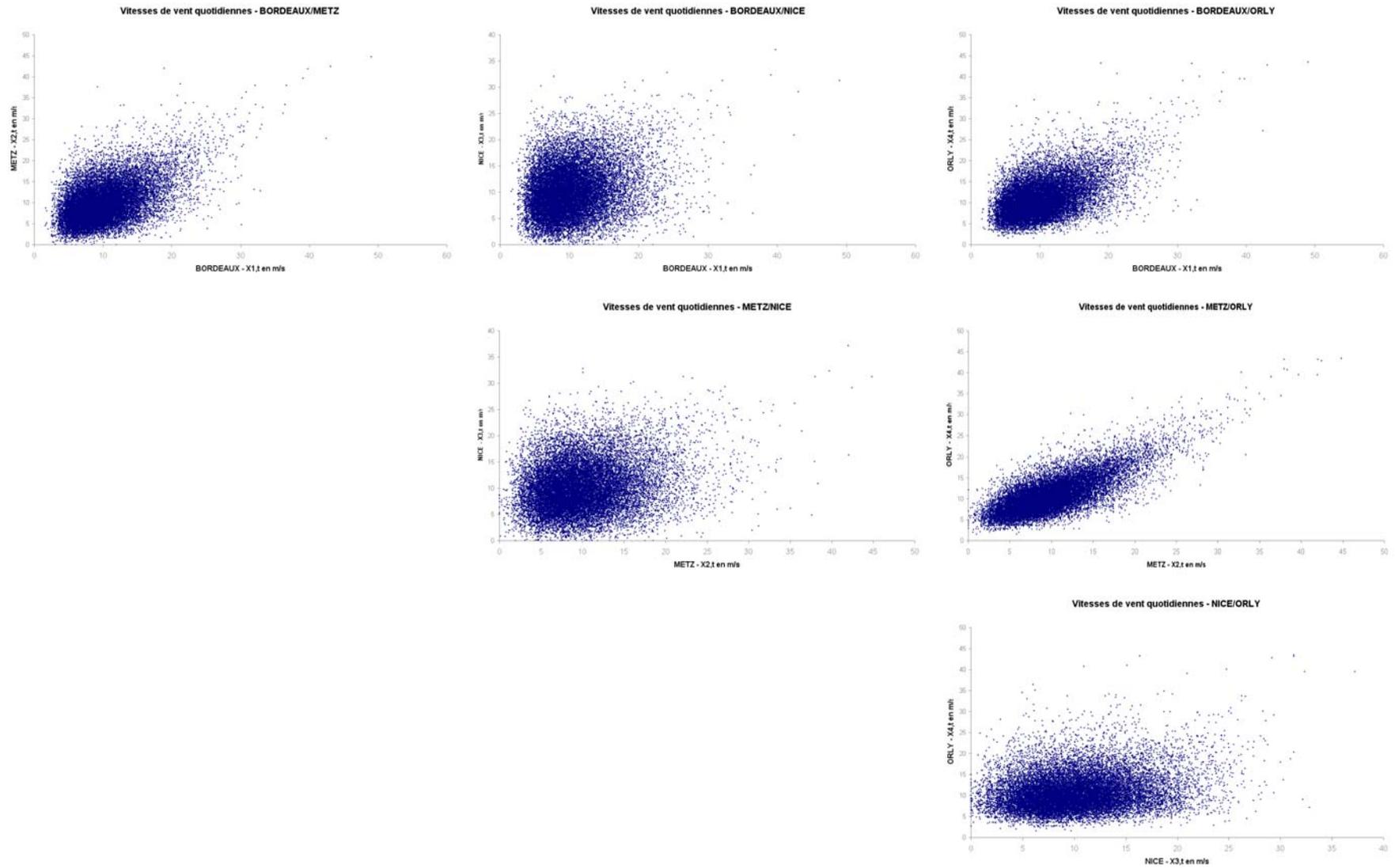
▪ Vitesses de vent des Couples ($X_{i,t}$, $X_{j,t}$) : simulation des corrélations :



