



Mémoire présenté
devant l'Institut de Science Financière et d'Assurances
pour l'obtention du diplôme d'Actuaire de l'Université de Lyon

le mercredi 18 septembre 2013

Par : Christophe DENISE-BAILLON

Titre: Modélisation statistique du risque de tempêtes tropicales dans le bassin
Atlantique Nord

Confidentialité : NON OUI (Durée : 1 an 2 ans)

Membres du jury de l'Institut des Actuaires

Olivier CAYOT

Gérald MONCHANIN

Entreprise :

Generali France

Membres du jury I.S.F.A.

M. Alexis BIENVENÛE

M. Areski COUSIN

Mme Diana DOROBANTU

Mme Anne EYRAUD-LOISEL

M. Nicolas LEBOISNE

M. Stéphane LOISEL

Mlle Esterina MASIELLO

Mme Véronique MAUME-DESCHAMPS

M. Frédéric PLANCHET

Mme Béatrice REY-FOURNIER

M. Pierre RIBEREAU

M. Christian-Yann ROBERT

M. Didier RULLIERE

M. Pierre THEROND

Directeur de mémoire en entreprise :

Romain BOYER CHAMMARD

Invité :

Autorisation de mise en ligne sur
un site de diffusion de documents
actuariels (après expiration de
l'éventuel délai de confidentialité)

Signature du responsable entreprise

Secrétariat

Mme Marie-Claude MOUCHON

Signature du candidat

Bibliothèque :

Mme Patricia BAROLO

*Je dédie ce travail à
la mémoire de ma mère
Françoise.*

*"Le simple est toujours faux.
Ce qui ne l'est pas est inutilisable."*

PAUL VALÉRY,
Œuvres, Tome II, 1942

Résumé

Pour les assureurs possédant un portefeuille dans les Antilles, les tempêtes tropicales et les cyclones constituent chaque année un risque majeur pouvant générer des dommages considérables.

Afin de répondre aux exigences de la nouvelle réforme européenne Solvabilité 2 en terme de ratio de couverture, les compagnies ayant opté pour un modèle interne doivent s'interroger quant à l'élaboration d'un outil de modélisation de ce risque.

Dans ce mémoire nous nous intéressons à une modélisation dite statistique couvrant le bassin Atlantique Nord avec pour objectif le calcul d'un capital de solvabilité (SCR) lié au risque de tempête. L'approche retenue se décompose en plusieurs étapes. En premier lieu les trajectoires des tempêtes sont simulées *via* une chaîne de Markov discrète et homogène. Puis, les caractéristiques des tempêtes sont associées aux trajectoires par des modèles de régression multiple. Enfin, un modèle de vulnérabilité assure la liaison entre les vitesses de vents atteignant une zone d'intérêt et les taux de dommages.

Mots-clefs :

Tempêtes tropicales, cyclones, ouragans, chaîne de Markov, modèle de régression multiple, modèle de désintégration, courbe de vulnérabilité, SCR

Abstract

Tropical storms and hurricanes are among the most costly hazard for Caribbean insurers.

In order to meet the Solvency 2 directive's capital requirement, european insurers applying for an internal model should focus on the development of a hurricane-modelling tool.

In this study we describe a statistical modelling of the North Atlantic area in order to compute the windstorm capital requirement (windstorm SCR). The calculation is split in several stages. Firstly, tropical storms tracks are simulated as a homogeneous discrete-time Markov chain. Then, intensity models are applied along these tracks using multiple linear regressions. Lastly, a vulnerability model is used to link wind speeds and damage ratios of buildings insured.

Keywords :

Tropical storms, hurricane, Markov's chains, multiple regression model, decay model, vulnerability curb, SCR

Avant-propos

Pour des raisons de confidentialité, les valeurs numériques résultant de l'activité de Generali France ont été normées. Néanmoins, les ordres de grandeurs ne s'en trouvent pas modifiées.

On précise aussi que toutes les modélisations présentes dans ce mémoire ont été réalisées à partir du logiciel libre de statistique R¹.

1. Page de téléchargement : <http://cran.r-project.org/bin/windows/base/>

Remerciements

Je tiens avant tout à remercier Generali France pour son accueil et toutes les personnes appartenant à la Direction Technique et des Risques qui, de près ou de loin, ont contribué à faire de cette année d'apprentissage une formidable expérience professionnelle dans une ambiance chaleureuse.

J'aimerais aussi adresser ma reconnaissance et ma gratitude à toute l'équipe du département Solvabilité et Modélisation des Risques et plus particulièrement à :

- Romain BOYER CHAMMARD, mon maître de stage, pour sa disponibilité, sa pédagogie et sa bonne humeur,
- Sylvestre FREZAL, mon manager, pour ses conseils et la confiance qu'il m'a accordé durant tout mon apprentissage,
- Christophe MOVREL et KhouLOUD MANDHOUI mes collègues de l'équipe non-vie du modèle interne.

J'adresse aussi mes remerciements à Stéphane LOISEL pour son suivi pédagogique.

Enfin, je remercie mon père, Gilles, et ma soeur, Estelle, pour leur indéfectible soutien tant dans mes études que dans les moments difficiles.

Sommaire

Introduction	2
I Généralités	3
1 Cadre, contexte et objectif de l'étude	4
1.1 Generali France dans les Antilles	5
1.2 Contexte réglementaire	6
1.3 Objectif et périmètre	7
2 Notions sur les tempêtes et cyclones tropicaux	9
2.1 Définition	10
2.2 Description du phénomène dans le bassin Atlantique Nord	11
2.2.1 Saison cyclonique	11
2.2.2 Cycle de vie	12
2.2.3 Caractéristiques	14
2.2.4 Quelques records	16
2.3 Nomenclature : l'échelle de SAFFIR-SIMPSON	16
3 Présentation générale de la modélisation	18
3.1 La base de données HURDAT	19
3.2 Hypothèses	19
3.3 Méthodologie	20
II Simulation des tempêtes	21
4 Modélisation des trajectoires	22
4.1 Périmètres	23
4.1.1 Périmètre temporel	23
4.1.2 Périmètre spatial	23
4.2 Méthodologie	23
4.3 Loi du nombre de tempête à horizon un an	24
4.4 Simulation des trajectoires	27
4.4.1 Préliminaires	27
4.4.2 Rappels concernant les chaînes de Markov	28
4.4.3 Transposition à l'étude	29
4.4.4 Initialisation des trajectoires	29
4.4.5 Détermination des probabilités de transition	30
4.4.6 Méthode de simulation	30
4.5 Résultats	31
4.5.1 Simulation de 10 000 saisons de tempêtes	31

4.5.2	Confrontation avec les données empiriques	31
5	Détermination des caractéristiques des tempêtes le long des trajectoires simulées	35
5.1	Méthodologie	36
5.2	Modèle de pression dans l'océan	38
5.2.1	Rappels sur les modèles de régression multiples	38
5.2.2	Le modèle de régression de Vickery	42
5.2.3	Application du modèle et analyse des résidus	46
5.2.4	Traitement des localisations avec peu ou pas d'historiques	47
5.2.5	Confrontation avec les données empiriques	48
5.3	Modèle de pression dans les terres	50
5.3.1	Modèle de désintégration exponentielle avec le temps	50
5.3.2	Ajustement d'une distribution sur la constante de désintégration	52
5.3.3	Confrontation avec les données empiriques	54
5.4	Détermination de v_{max} , R_{max} et d'un gradient de vitesse des vents	55
5.4.1	Vitesse maximale des vents	55
5.4.2	Rayon de vitesse maximale des vents	57
5.4.3	Gradient de vitesse	60
5.5	Vitesse d'atteinte des îles	61
III	Simulation du coût pour l'assureur	64
6	Modélisation des coûts bruts de réassurance	65
6.1	Lien entre vitesse du vent et taux de dommage	66
6.1.1	Présentation du portefeuille concerné	66
6.1.2	Modèle de vulnérabilité	66
6.1.3	Distribution des coûts bruts de réassurance	69
6.2	Calcul du SCR de tempêtes brut de réassurance	70
6.2.1	Utilisation de l'outil de simulation de tempêtes	70
6.2.2	Comparaison avec la formule standard	71
7	Introduction de la réassurance	73
7.1	Généralités sur la réassurance	74
7.1.1	Principe de la réassurance	74
7.1.2	Les traités de réassurance classiques	74
7.1.3	Le cas particulier de la Caisse Centrale de Réassurance (CCR)	76
7.2	Le programme de réassurance de GFA Caraïbes	77
7.3	Calcul du SCR de tempêtes net de réassurance	78
7.3.1	Utilisation de l'outil de simulation de tempêtes	78
7.3.2	Détermination d'un intervalle de confiance par rééchantillonnage	79
7.3.3	Comparaison avec la formule standard	79
8	Incorporation de l'outil dans le modèle interne	82
8.1	Loi du nombre de tempête impactant GFA Caraïbes à horizon un an	83
8.2	Loi des coûts bruts générés par les tempêtes	84
8.3	Mise en œuvre de l'outil	87

Conclusion	91
Pistes d'amélioration	93
Annexes	95
A Calcul de l'intensité relative par la méthode de DARLING	95
B Détermination de la distance orthodromique entre deux points	97
C Régressions retenues dans le cadre du modèle de désintégration	99
D Equations associées aux courbes de vulnérabilité	105
D.1 Classe 1	105
D.2 Classe 2	105
D.3 Classe 3	106
E Ajustement d'une distribution sur les coûts bruts dans le cadre du modèle simplifié	107
Bibliographie	111

Introduction

Chaque année, l'océan Atlantique Nord voit se développer des phénomènes cycloniques menaçant les régions des Caraïbes, du golfe du Mexique et de la côte Est américaine. Ce risque est un enjeu majeur pour les entreprises d'assurances couvrant la zone car susceptible de provoquer des coûts de sinistres extrêmement lourds en une seule occurrence. Ainsi, en 1989 le cyclone Hugo frappait les Antilles françaises et notamment la Guadeloupe à son maximum d'intensité, provoquant plus de 600 millions d'euros de dommages et restant dans les mémoires comme le pire phénomène ayant frappé l'île depuis le milieu du XX^{ème} siècle.

Pour les compagnies d'assurances européennes telles que Generali France possédant une filiale caribéenne, la compréhension du risque de tempête tropicale apparaît donc comme essentielle en particulier depuis l'arrivée de la nouvelle directive Solvabilité 2. Cette dernière a pour but de permettre aux assureurs une meilleure maîtrise de leurs risques et réclame notamment le calcul d'un capital réglementaire appelé SCR pour chacun d'eux. Dans ce mémoire, notre objectif sera le calcul du SCR de tempêtes *via* l'élaboration d'un outil de modélisation statistique couvrant le bassin Atlantique Nord.

Dans cette optique, la présente étude s'articulera autour de trois parties.

Dans un premier temps, nous poserons les éléments préliminaires à la modélisation en répondant aux trois questions suivantes en autant de chapitres :

- Quel est le cadre, le contexte et l'objet d'une telle étude ?
- Qu'est-ce qu'un phénomène cyclonique, quelles sont ses particularités ?
- Quelle méthodologie retenir pour modéliser le risque de tempête étant donné l'historique disponible ?

La seconde partie sera consacrée à la modélisation des tempêtes à travers deux chapitres. Dans le premier chapitre nous développerons un modèle stochastique de simulations de trajectoires basé sur une chaîne de Markov. Puis, dans le second chapitre, nous associerons les caractéristiques des tempêtes le long des trajectoires simulées en s'appuyant sur des modèles de régression multiple. Deux modèles différents seront utilisés selon que le phénomène traverse une étendue d'eau ou une étendue terrestre. Il sera alors possible de suivre l'intensité des tempêtes au cours de leur déplacement et ainsi obtenir une distribution de vitesse des vents atteignant les Antilles françaises.

Enfin, la troisième partie permettra de déterminer l'impact financier des périls simulés. Pour cela, nous utiliserons un modèle de vulnérabilité faisant la liaison entre l'intensité des tempêtes et les dommages causés. Ensuite, l'introduction des traités de réassurance assurera le passage des coûts bruts aux coûts nets. L'étude s'achèvera par la mise en œuvre de l'outil de modélisation afin de déterminer le capital réglementaire associé au risque tempête et une comparaison avec le résultat obtenu par la formule standard de Solvabilité 2 sera effectuée.

Première partie

Généralités

Chapitre 1

Cadre, contexte et objectif de l'étude

Les assureurs français sont présents dans la zone tropicale du bassin Atlantique Nord par l'intermédiaire des départements et communautés d'Outre-mer (DOM-COM) situés dans l'arc d'îles des Caraïbes. Le marché de l'assurance dans les Antilles françaises se caractérise par une exposition aux risques de catastrophes naturelles accru par rapport à la métropole en terme de fréquence de survenance et de multiplicité. Ainsi, on trouve notamment les risques sismiques, volcaniques et de tempêtes mais aussi les risques de mouvement de terrain, de tsunamis et d'inondations. Dans le cadre de ce mémoire nous allons nous pencher plus particulièrement sur le risque de tempêtes tropicales, phénomène qui touche régulièrement la zone.

Dans ce chapitre nous commençons par présenter l'implantation de Generali France dans les départements d'Outre-mer antillais ainsi que les types de produits concernés par le risque étudié. Puis, nous rappelons le contexte réglementaire dans lequel se trouvent les assureurs européens et qui est à l'origine du développement d'une modélisation interne du risque tempête. Enfin, nous posons les objectifs et le périmètre de l'étude.

1.1 Generali France dans les Antilles

Generali France est présente dans les Antilles françaises par l'intermédiaire de sa filiale GFA Caraïbes. Cette dernière propose des produits d'assurance non-vie à un marché composé de la Guadeloupe, la Martinique, Saint-Barthélemy, Saint Martin et la Guyane. Cette dernière n'étant pas concernée par le risque cyclonique nous ne l'aborderons pas dans la suite de ce mémoire. En ce qui concerne les 4 autres territoires ce sont des îles dont nous donnons la localisation géographique dans la figure ci-après.

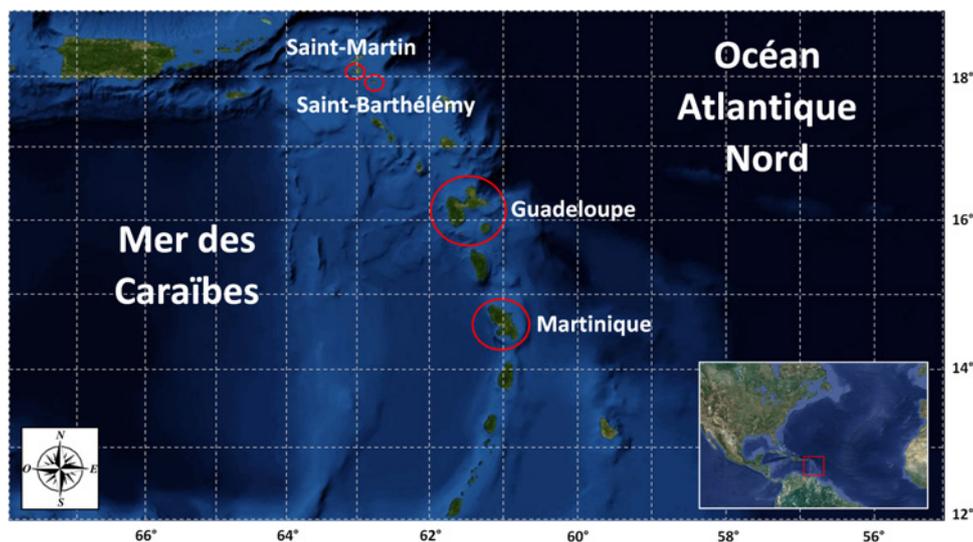


FIGURE 1 – Implantation de GFA Caraïbes

L'assureur antillais réalise une très grande partie de son chiffre d'affaires grâce aux contrats couvrant les risques liés à l'utilisation d'une automobile (responsabilité civile (RC) et dommages) et aux dommages aux biens (multirisques habitations (MRH), risques industriels (RI), etc.). A titre d'illustration nous donnons le tableau des chiffres d'affaires de GFA Caraïbes pour l'année 2011 et selon trois branches d'assurances, branches représentant à elles seules 99,6% de son chiffres d'affaires total qui est d'environ 52 millions d'euros.

	Automobile	Dommages aux biens	Frais de santé
Chiffre d'affaires en k€	28 000	13 600	10 300
Pourcentage du total	53,7%	26,1%	19,8%

La branche concernée par le risque cyclonique est essentiellement la branche dommages aux biens. Ainsi, si on considère toutes les tempêtes survenues depuis 1995, GFA Caraïbes a enregistré un coût total de sinistres bruts de réassurance d'environ 74 millions d'euros pour cette branche. La répartition de ces coûts est donnée ci-dessous.

Année	Nom de la tempête	Coût bruts en k€	Produit
1995	Luis	56 000	MRH / RI
1995	Marylin	2600	MRH / RI
1999	Jose	800	MRH / RI
1999	Lenny	2 300	MRH / RI
2007	Dean	12 000	MRH / RI

Le cyclone Luis a été particulièrement destructeur occasionnant plus de 4 ans de chiffre d'affaires dommages aux biens. Précisons que la tempête est passée sur la Guadeloupe alors qu'elle

était à son maximum d'intensité. Cet exemple à lui seul montre à quel point l'étude du risque de tempête est essentielle notamment dans une optique de transfert.

La particularité des contrats d'assurance habitations souscrit en France et en particulier dans les DOM-COM réside dans le fait que la couverture des catastrophes naturelles est obligatoire en vertu d'un régime spécial dit CAT-NAT. Le champ d'application de ce régime est défini dans l'article 1 de la loi du 13 juillet 1982 en les termes suivants :

"Les contrats d'assurance, souscrits par toute personne physique ou morale autre que l'Etat et garantissant les dommages d'incendie ou tous autres dommages à des biens situés en France, ainsi que les dommages aux corps de véhicules terrestres à moteur, ouvrent droit à la garantie de l'assuré contre les effets des catastrophes naturelles sur les biens faisant l'objet de tels contrats."

En retour, l'Etat propose aux assureurs sa garantie illimitée par l'intermédiaire de la Caisse Centrale de Réassurance (CCR) moyennant une prime². Néanmoins, une étude de l'INSEE réalisée en 2011 (voir CALVET & GRISLAIN-LETRÉMY [2]) indique que seulement 50% des ménages martiniquais et 44% des ménages guadeloupéens ont souscrit à ce type d'assurance contre 99% des ménages métropolitains. Bien que le niveau de vie des îles soit une des explications à ce phénomène, la qualité des bâtiments en est aussi une cause importante. Dans le tableau suivant on présente la typologie des structures choisies par les guadeloupéens et les martiniquais pour leur résidence principale en 2007.

	Guadeloupe	Martinique
Structures en dur	89,6%	93,7%
Logements en bois	5,5%	3,6%
Cases traditionnelles	3,6%	1,1%
Habitations de fortune	1,3%	1,6%

En général, seuls les bâtiments construits en dur sont acceptés par les assureurs et de nombreux propriétaires se retrouvent ainsi sans couvertures. Dans le cadre de cette étude nous considérons donc que les biens assurés se rapprochent de ceux que l'on trouve en métropole. Nous reviendrons sur ce point dans la troisième partie.

1.2 Contexte réglementaire

Actuellement, les entreprises d'assurances européennes sont dans une période de mutation en termes de règles prudentielles avec l'arrivée de la nouvelle directive Solvabilité 2 qui remplace progressivement Solvabilité 1.

Cette réforme fixe une nouvelle réglementation des normes comptables harmonisées au niveau européen. Son objectif est de s'assurer de la bonne santé des compagnies, ces dernières devant notamment justifier de leur solvabilité, c'est-à-dire de la détention d'un capital réglementaire suffisant afin de faire face aux engagements pris envers leurs assurés en cas d'évènements imprévus. Un indicateur classique de cette suffisance est le ratio de solvabilité définie comme le rapport entre les fonds propres détenus et le capital réglementaire. Lorsque ce ratio est inférieur à 100% l'entreprise est considérée comme non suffisamment capitalisée et le régulateur³ intervient.

Sous Solvabilité 1, ce capital réglementaire était calculé selon un pourcentage d'éléments comptables du bilan tels que les provisions mathématiques en assurance vie ou encore le chiffre d'affaires en assurance non-vie. Dans le référentiel Solvabilité 2 il apparaît deux montants :

². Pour plus de détails voir la section 7.1.3 de la troisième partie

³. En France, le régulateur est l'Autorité de Contrôle Prudenciel (ACP)

- Le capital de solvabilité requis (SCR) : défini comme le niveau de fonds propres que devra détenir la compagnie d'assurance pour assurer sa solvabilité à horizon un an avec une probabilité de 99,5%. C'est un capital cible qui pourra être calculé soit *via* une formule standard commune à toutes les compagnies soit *via* un modèle interne propre à chaque compagnie. En dessous de ce seuil l'intervention de l'autorité de contrôle sera progressive.
- Le capital minimum requis (MCR) : il correspond au montant minimum de fonds propres en dessous duquel l'autorité de contrôle interviendra automatiquement avec comme risque un retrait d'agrément.

En pratique, l'entreprise doit identifier les risques auxquels elle est exposée (par exemple un risque de chute des marchés actions ou encore un risque de défaut d'une contrepartie) et associer à chacun d'entre eux un SCR intermédiaire. L'agrégation de ces SCR intermédiaires par des matrices de corrélation permet alors d'obtenir le SCR global.

Le Groupe Generali a choisi d'opter pour un modèle interne afin de calculer son SCR et GFA Caraïbes est concernée par ce choix en tant que filiale de Generali France. Les risques auxquels GFA Caraïbes doit faire face sont spécifiques à un assureur non-vie. Nous présentons dans la figure suivante l'architecture des risques choisie afin d'obtenir le SCR de la compagnie.

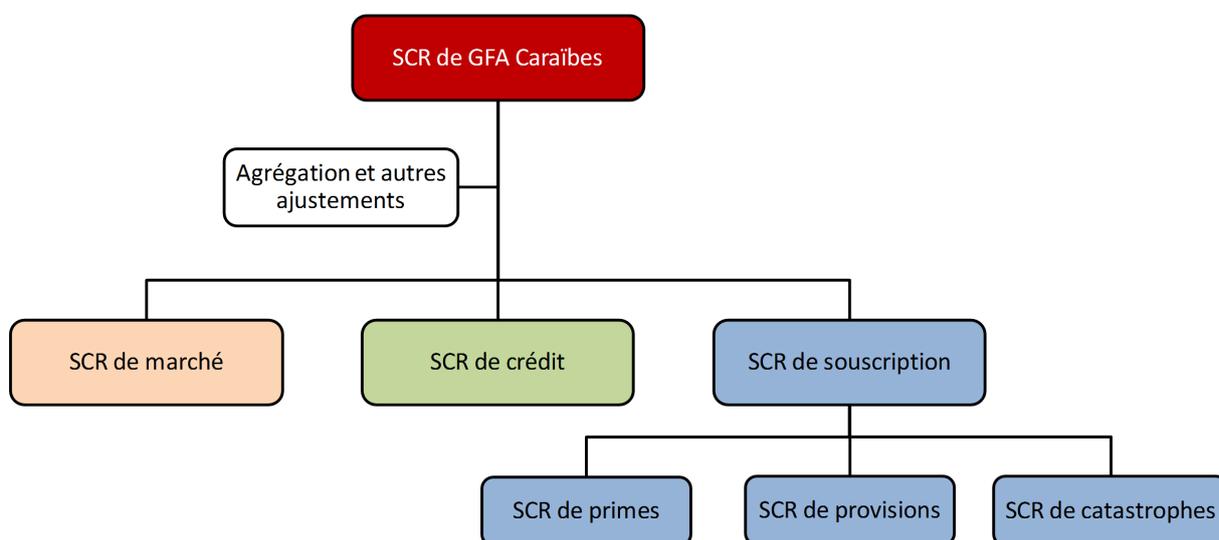


FIGURE 2 – Architectures des risques de GFA Caraïbes

Dans le cadre de ce mémoire nous nous intéresserons au SCR de tempêtes qui est un sous module du SCR de catastrophe.