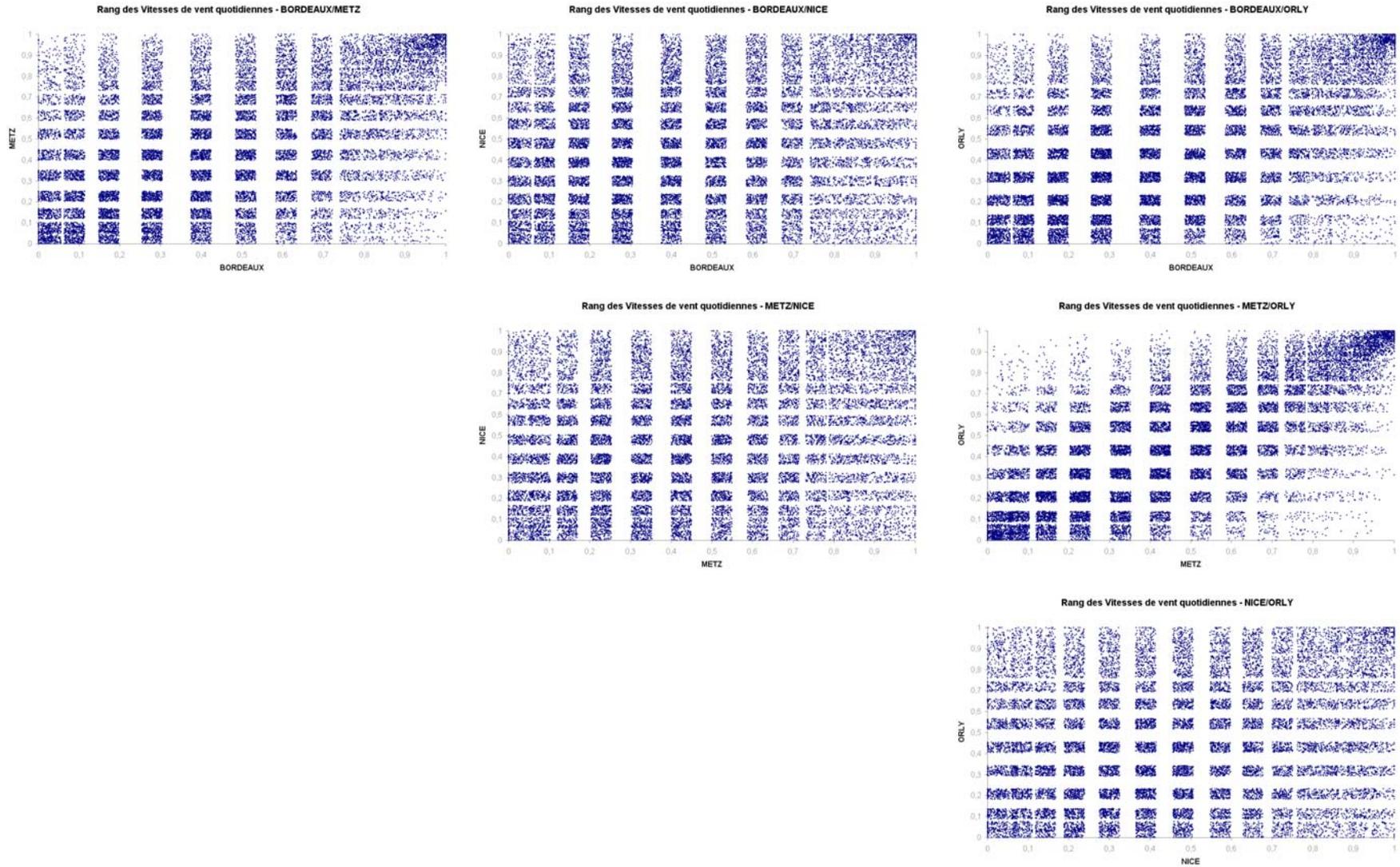
- Q-Q plots des Couples $(X_{i,t}, X_{j,t})$: simulation des corrélations :