1.3 Objectif et périmètre

L'objectif de l'étude est de développer un outil de modélisation stochastique permettant le calcul du SCR associé au risque de tempêtes pour GFA Caraïbes. L'horizon temporel de un an fixé par Solvabilité 2 impose de simuler ce que l'on nomme des saisons cycloniques, chaque saison représentant une année de tempêtes.

Le capital réglementaire défini par la directive correspond à une perte d'une ampleur ayant une probabilité d'occurrence de 0,5%, autrement dit la saison cyclonique la plus couteuse pour

la compagnie ayant une période de retour de 200 ans. Ainsi, l'estimation de ce coût nécessite un nombre de simulations bien supérieur à 200 afin d'obtenir un quantile le plus fiable possible. Par ailleurs, l'outil doit intégrer les traités de réassurance propres à la compagnie ces derniers étant primordiaux dans le domaine du risque catastrophe en absorbant une grosse partie des pertes.

Enfin, comme nous le verrons dans la suite, les tempêtes tropicales qui touchent les départements d'Outre-mer couverts par GFA Caraïbes se forment dans ce que l'on appelle le bassin Atlantique Nord. Ce bassin sera le périmètre de l'étude et il regroupe l'océan Atlantique Nord, la mer des Caraïbes et le Golfe du Mexique. Cette zone d'étude est représentée ci-dessous.



FIGURE 3 – Le bassin Atlantique Nord

Chapitre 2

Notions sur les tempêtes et cyclones tropicaux

Historiquement, le terme de cyclone tropical a été introduit en 1848 dans les mémoires du capitaine de marine anglais PIDDINGTON⁴ sous le nom de *kyklos* en grec. Le marin est un des pionniers dans le domaine de l'étude des tempêtes tropicales et la qualité de ses travaux fut mondialement reconnue à l'époque.

Aujourd'hui, le phénomène est un enjeu humain et matériel majeur dans de nombreuses zones du globe et fait l'objet d'intenses recherches par les scientifiques. Nous pouvons ajouter à cela les études concernant les effets du réchauffement climatique sur l'activité cyclonique, sujet de débat qui divise les experts du monde entier.

Contrairement à d'autres catastrophes naturelles telles que les éruptions volcaniques ou les tremblements de terre, les cyclones ne surviennent pas de manière complètement aléatoire et les météorologues savent que, chaque année, un certain nombre de cyclones vont se former dans les océans du globe dès lors que certaines conditions géographiques et atmosphériques sont réunies.

L'objet de ce chapitre est de décrire brièvement le cycle de vie du phénomène dans le bassin Atlantique Nord et de définir les principales caractéristiques physiques qui seront utiles pour la suite. Nous présentons aussi le système de classification des tempêtes en vigueur dans cette région de la planète.

4. *The Sailor's Horn-book for the Law of Storms* (Guide du marin sur la loi des tempêtes)

2.1 Définition

En météorologie, les tempêtes tropicales sont des perturbations atmosphériques dépressionnaires qui se développent dans les zones océaniques tropicales de l'hémisphère nord ou de l'hémisphère sud et occasionnant des vents et des précipitations. On parle abusivement de tempêtes tropicales pour désigner l'ensemble des phénomènes dépressionnaires présents dans les tropiques mais en réalité elles désignent les précurseurs des cyclones tropicaux qui en sont la forme la plus évoluée. Dans le bassin Atlantique Nord ces derniers sont aussi appelés ouragan⁵ alors que l'on parle de typhon dans le Pacifique Nord. Traditionnellement, un nom alternativement féminin et masculin est attribué aux tempêtes selon six listes préétablies que l'on retrouve périodiquement tous les six ans.

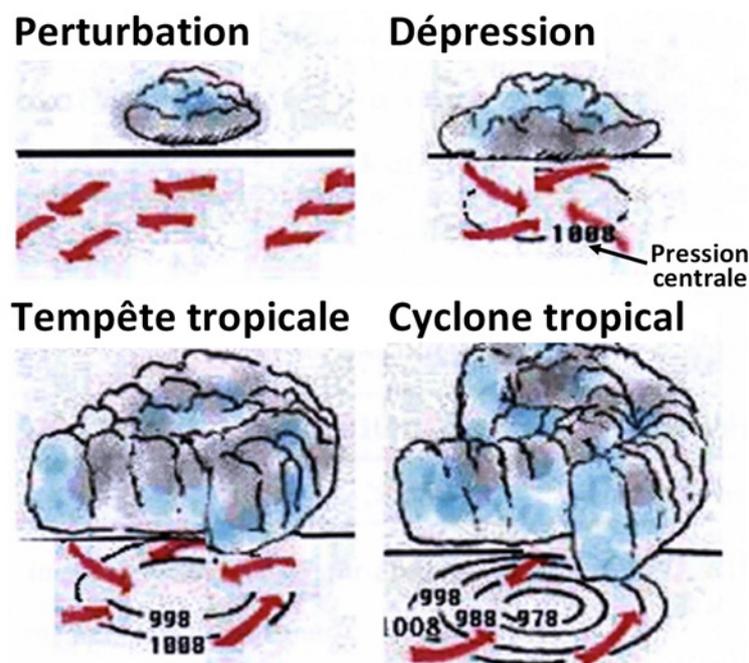


FIGURE 4 – Les différents stades d'une perturbation tropicale

D'un point de vue physique une tempête tropicale est un véritable système thermodynamique dont la formation nécessite un certain nombre de conditions météorologiques et géographiques, les principales étant la température de l'océan et la latitude. Leur source d'énergie principale est le dégagement de chaleur latente causé par la condensation de vapeur d'eau en altitude. Lorsqu'une tempête gagne assez d'énergie elle passe au stade de cyclone, un des phénomènes climatiques les plus puissants connu par les scientifiques. Les cyclones sont caractérisés par une forme en spirale composée de bandes nuageuses en rotation autour d'un œil en leur centre. La rotation est due à la force de Coriolis induite par la rotation de la Terre. Dans l'hémisphère nord le sens de rotation des cyclones est l'inverse de celui des aiguilles d'une montre. Plus de détails seront apportés dans la section suivante.

⁵. Terme francisé de *hurricane* en anglais



FIGURE 5 – Image satellitaire du cyclone Katrina (2005) dans le golfe du Mexique

Les cyclones sont parmi les phénomènes naturels les plus courants et une attention particulière leur est portée afin de protéger les populations de ses effets dévastateurs. Au niveau mondial la coordination des mesures de surveillance est organisée par l'Organisation Météorologique Mondiale (OMM), organisme des Nations Unies. Malgré tout, on dénombre encore chaque année des milliers de victimes notamment dans les pays en développement.

2.2 Description du phénomène dans le bassin Atlantique Nord

2.2.1 Saison cyclonique

Dans le bassin Atlantique Nord la saison cyclonique commence habituellement en juin et prend fin au mois de novembre. Cependant, il peut arriver qu'un phénomène se forme à d'autres mois de l'année. L'historique des mois de formation des tempêtes tropicales pour la période 1982 – 2011 est présenté dans l'histogramme ci-après.

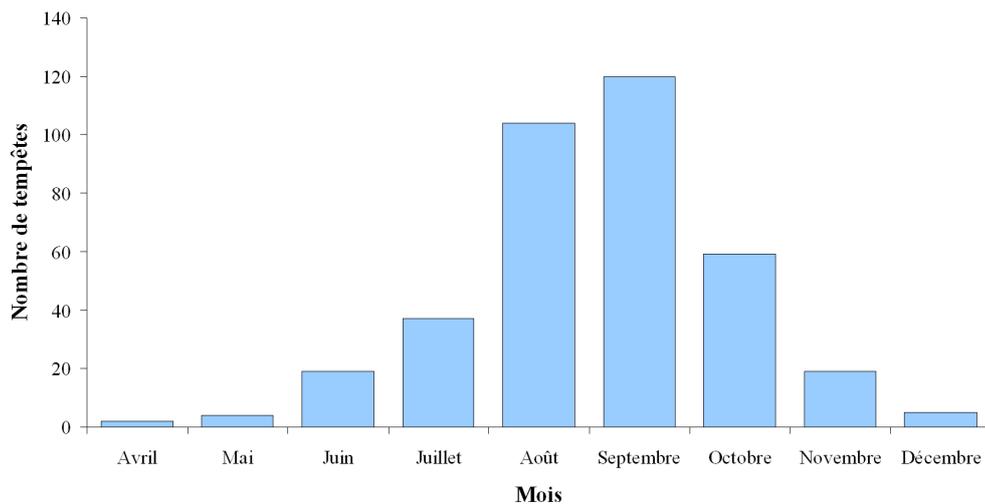


FIGURE 6 – Mois de formation des tempêtes tropicales (1982-2011)

Nous observons que la saison cyclonique bat son plein à la fin de l'été alors que la température des eaux de l'océan Atlantique atteint son niveau le plus haut de l'année. En revanche, comme nous le verrons dans la partie suivante les mois de janvier, février et mars ne sont pas propices à la formation de cyclones en raison de la trop faible température de l'eau.

2.2.2 Cycle de vie

2.2.2.1 Formation

De nombreuses conditions sont nécessaires pour qu'un cyclone tropical se développe. Les scientifiques travaillent encore sur le sujet mais en règle générale il est admis que les conditions cumulatives suivantes sont requises :

- **Conditions atmosphériques :** L'atmosphère doit être favorable à une dépression et à la formation d'orage. L'humidité de l'air doit être supérieure à 40% (typiquement 70%), indispensable à la formation de cumulonimbus et la température atmosphérique doit diminuer rapidement avec l'altitude. Enfin, la présence de vent en altitude est indispensable (jusqu'à 15 km environ). Ces vents ne doivent pas présenter de cisaillement c'est-à-dire une force et une direction qui varie selon l'altitude pouvant détruire la structure verticale du cyclone.
- **Conditions géographiques :** Le cyclone se forme nécessairement au dessus de l'eau et à une latitude suffisamment éloignée de l'équateur, c'est-à-dire environ 10° au dessus de ce dernier dans l'hémisphère nord de sorte que la force de Coriolis⁶ ne soit pas nulle. En effet, c'est cette force qui, combinée à la force du gradient de pression, va engendrer la rotation du cyclone et assurer son maintien.

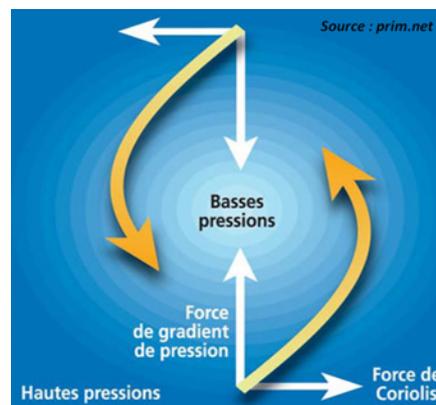


FIGURE 7 – Forces engendrant la rotation d'un cyclone

- **Conditions thermiques :** La température de l'eau doit être supérieure à 26°C jusqu'à une profondeur de 50 mètres. La température en surface de l'eau⁷ doit elle atteindre 28°C . L'eau chaude est la principale source d'énergie d'un cyclone

Il peut arriver qu'un cyclone tropical se forme en dehors de ces conditions à de rares occasions. Par exemple, le cyclone Vince survenu en 2005 s'est formé alors que la température de surface n'était que de 25°C . Cela montre que certains points sont encore à éclaircir dans la théorie des cyclones initiée par EMMANUEL [7].

6. Pour une latitude Ψ la force de Coriolis est donnée par : $f = 1,454 \times 10^{-4} \times \sin(\Psi)$

7. Cette notion sera développée dans la section 5.2.2.2 de la deuxième partie de ce mémoire

2.2.2.2 Maturité

Une fois formé un cyclone se comporte comme une gigantesque machine thermique qui puise son énergie du déséquilibre thermodynamique entre l'atmosphère (au niveau de la tropopause, limite supérieure de la troposphère elle-même située entre 8 et 15 km d'altitude) et l'océan qui jouent respectivement le rôle de source froide et de source chaude. On peut simplifier le bilan d'énergie en considérant que le système transforme l'énergie thermique extraite de l'océan en énergie cinétique. Ainsi, plus un cyclone se déplace dans des eaux chaudes et plus il gagne en énergie et devient dévastateur alors qu'au contraire il perd en intensité lorsqu'il traverse des eaux plus froides.

Un cyclone tropical à maturité se compose d'une énorme masse nuageuse enroulée en spirale pouvant atteindre un diamètre de 1000 km pour les plus importants et une altitude de 10 à 15 km. Cette enroulement se fait autour d'un noyau appelé œil du cyclone et au sein duquel les conditions climatiques sont calmes et le plus souvent sans nuages. Sa forme est circulaire et son diamètre peut varier de quelques kilomètres à plus de 200 km. A l'intérieur les courants sont descendants et on trouve les pressions les plus faibles du système. La bande nuageuse qui forme l'aspect circulaire de l'œil est appelé mur du cyclone ou mur de l'œil et c'est à cette endroit que l'on trouve les conditions climatiques les plus extrêmes.

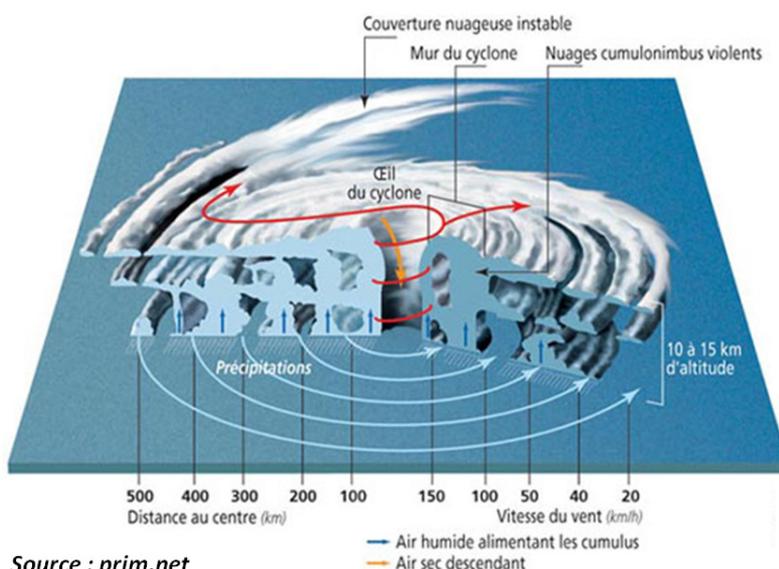


FIGURE 8 – Structure d'un cyclone tropical à maturité

Dans le bassin Atlantique Nord les trajectoires cycloniques sont généralement paraboliques en raison de deux types de circulations différentes :

- **Aux latitudes tropicales** le cyclone se déplace d'est en ouest poussé par l'alizé, un vent des régions intertropicales.
- **A l'entrée dans les zones tempérées** il change brusquement de direction pour se diriger vers le nord-est. Cet effet est due à la force de Coriolis qui augmente avec la latitude.



FIGURE 9 – Trajectoires habituelles des cyclones dans le bassin Atlantique Nord

En réalité, ce type de trajectoire n'est pas toujours constatée et on peut trouver des boucles ou encore des retours en arrière. La vitesse de déplacement d'un cyclone est en général inférieure à 20 km/h en début de parcours puis s'accélère lorsque ce dernier se dirige vers le nord. Il peut alors atteindre une vitesse supérieure à 40 km/h.

2.2.2.3 Dissipation

Lorsqu'un cyclone s'affaiblit ou meurt on parle de dissipation. Cette phase de la vie d'un cyclone apparaît lorsque l'une des conditions essentielles à son existence se détériore ou disparaît. Les causes les plus fréquentes sont les suivantes :

- **Lorsque l'eau de l'océan se refroidit** : le cyclone perd alors de l'énergie. Lorsque cette énergie devient insuffisante le phénomène meurt. C'est notamment le cas lorsqu'il prend la direction du pôle nord.
- **Lorsqu'il traverse des étendues terrestres** : on parle alors d'atterrissage du cyclone. Dans ce cas le cyclone est privé de sa principale source d'énergie que sont les eaux chaudes et il s'affaiblit très rapidement. La pression au centre de l'œil augmente alors très rapidement pour se rapprocher de la pression atmosphérique⁸. C'est notamment le cas lorsqu'un cyclone atterrit sur les grandes îles des Caraïbes dont il ressort très affaiblit.
- **Lorsque sa structure verticale subit les effets du cisaillement des vents** : cela se traduit par une détérioration de la colonne qui forme l'œil.
- **Lorsque sa trajectoire se rapproche de l'équateur** : la force de Coriolis est alors très faible et l'enroulement de la bande nuageuse n'est plus possible ce qui nuit au bon développement du cyclone.

Cette liste de causes n'est pas exhaustive et d'autres phénomènes sont à l'origine de la mort des cyclones tropicaux.

2.2.3 Caractéristiques

Pour la suite de l'étude il est nécessaire de définir quelques caractéristiques propres à une tempête ou un cyclone tropical. On trouve notamment :

⁸. Cet aspect sera étudié dans la section 5.3 de la deuxième partie

- **La pression centrale** : déjà évoquée précédemment, c'est la pression qui règne au centre du cyclone. En météorologie on a l'habitude de la mesurer en hectopascal (hPa)⁹. Pour les précurseurs nous considèrerons que c'est la pression au niveau central de la tempête.
- **La vitesse maximale des vents** : c'est la vitesse maximale atteinte par une rafale de vent à l'intérieur d'un cyclone pendant une période donnée. Généralement, cette vitesse est atteinte au mur de l'œil et on l'exprime en mètre par seconde (m/s) ou kilomètre par heure (km/h). Selon la convention choisie, cette vitesse est définie comme étant le maximum de la vitesse moyenne des vents soutenus pendant 1 minute ou pendant 10 minutes à une altitude prédéfinie. Dans le cadre de ce mémoire c'est la première qui sera retenue. Notons que l'on trouve aussi la notion de vitesse maximale d'une rafale sur 3 secondes. Cependant, les rafales étant des phénomènes très localisés on préfère déterminer l'intensité des vents sur des intervalles de temps plus longs. En règle générale, on considère que la vitesse d'une rafale sur 3 secondes est 30% plus élevée qu'une vitesse moyenne soutenue pendant 1 minute cette dernière étant elle-même 12% plus élevée qu'une vitesse soutenue pendant 10 minutes¹⁰.
- **Le rayon de vitesse maximale des vents** : c'est la distance entre le centre de la tempête et le mur du cyclone, c'est-à-dire là où est atteinte la vitesse maximale des vents. Cela correspond donc en gros au rayon de l'œil. Ainsi, pour les précurseurs cette notion n'existe pas.

La figure suivante montre l'évolution habituelle de la pression et de la vitesse des vents en fonction de la distance r au centre de la tempête.

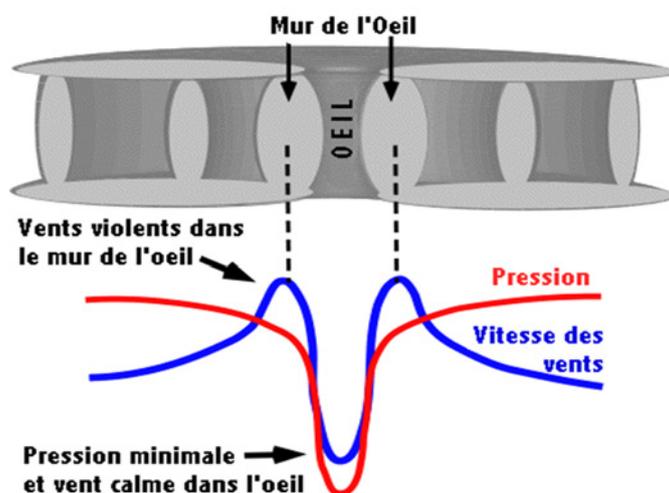


FIGURE 10 – Evolution de la pression et de la vitesse des vents dans un cyclone

L'augmentation de la vitesse des vents depuis la centre du cyclone est proportionnelle à la distance r jusqu'au mur de l'œil puis décroît en $1/r$. On caractérise souvent l'intensité d'une tempête par la vitesse maximale des vents, vitesse qui est intimement liée à la valeur de la pression centrale. En général, plus la pression centrale est faible et plus le phénomène sera dévastateur.

⁹. Typiquement la pression atmosphérique dans des conditions normales se situe autour de 1013 hPa

¹⁰. Source : <http://www.aoml.noaa.gov/hrd/tcfaq/D4.html>

2.2.4 Quelques records

Nous présentons dans le tableau suivant quelques chiffres records concernant les cyclones tropicaux et tempêtes tropicales survenus dans le bassin Atlantiques depuis le début du XX^{ème} siècle.

Record	Chiffre	Nom du cyclone	Année
Plus faible pression centrale	882 hPa	Wilma	2005
Vitesse maximale la plus élevée	≈ 300 km/h	Wilma	2005
Record de longévité	27, 25 jours	Ginger	1971
Trajet le plus long	12 500 km	Faith	1966
Cyclone le plus coûteux	> 110 milliards de \$ US	Katrina	2005
Cyclone le plus meurtrier	11 000 – 18 000 morts	Mitch	1998
Nombre record pour une année	28 tempêtes nommées	–	2005

Le bilan de Mitch et Katrina, survenus il y a moins de 15 ans, montre à lui seul l'étendu des dégâts que peut provoquer un ouragan, à la fois humains et matériels. Pourtant, ces deux phénomènes n'ont pas atteint le record d'intensité détenu par Wilma mais leurs trajectoires ont traversé des zones particulièrement peuplées et vulnérables.

2.3 Nomenclature : l'échelle de Saffir-Simpson

Dans le bassin Atlantique Nord les tempêtes sont classifiées selon leur intensité grâce à une échelle proposée par l'ingénieur américain SAFFIR et le météorologue SIMPSON, deux américains spécialistes des ouragans. Cette échelle est formée de 2 catégories de précurseurs et de 5 catégories de cyclones. La vitesse maximale correspond au maximum de la vitesse moyenne des vents soutenus pendant 1 minute à 10 mètres d'altitude (hauteur standard d'observation en météorologie).