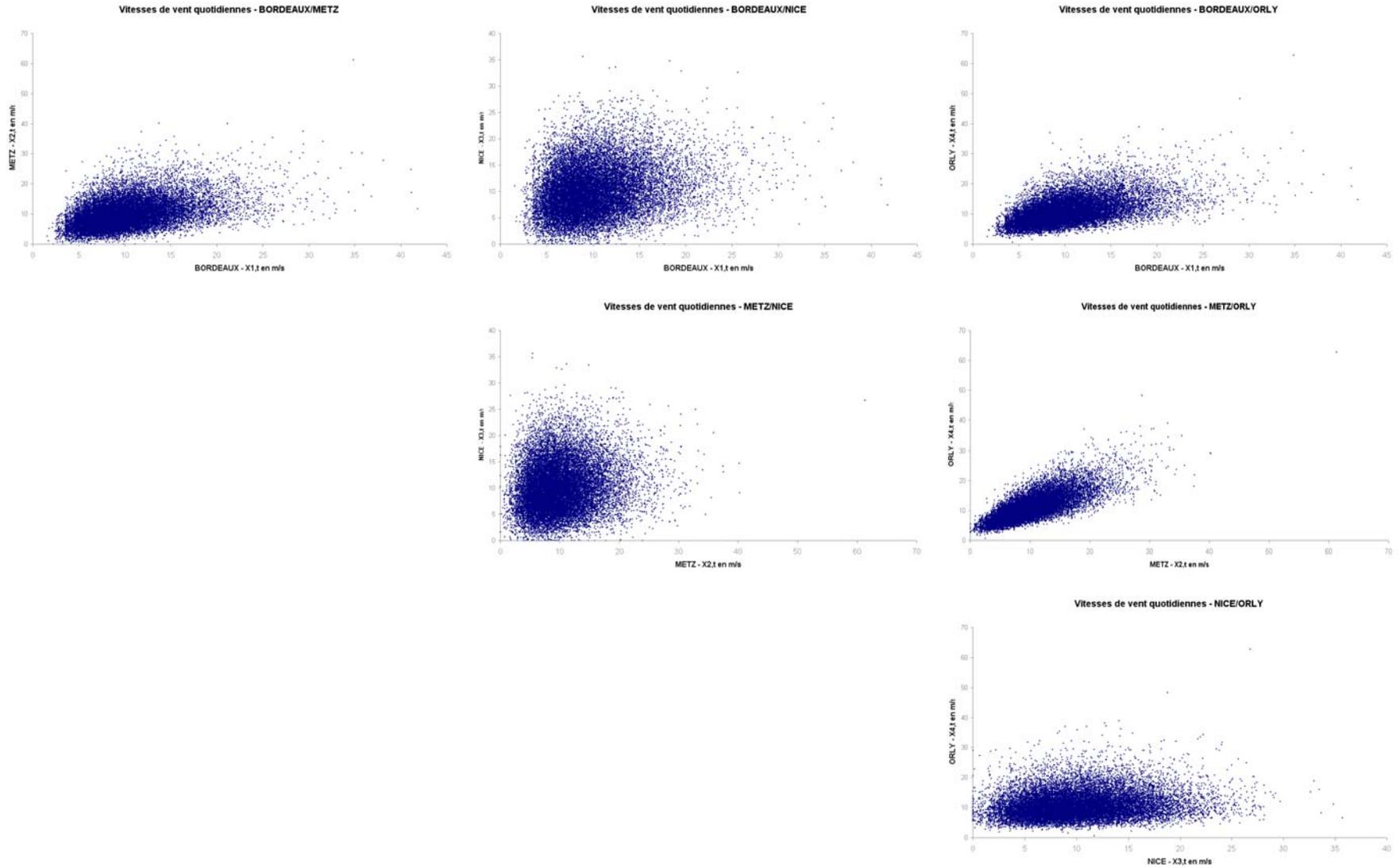
- Vitesses de vent des Couples ($X_{i,t}$, $X_{j,t}$) : simulation avec la copule de Gumbel :