Dénomination	Vitesse maximale des vents	Pression centrale habituelle	Dommmages potentiels
Dépression tropicale	< 62 km/h	> 995 hPa	Pas de dommmage.
Tempête tropicale	63 - 118 km/h	> 995 hPa	Dommmages négligeables
Cyclone - catégorie 1	119 - 153 km/h	980 - 994 hPa	Dommmages aux maisons mobiles et à la végétation. Dommmages faibles à la structure des bâtiments.
Cyclone - catégorie 2	154 - 177 km/h	965 - 979 hPa	Dommmages possibles aux toitures, portes et fenêtres des habitations.
Cyclone - catégorie 3	178 - 208 km/h	945 - 964 hPa	Dommmages moyens à la structure des bâtiments. Les maisons mobiles sont détruites.
Cyclone - catégorie 4	209 - 251 km/h	921 - 944 hPa	Dommmages importants à la structure des bâtiments.
Cyclone - catégorie 5	> 252 km/h	< 920 hPa	Destruction des toitures, des portes et des fenêtres. Des habitations sont complètement détruites.

FIGURE 11 – Echelle de SAFFIR-SIMPSON

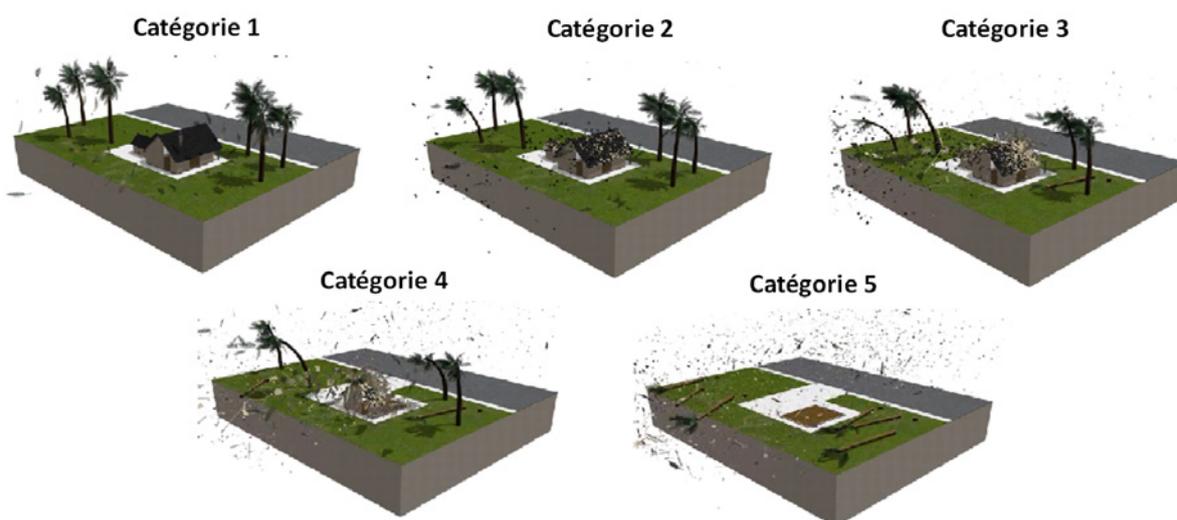
Un niveau de détails plus importants concernant les dommmages est donné dans [14]. Cette échelle est considérée trop simpliste par certains spécialistes car elle ne prend pas en compte certains aspects importants du phénomène tels que sa taille ou son taux de précipitation. Cependant, c'est l'échelle officielle qui est utilisée dans l'Atlantique Nord. Notons qu'il existe d'autres classifications utilisées dans les autres bassins du globe comme par exemple l'échelle de BEAUFORT basée sur la vitesse moyenne des vents soutenue pendant 10 minutes.

Pour fixer les idées nous donnons dans le tableau ci-dessous les nombres et pourcentages de cyclones ayant atteint chacune des catégories au cours de la période 1982 – 2011.

	Catégorie 1	Catégorie 2	Catégorie 3	Catégorie 4	Catégorie 5
Nombre	82	32	31	39	12
Pourcentage	41,8%	16,3%	15,8%	19,9%	6,2%

Les cyclones de catégorie 1 sont les plus courants alors que ceux de la catégorie 5 sont heureusement plutôt rares. Cependant, la tendance est à la hausse des cyclones de forte intensité. Notons qu'au cours de cette période on dénombre aussi 173 précurseurs ayant reçu un nom.

Pour illustrer cette nomenclature, nous donnons dans la figure suivante les dommages potentiels qui peuvent être causés par les cyclones tropicaux en fonction de la catégorie d'intensité à travers l'exemple d'une résidence pavillonnaire avec une structure en maçonnerie et un toit en tuile.



Source : National Oceanic and Atmospheric Administration (NOAA)

FIGURE 12 – Illustration des dommages par catégorie

Chapitre 3

Présentation générale de la modélisation

Ce chapitre présente la base de données qui servira d'*input* pour la plupart des modèles développés, base qui recense l'ensemble des phénomènes cycloniques depuis les dernières années. Il a aussi pour but d'explicitier la démarche que nous adopterons pour la modélisation du risque de tempête ainsi que les hypothèses sous-jacentes qui seront retenues tout au long de l'étude.

3.1 La base de données HURDAT

Depuis quelques années, le *National Hurricane Center* (NHC), l'agence américaine en charge de la localisation et de la prédiction des tempêtes dans l'océan Atlantique Nord, fournit gratuitement sur internet une base de données nommées HURDAT (*North Atlantic Hurricane Database*) répertoriant toutes les tempêtes tropicales du bassin Atlantique Nord depuis l'année 1851¹¹. Cette dernière est considérée comme le meilleur relevé de trajectoires historiques ou en anglais *historical best track*. Dans la figure 13 nous donnons à titre d'exemple un extrait des caractéristiques que l'on peut trouver sur la tempête Katrina survenue en 2005.

Année	Mois	Jour	Heure	Numéro	Nom	Latitude	Longitude	Vitesse maximale des vents	Pression au centre de l'œil
2005	8	23	18	11	KATRINA	23,1	75,1	30	1008
2005	8	24	0	11	KATRINA	23,4	75,7	30	1007
2005	8	24	6	11	KATRINA	23,8	76,2	30	1007
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

FIGURE 13 – Allure de la base HURDAT

Précisons que les relevés sont donnés à 6 heures d'intervalle en UTC (temps universel coordonné), que les latitudes sont en degrés nord avec une précision de un dixième, que les longitudes sont en degrés ouest avec la même précision et que les pressions centrales sont des entiers exprimés en millibars¹². Les vitesses maximales des vents correspondent au maximum de la moyenne des rafales soutenues pendant 1 minute et sont des entiers exprimés en nœuds¹³.

Par ailleurs, cette base ne contient pas seulement les cyclones de catégorie 1 à 5 sur l'échelle de SAFFIR-SIMPSON puisqu'elle recense également les précurseurs. Ces derniers seront pris en compte dans l'étude puisque, outre l'historique supplémentaire qu'ils apportent, ce sont des phénomènes qui se sont dissipés prématurément mais qui auraient très bien pu se transformer en cyclones tropicaux.

Enfin, il existe une version étendue de la base HURDAT qui donne plusieurs caractéristiques supplémentaires dont le rayon de vitesse maximale des vents depuis 1988¹⁴. Les données relatives à ce rayon ont été répertoriées par DEMARIA & al. à partir de divers rapports concernant les ouragans et sont données avec une précision de 5 kilomètres.

3.2 Hypothèses

Dans toute la suite nous ferons les hypothèses suivantes :

- A l'instar de la base HURDAT, une trajectoire sera formée par un ensemble de points repérés par leurs latitudes et leurs longitudes et espacés d'un intervalle de temps de 6 heures.
- Le déplacement d'une tempête entre deux points sera considéré comme étant un mouvement de translation à vitesse constante.
- Nous considérerons que chaque cyclone forme un disque parfait ce qui est rarement vérifié dans la réalité¹⁵. En effet, le phénomène présente souvent une asymétrie plus ou moins

11. <http://weather.unisys.com/hurricane/>

12. 1 millibars = 10^2 Pascals = 1 hPa

13. 1 nœud \approx 1,85 km/h

14. http://rammb.cira.colostate.edu/research/tropical_cyclones/tc_extended_best_track_dataset/index.asp

15. Les cyclones présentant une symétrie quasi-parfaite sont appelés cyclones annulaires

importante du fait de son déplacement et de sa composition nuageuse enroulée en spirale. Cependant, ce n'est pas une approximation préjudiciable car cela concerne une zone éloignée de l'œil, là où les vents sont relativement calmes.

- Nous assimilerons la Terre à une sphère de rayon $R_T = 6\,400$ km notamment pour faciliter le calcul de distance entre deux localisations repérées par leurs coordonnées géographiques.
- Le risque étudié se limite au risque de destruction de biens provoquée par les vents violents générés par une tempête. Ainsi, ce mémoire ne traite pas du risque d'inondation dont une corrélation existe avec le risque de tempêtes¹⁶ en particulier en raison des fortes précipitations qui accompagnent les phénomènes cycloniques.

3.3 Méthodologie

L'approche retenue pour modéliser le risque tempête est une approche classique dite fréquence-coût. Cependant, le nombre d'évènements répertoriés dans les Antilles françaises n'est pas suffisant pour leur ajuster une distribution de probabilité. C'est pour cela que notre choix s'est porté sur un outil de simulation couvrant non pas la zone couverte par GFA Caraïbes mais le bassin Atlantique Nord dans son ensemble afin d'avoir un nombre de données historiques bien plus important pour l'élaboration des différents modèles. La démarche retenue pour la construction de l'outil est donnée dans la figure ci-après.

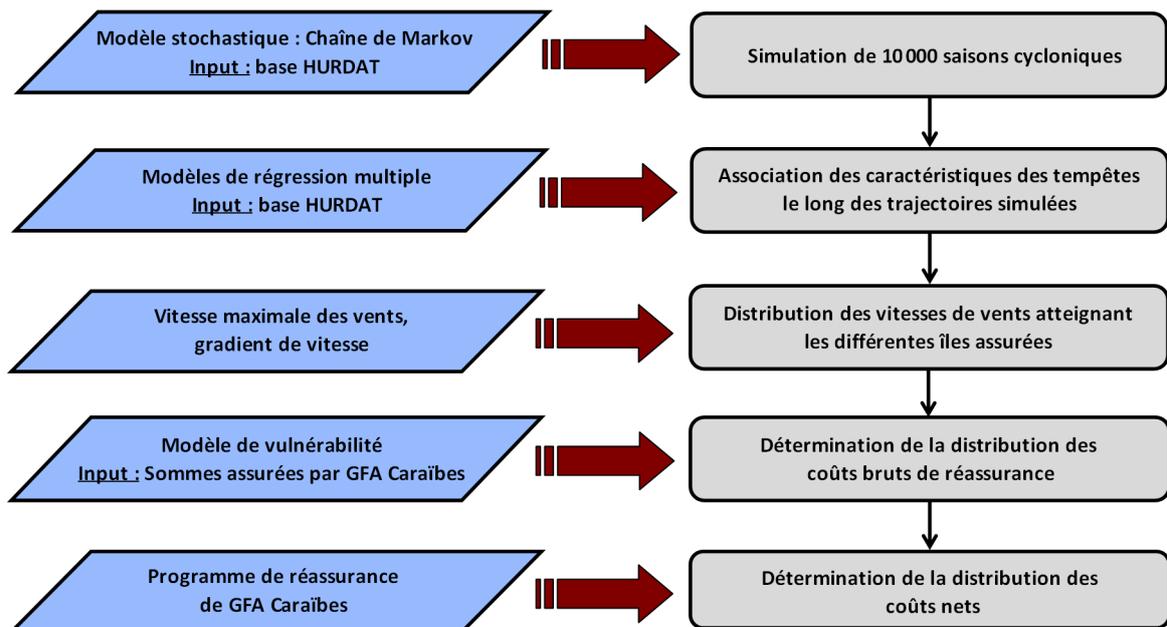


FIGURE 14 – Modélisation du risque tempête

Les trois premières étapes correspondent à la simulation des tempêtes, c'est la partie fréquence de l'étude et elle fera l'objet de la deuxième partie de ce mémoire. Quant à la partie coût, correspondant aux deux dernières étapes de la modélisation, elle sera étudiée dans la troisième et dernière partie. Cette dernière s'achèvera par le calcul du capital réglementaire associé au risque de tempête pour GFA Caraïbes.

¹⁶. Cette corrélation est fixée à 25% dans les spécifications techniques de la formule standard

Deuxième partie

Simulation des tempêtes

Chapitre 4

Modélisation des trajectoires

La modélisation des trajectoires des tempêtes est une étape délicate du fait du nombre importants de paramètres physiques qui influent sur la direction et la vitesse du système à chaque instant. Parmi ces paramètres nous pouvons citer l'orientation des vents, la force de Coriolis (qui varie en fonction de la latitude) ou encore un déplacement dans les terres plutôt que dans l'océan.

Ainsi, les chercheurs ont développé de très nombreux modèles qui peuvent être classés en deux grandes catégories :

- Les **modèles dynamiques** qui prédisent la trajectoire d'une tempête en fonction de variables météorologiques régies par des équations aux dérivées partielles de mécanique des fluides. Souvent, une correction est prise en compte pour considérer ce que l'on nomme communément "l'effet bêta", c'est-à-dire la variation de la force de Coriolis due à un changement de latitude. Ce type de modèles n'est applicable qu'à *posteriori*, c'est-à-dire une fois le phénomène formé car il requiert la connaissance et le suivi de l'environnement atmosphérique du phénomène cyclonique à chaque instant.
- Les **modèles statistiques** qui permettent de générer des trajectoires en fonctions de l'étude statistique des trajectoires passées. Dans cette approche on fait de l'inférence statistique, c'est-à-dire que l'on généralise ce qui a été observé dans un échantillon de trajectoires passées à toutes les futures trajectoires. On trouve par exemple des modèles de régression basés sur le lieu géographique comme celui de VICKERY & al. [15] dans lequel la direction et la vitesse de la tempête sont mises à jour en fonction de la latitude, de la longitude mais aussi de la direction et de la vitesse à deux instants précédents. D'autres ont utilisé une approche stochastique en ajustant une loi de probabilité sur la vitesse maximale des vents dans une zone d'intérêt (voir NEUMANN [13]).

Dans le cadre de l'étude, c'est un modèle statistique qui a été développé *via* l'utilisation d'une chaîne de Markov homogène à temps discret. L'objet de ce chapitre est d'explicitier la méthodologie adoptée, les hypothèses retenues et les résultats obtenus par une telle approche.

4.1 Périmètres

4.1.1 Périmètre temporel

Bien que la base HURDAT remonte assez loin dans le temps, un historique de 30 ans seulement sera retenu (1982 – 2011) en raison de deux principaux facteurs. D’une part, les études récentes montrent une tendance à l’augmentation de l’intensité des tempêtes en raison du réchauffement climatique. Ainsi, une profondeur d’historique trop importante aurait tendance à atténuer cet effet ce qui serait préjudiciable pour l’étude qui sera faite dans le chapitre 5 (en revanche nous considérerons que le nombre de tempête n’est pas affecté par ce réchauffement). D’autre part, avant le début des relevés par satellites dans les années 70, une tempête était considérée en début de formation seulement à partir du moment où un bateau ou une station météorologique avait réussi à la détecter. Il n’est donc pas exclu que des portions de trajectoires soient manquantes voir même que certaines tempêtes n’aient jamais été détectées.

4.1.2 Périmètre spatial

Les données de la base HURDAT pour la période 1982 – 2011 sont toutes contenues dans une zone allant de 0 à 59,9 degrés nord en latitude et de 0 à 109,9 degrés ouest en longitude. Nous pourrions donc nous limiter au périmètre spatial $[0^\circ; 60^\circ] \times [0^\circ; 110^\circ]$. Dans la figure suivante nous avons représenté l’ensemble des points d’impacts à 6 heures d’intervalle au cours de cette période ce qui représente 11536 points. Nous avons également représenté leur densité en terme de latitude et longitude.

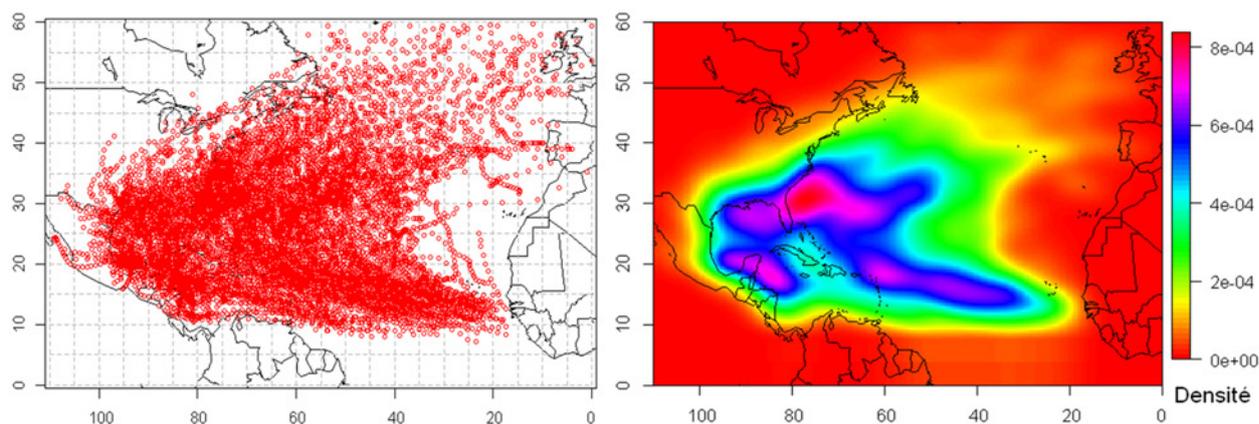


FIGURE 15 – Points d’impact à 6 heures d’intervalle (1982-2011)

Nous retrouvons l’aspect parabolique des trajectoires avec une tendance à naître dans les eaux chaudes à 10° au dessus de l’équateur avant de percuter les Caraïbes puis de mourir dans les eaux froides de l’Atlantique Nord. Nous observons aussi que la mer des Caraïbes, le golfe du Mexique ainsi que la côte Est américaine sont les zones les plus exposées.

4.2 Méthodologie

Les tempêtes tropicales sont des phénomènes courants à l’échelle du bassin Atlantique puisque, d’après les statistiques de la base HURDAT pour les 30 dernières années, on recense une moyenne de 12,3 tempêtes par an. Cependant, si l’on s’intéresse à un site en particulier (par exemple une zone autour de la Guadeloupe) on se rend compte que la fréquence de survenance de ces événements reste heureusement très faible à horizon un an. Afin d’obtenir un échantillon

d'évènements plus important une méthode classique est alors d'utiliser une approche par simulation dite de Monte-Carlo. L'idée d'une telle approche est de modéliser par un modèle statistique les trajectoires des tempêtes tropicales en utilisant l'historique disponible, puis de générer un très grand nombre de simulations *via* ce modèle afin de récupérer un échantillon plus important.

Dans le cadre de l'étude, et pour respecter la convention temporelle de la directive Solvabilité 2, nous chercherons à modéliser des saisons complètes de tempêtes tropicales ce qui revient à considérer un horizon temporel de une année. Ainsi, nous commencerons par ajuster une loi de probabilité sur le nombre d'occurrences par an avec pour chacune de ces occurrences une initialisation de la position initiale en terme de latitude, longitude, mois de survenance et pression centrale. Puis, partant de chacune de ces positions initiales les trajectoires des tempêtes seront simulées *via* une chaîne de Markov discrète et homogène. Dans la figure 16 ci-après nous présentons l'algorithme résumant cette démarche. En utilisant cet algorithme un grand nombre de fois nous serons alors en mesure d'obtenir un échantillon très conséquent de saisons cycloniques.

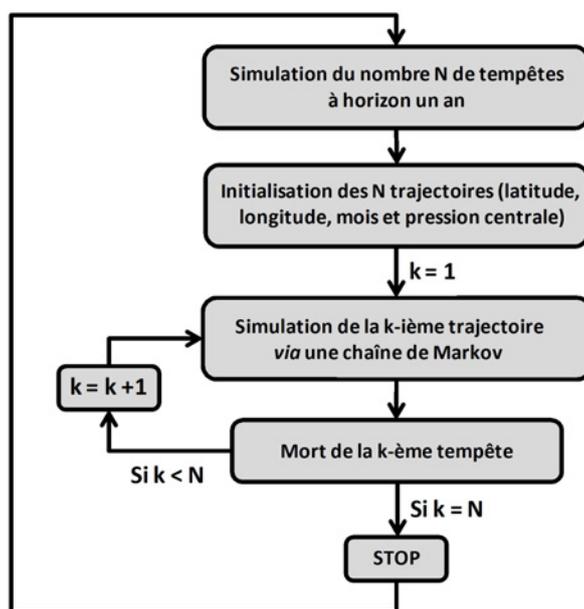


FIGURE 16 – Algorithme de simulation d'une saison complète de trajectoires

4.3 Loi du nombre de tempête à horizon un an

La période qui s'étale de 1982 à 2011 comporte $n = 369$ tempêtes tropicales avec une moyenne annuelle de 12,3 et une variance de 25,9 soit un écart-type de 5,1. L'année 1983 possède le plus faible nombre d'évènements sur la période avec seulement 4 tempêtes alors que l'année 2005 a été une année record avec 28 tempêtes répertoriées. Dans la figure ci-dessous nous avons tracé la série temporelle et l'histogramme du nombre d'occurrence par an au cours des 30 dernières années.