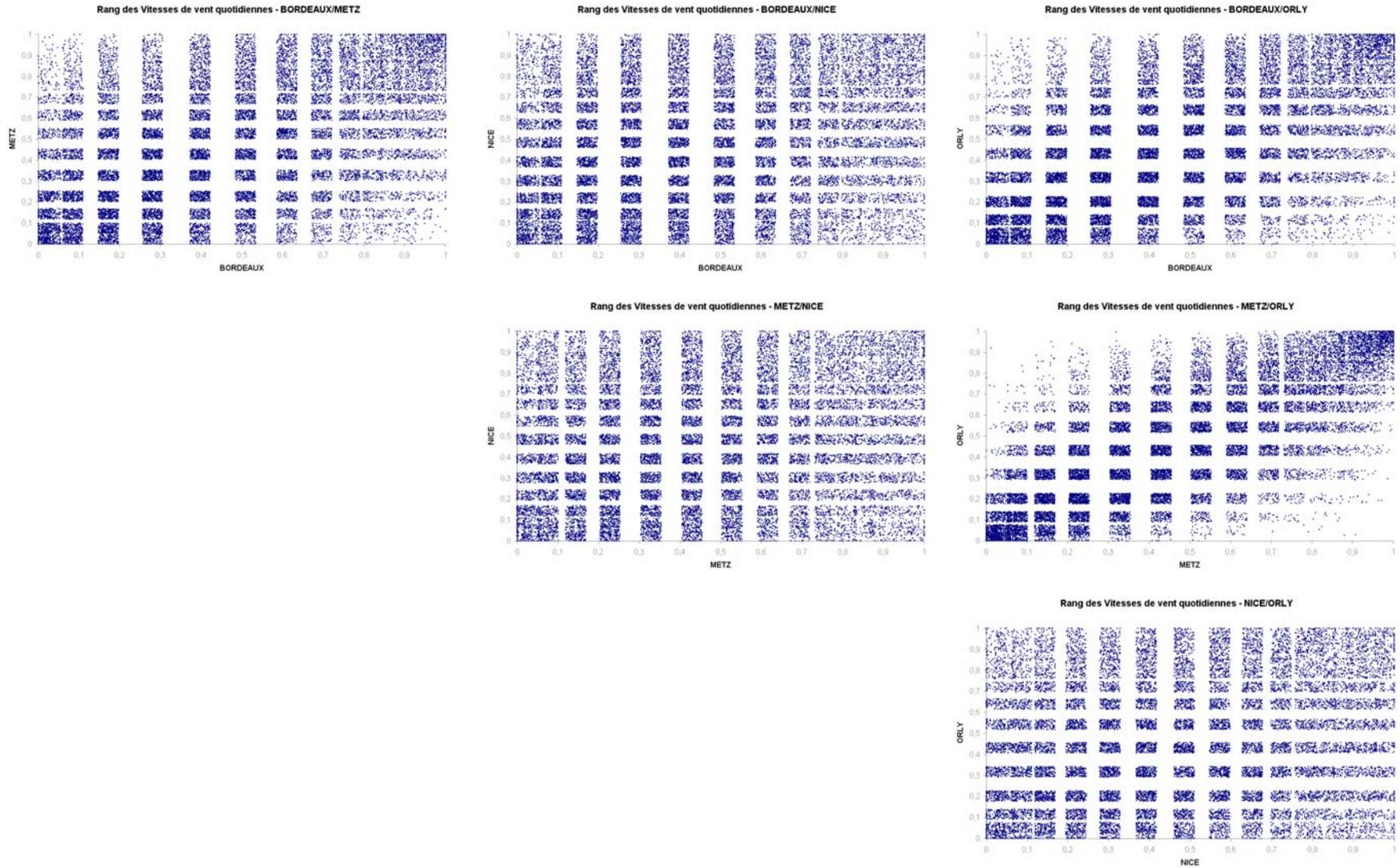
- Q-Q plots des Couples $(X_{i,t}, X_{j,t})$: simulation avec la copule de Gumbel :



- Vitesses de vent des Couples ($X_{i,t}$, $X_{j,t}$) : simulation avec la copule de Student :



- Q-Q plots des Couples $(X_{i,t}, X_{j,t})$: simulation avec la copule de Student :



Annexe 5 : Procédure de calcul des valeurs z_i sous R

La fonction R suivante permet d'obtenir le vecteur des valeurs z_i que l'on peut ensuite utiliser dans le calcul empirique de la fonction $K(z)$.

```
calczi<-function(x,y)
{
z<-x
for(i in 1:length(x)){
val1<-x[i]
val2<-y[i]
z[i]<-0
for(j in 1:length(x)){
if (x[j]<=val1){
if (y[j]<=val2){
z[i]<-z[i]+1
}}}}
return(z)
}
```