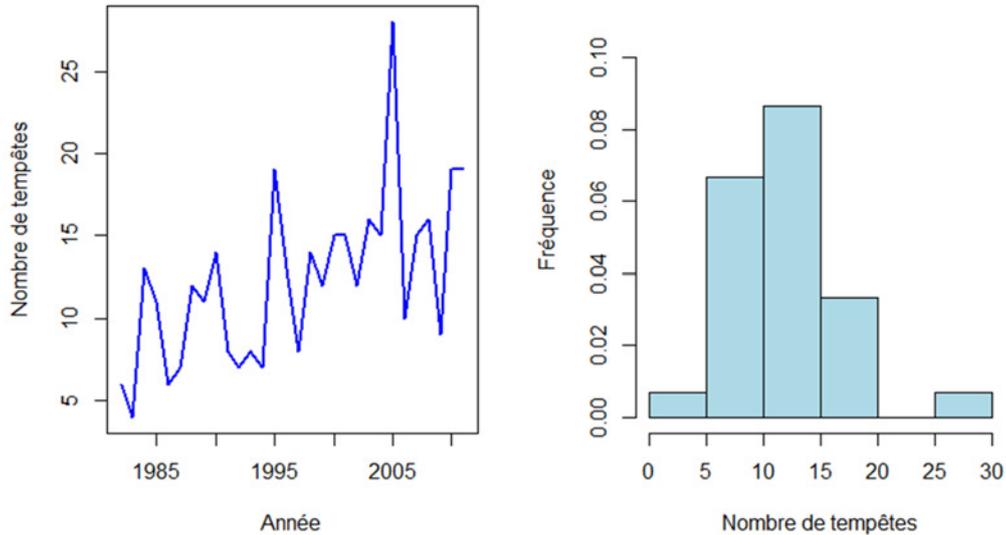


FIGURE 17 – Nombre d’occurrence à horizon un an et histogramme associé (1982 – 2011)

Nous remarquons une tendance à l’augmentation du nombre d’évènements durant la période que nous pourrions être tenté de prendre en considération. Néanmoins, les experts de l’OMM considèrent que cette évolution de l’activité cyclonique est une variation naturelle cyclique et multi-décennale due à une modification cyclique du courant océanique de l’Atlantique. Paradoxalement, un réchauffement climatique stabiliserait voir diminuerait le nombre de cyclones mais ces derniers seraient bien plus dévastateurs¹⁷. Ainsi, cette tendance ne sera pas prise en compte et nous retiendrons les nombres bruts.

Concernant l’histogramme, il semblerait que le nombre de tempêtes suive une loi binomiale négative du fait de l’existence d’une queue de distribution épaisse, cependant nous allons aussi tester la loi de Poisson qui est la loi discrète la plus classique. Le choix de la loi se fera en déterminant la distribution qui maximise la log-vraisemblance, puis un test du chi-deux d’adéquation sera effectué pour validation.

Soit N la variable aléatoire représentant le nombre annuel de tempêtes et (n_1, \dots, n_{30}) l’échantillon empirique à disposition avec $\sum_{i=1}^{30} n_i = 369$. En posant $n = 30$ le nombre d’année d’historique nous rappelons que :

- Si N suit une loi de Poisson de paramètre λ (notée $\mathcal{P}(\lambda)$) alors :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}[N = k] = e^{-\lambda} \frac{\lambda^k}{k!}$$

Cette loi est caractérisée par une équi-dispersion ($\mathbb{E}[N] = \text{Var}[N] = \lambda$) ce qui semble inadapté à notre échantillon et l’estimateur $\hat{\lambda}$ de λ du maximum de vraisemblance et des moments est donnée par la moyenne empirique : $\hat{\lambda} = \bar{N} = \frac{1}{n} \sum_{i=1}^n n_i$

- Si N suit une loi binomiale négative de paramètres m et p (notée $\mathcal{BN}(m, p)$) alors :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}[N = k] = C_{m+k-1}^k p^m (1-p)^k$$

¹⁷. Source : http://www.wmo.int/pages/mediacentre/infonotes/infonote62_en.html

Cette loi à le mérite d'être sur-dispersée ($\mathbb{E}[N] < \text{Var}[N]$) avec une queue épaisse ce qui correspond à notre situation. La méthode des moments donne pour estimateurs : $\tilde{p} = \frac{\bar{N}}{\hat{\sigma}^2}$ et $\tilde{m} = \frac{\bar{N}^2}{\hat{\sigma}^2 - \bar{N}}$ avec $\hat{\sigma}^2$ la variance empirique. Il n'existe pas de formules fermées pour les estimateurs du maximum de vraisemblance mais ils vérifient :

$$\hat{p} = \frac{\hat{m}}{\hat{m} + \bar{N}} \quad \text{et} \quad \sum_{j=0}^{l-1} \frac{1}{\hat{m} + j} \left[1 - \sum_{k=0}^j f_k \right] - \ln \left(1 + \frac{\bar{N}}{\hat{m}} \right) = 0$$

où l est la plus grande valeur observée parmi (n_1, \dots, n_{30}) et f_k la proportion du nombre n_i égaux à k ($k \in \{0, \dots, l\}$). Ces équations peuvent être résolues par une méthode de Newton en prenant comme initialisation les estimateurs des moments.

Dans le tableau suivant nous donnons la log-vraisemblance obtenue grâce au logiciel R pour la loi de Poisson et la loi binomiale négative ainsi que les estimateurs des paramètres.

Distribution	Log-vraisemblance	Paramètres estimés
$\mathcal{P}(\lambda)$	-93,9	$\hat{\lambda} = 12,3$
$\mathcal{BN}(m, p)$	-89,3	$\hat{m} = 12,5$ et $\hat{p} = 0,5$

Comme prévu, c'est la distribution binomiale négative qui maximise la log-vraisemblance. Nous allons alors tester l'hypothèse suivante :

$$H_0 : N \sim \mathcal{BN}(\hat{m}, \hat{p})$$

Un outil classique pour les lois discrètes est le test d'adéquation du chi-deux décrit brièvement ci-après.

Principe du test du chi-deux : Soit (x_1, \dots, x_ν) un échantillon de ν données issues d'une variable aléatoire discrète X . On répartie l'échantillon en J classes notées $(\mathcal{C}_j)_{j=1, \dots, J}$. Le but est de tester l'hypothèse nulle selon laquelle les probabilités que X prenne ses valeurs dans $\mathcal{C}_1, \dots, \mathcal{C}_J$ sont respectivement p_1, \dots, p_J , ces dernières étant issues d'une distribution théorique. Si on note \hat{p}_j la probabilité empirique que X appartienne à \mathcal{C}_j on a :

$$\hat{p}_j = \frac{1}{\nu} \sum_{i=1}^{\nu} \mathbb{1}_{\{x_i \in \mathcal{C}_j\}} \quad \text{avec} \quad \mathbb{1}_{\{x_i \in \mathcal{C}_j\}} = \begin{cases} 1 & \text{si } x_i \in \mathcal{C}_j \\ 0 & \text{sinon} \end{cases}$$

On définit alors la statistique :

$$D = \nu \times \sum_{j=1}^J \frac{(\hat{p}_j - p_j)^2}{p_j}$$

Sous l'hypothèse nulle cette statistique suit asymptotiquement une loi du chi-deux à $J-1$ degrés de liberté. Ainsi, on construit un test asymptotique de niveau α en comparant la valeur de D calculée à partir d'un échantillon et le quantile d'ordre $1-\alpha$ d'une loi du chi-deux à $J-1$ degrés de liberté noté $\chi_{1-\alpha}^2$. On rejetera l'hypothèse nulle si $D > \chi_{1-\alpha}^2$.

Dans notre cas nous avons un échantillon de taille $\nu = 30$ assez important pour effectuer le test, les p_j correspondent aux probabilités théoriques d'une loi binomiale négative de paramètres \hat{m} et \hat{p} et le nombre de classe a été déterminé à partir de la règle empirique de Yates selon laquelle un bon choix pour J est $1 + 3,33 \log(\nu) \approx 6$. Le logiciel R donne le résultat suivant pour un test de niveau 5% :

Statistique	Seuil critique	p-valeur
5,8727	11,0705	0,3188

Au vu du test, la p-valeur, c'est-à-dire la probabilité de rejeter à tort H_0 , est de l'ordre de $32\% > 5\%$ donc nous retiendrons l'hypothèse nulle : **le nombre de tempêtes tropicales à horizon un an suit une loi binomiale négative de paramètres $m = 12,5$ et $p = 0,5$.**

Nous pouvons aussi nous assurer que, graphiquement, la distribution choisie est bien en accord avec les données empiriques. Pour cela on a représenté dans la figure suivante l'histogramme théorique d'une $\mathcal{BN}(12,5; 0,5)$ ainsi que l'histogramme empirique du nombre de tempêtes. Sur ce dernier nous avons aussi tracé le diagramme en bâton théorique.

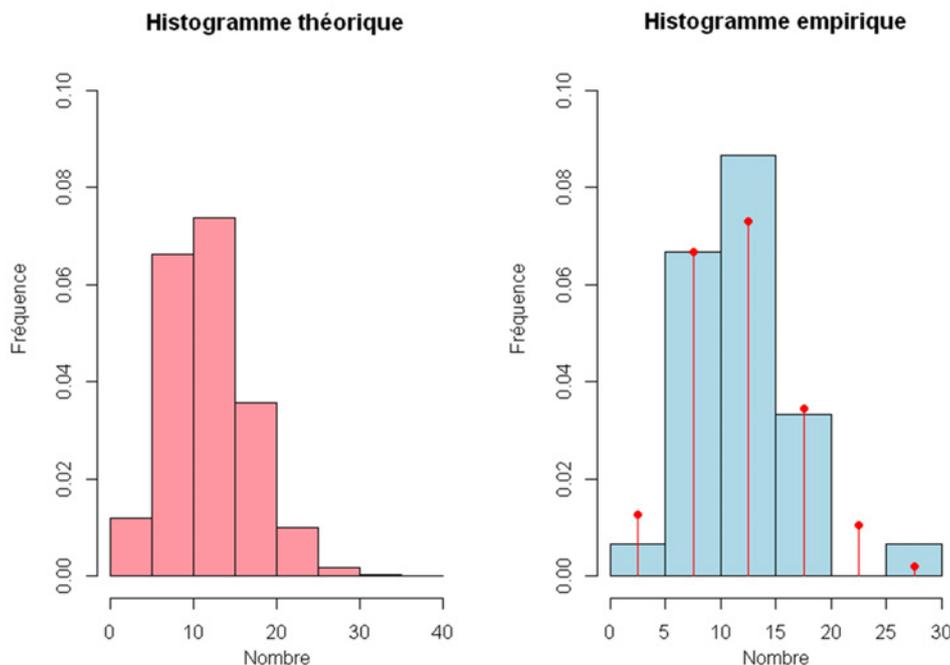


FIGURE 18 – Histogrammes théorique et empirique du nombre d'occurrences

4.4 Simulation des trajectoires

4.4.1 Préliminaires

Nous considérons le périmètre spatial défini précédemment que l'on discrétise selon des carrés d'aire $r_x \times r_y$ où r_x et r_y représentent respectivement la résolution en latitude (nord) et en longitude (ouest). Dans le cadre de ce mémoire nous avons choisi une résolution de 1° en latitude et longitude ce qui permet d'avoir un quadrillage assez fin ($60 \times 110 = 6600$ cases) tout en restant proche de la résolution choisie dans la base HURDAT (un dixième de degrés). Précisons qu'en considérant la Terre parfaitement sphérique un degrés de latitude ou de longitude correspond à environ 110 km¹⁸. On attribue ensuite à chaque case du quadrillage un numéro de 1 à 6600 en partant de l'extrémité sud-est et en balayant la zone de gauche à droite comme l'illustre la figure suivante.

¹⁸. En prenant pour rayon de la Terre $R_t = 6400$ km nous obtenons bien $\frac{\pi}{180} \times R_t \approx 110$ km

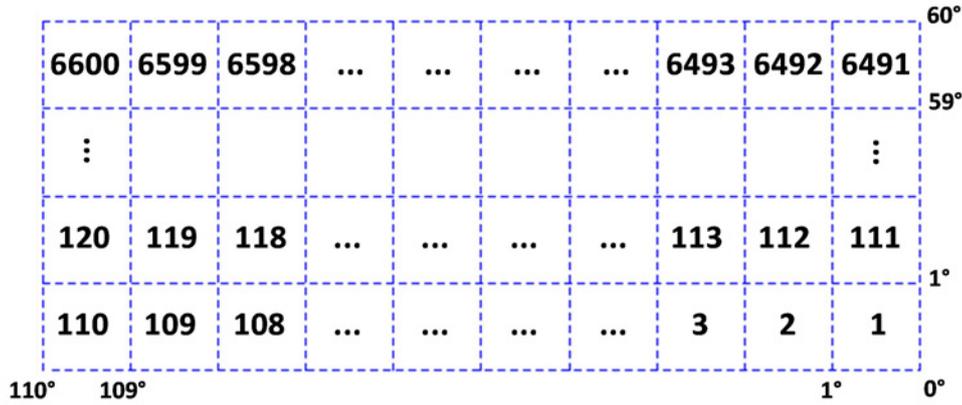


FIGURE 19 – Quadrillage du bassin Atlantique selon une grille $1^\circ \times 1^\circ$

On affecte alors une localisation aux points de chaque trajectoires historiques en utilisant la formule suivante :

$$\text{Localisation} = 110 \times [\text{Latitude}] + [\text{Longitude}] + 1$$

où $x \mapsto [x]$ désigne la fonction partie entière. Réciproquement, chaque localisation sera repérée par la latitude et la longitude correspondant au milieu du carré qui lui est associé. Par exemple, si nous appliquons cette méthodologie à la position initiale du cyclone Alberto survenu en 1982 nous obtenons le tableau de correspondance suivant :

Latitude	Longitude	Localisation $1^\circ \times 1^\circ$	Nouvelle latitude	Nouvelle longitude
$22, 2^\circ$	$86, 5^\circ$	2507	$22, 5^\circ$	$86, 5^\circ$

4.4.2 Rappels concernant les chaînes de Markov

Avant d'aller plus loin nous rappelons quelques résultats utiles pour l'étude concernant les chaînes de Markov.

Définition 1 (Chaîne de Markov). Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires discrètes à valeur dans un espace d'état E . On dit que $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov discrète de matrice de transition P de terme général $(p_{i,j})_{(i,j) \in E^2}$ si pour tout $n \in \mathbb{N}$ et tout $(x_0, \dots, x_{n+1}) \in E^{n+2}$:

$$\mathbb{P}[X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0] = \mathbb{P}[X_{n+1} = x_{n+1} | X_n = x_n] = p_{x_n, x_{n+1}}$$

Cette définition implique que la probabilité de survenance d'un état futur est entièrement déterminée par l'état présent : "le futur ne dépend du passé que par l'état présent". Remarquons que la loi initiale, c'est-à-dire la loi de X_0 n'est pas fixée par la définition.

Définition 2 (Chaîne de Markov homogène). Si le mécanisme de transition ne change pas au cours du temps on parle de chaîne de Markov homogène et la définition précédente devient pour tout $n \in \mathbb{N}$ et tout $(x_0, \dots, x_{n+1}) \in E^{n+2}$:

$$\mathbb{P}[X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0] = \mathbb{P}[X_1 = x_{n+1} | X_0 = x_n] = p_{x_n, x_{n+1}}$$

Et la propriété suivante est vérifiée :

Propriété 1. On a :

(i) P est une matrice stochastique, c'est-à-dire que la somme des termes de n'importe quelle ligne est toujours égale à 1 :

$$\forall i \in E, \sum_{j \in E} p_{i,j} = 1$$

(ii) La loi d'une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ est entièrement déterminée par sa matrice de transition P et la loi de X_0 , formellement pour tout $n \in \mathbb{N}$ et tout $(x_0, \dots, x_{n+1}) \in E^{n+2}$:

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x_n] = \mathbb{P}[X_0 = x_0] \times p_{x_0, x_1} \times \dots \times p_{x_{n-1}, x_n}$$

4.4.3 Transposition à l'étude

Dans le cadre de notre modélisation, nous ferons l'hypothèse que la trajectoire d'une tempête est une chaînes de Markov discrète et homogène notée $(T_n)_n$. Le choix d'une telle modélisation repose sur le fait que la trajectoire future d'une tempête tropicale n'est pas influencée par son passé mais seulement par sa localisation présente. En effet, comme nous avons pu le voir en section 2.2.2.2 de la première partie, ce sont essentiellement la force de Coriolis ainsi que l'orientation des vents qui guide le phénomène.

A l'instar de la base HURDAT, le passage d'un état à un autre se fera avec un pas temporel de 6 heures ce qui justifie le caractère discret de la chaîne. L'instant $n = 0$ représentera donc l'heure 0 puis l'instant $n = 1$ représentera l'heure 6 et ainsi de suite. Par ailleurs, il semble raisonnable de penser que le temps n'influence pas la probabilité de passer d'une localisation à une autre, d'où l'hypothèse d'homogénéité. Enfin, l'espace des états E sera formé par les 6600 localisations du bassin Atlantique définies en préliminaire 4.4.1 plus un état fictif représentant la mort de la tempête.

4.4.4 Initialisation des trajectoires

Le point (ii) de la propriété 1 indique que la connaissance de la loi initiale de la chaîne est un élément nécessaire (mais non suffisant) pour déterminer la loi de la chaîne. A partir de la base HURDAT on peut extraire les 369 positions initiales des tempêtes survenues lors de la période. Nous avons représenté ces positions dans la figure suivante.

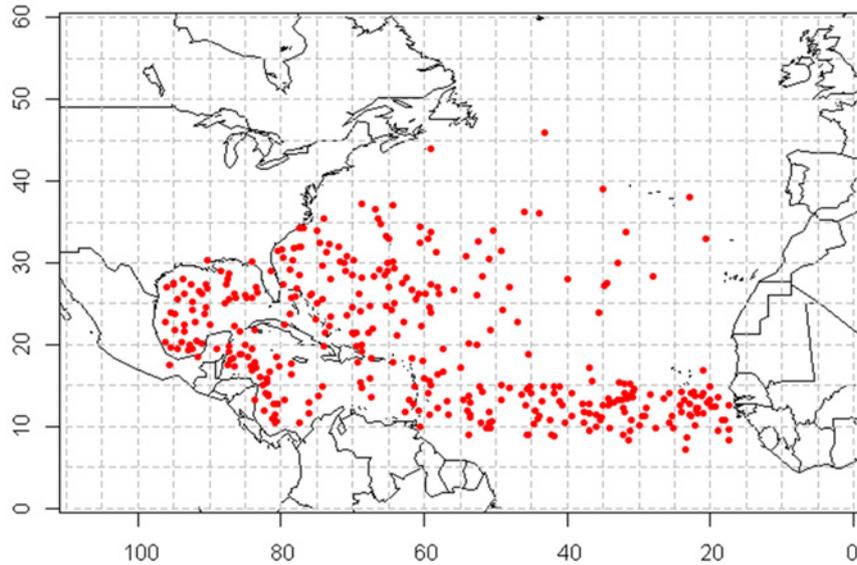


FIGURE 20 – Positions initiales des tempêtes tropicales (1982-2011)

Sur ce schéma nous retrouvons la condition géographique de formation des cyclones qui se forment au moins à 10° de latitude au dessus de l'équateur et dans les eaux chaudes de la mer des Caraïbes, du golfe du Mexique et de l'océan Atlantique même certains cas particuliers dérogent à la règle.

Après conversion des positions initiales empiriques en localisations selon la résolution $1^\circ \times 1^\circ$, la position initiale de la chaîne sera alors tout simplement tirée aléatoirement parmi cet échantillon. De plus, nous verrons dans le chapitre 5 qu'il est nécessaire d'initialiser le mois au cours duquel la tempête apparaît ainsi que la pression centrale aux instants $n = 0, 1$ et 2 . Concernant ces dernières nous prendrons les trois premières pressions centrales empiriques de la trajectoire tirée. Quant au mois, nous le génèrerons selon la distribution historique des mois de survenance (voir l'histogramme en figure 6).

4.4.5 Détermination des probabilités de transition

Le deuxième élément à estimer pour connaître la loi de $(T_n)_n$ est la matrice de transition $P = (p_{i,j})_{(i,j) \in E^2}$ de dimension 6601×6601 , soit potentiellement $43\,573\,201$ éléments. Dans la pratique une grande partie de la matrice est composé d'éléments nuls et il est difficile de stocker une matrice aussi importante dans un logiciel informatique. En outre, plusieurs lignes et colonnes sont vides du fait de l'absence d'historique pour certaines localisations.

Ainsi, nous allons uniquement se focaliser sur les localisations pour lesquelles une tempête est effectivement passée historiquement. A partir des données historiques, nous avons accès, pour un point i d'une trajectoire donnée, à la localisation suivante j de la tempête. Cette nouvelle localisation pouvant très bien correspondre à la localisation i elle-même ou à la mort de la tempête. Pour chaque état i on calcule alors la probabilité de passer à l'état j par :

$$p_{i,j} = \frac{\text{Nombre de localisations suivantes correspondant à } j}{\text{Nombre total de localisations suivantes}}$$

Par exemple pour la localisation numéro 1595, c'est-à-dire une latitude de $14,5^\circ$ et une longitude de $54,5^\circ$, nous avons :

Localisation suivante	Nombre empirique	Probabilité de transition
1596	2	0,4
1597	1	0,2
1816	1	0,2
Mort	1	0,2

Nous obtenons de cette manière, et pour chaque localisation i , une série de vecteurs de probabilités de transition $V_i = (p_{i,j_1^i}, \dots, p_{i,j_{t_i-1}^i}, p_{i,j_{t_i}^i})$ vers les localisations potentielles suivantes $(j_1^i, \dots, j_{t_i-1}^i, mort)$ où t_i est la taille du vecteur, autrement dit le nombre de localisations suivantes répertoriées, et où $\sum_{j \geq j_1^i} p_{i,j} = 1$. Une probabilité de transition nulle est associée pour toutes les localisations suivantes non répertoriées et la probabilité de passage de l'état *mort* à n'importe quel autre état est fixée à 0 ce qui caractérise la fin de la trajectoire.

4.4.6 Méthode de simulation

Nous donnons ici la méthode la plus classique afin de simuler la chaîne $(T_n)_n$. Supposons que l'on parte d'une position initiale $T_0 = i$ tirée parmi les historiques et associée au vecteur de transition V_i . On subdivise l'intervalle $]0, 1]$ en t_i intervalles de la façon suivante :

$$- I_{j_1^i} =]0, p_{i,j_1^i}]$$

- $I_{j_2^i} =]p_{i,j_1^i}, p_{i,j_1^i} + p_{i,j_2^i}]$
- ...
- $I_{mort} =]p_{i,j_1^i} + \dots + p_{i,j_{t_i-1}^i}, \underbrace{p_{i,j_1^i} + \dots + p_{i,j_{t_i}^i}}_{=1}]$

On simule alors une variable aléatoire U de loi uniforme sur $]0, 1]$.

- Si $U \in I_{j_1^i}$ alors on pose $T_1 = j_1^i$
- Sinon si $U \in I_{j_2^i}$ alors on pose $T_1 = j_2^i$
- ...
- Sinon si $U \in I_{mort}$ alors on pose $T_1 = mort$

Si T_1 correspond à l'état *mort* la simulation de la chaîne est terminée. Sinon on reprend les étapes précédentes en partant de T_1 jusqu'à atteindre l'état *mort*.

4.5 Résultats

4.5.1 Simulation de 10 000 saisons de tempêtes

La méthode décrite précédemment à été mise en œuvre afin de simuler 10 000 années de tempêtes tropicales dans le bassin Atlantique ce qui correspond à environ 123 299 trajectoires. Ce nombre correspond à une moyenne d'environ 370 trajectoires tous les trente ans, chiffre à rapprocher du nombre de tempêtes répertoriées depuis les trente dernières années (369). Nous donnons ci-dessous les tracés de 50 trajectoires tirées au hasard dans la base HURDAT et les tracés de 50 trajectoires tirées au hasard parmi les simulations.

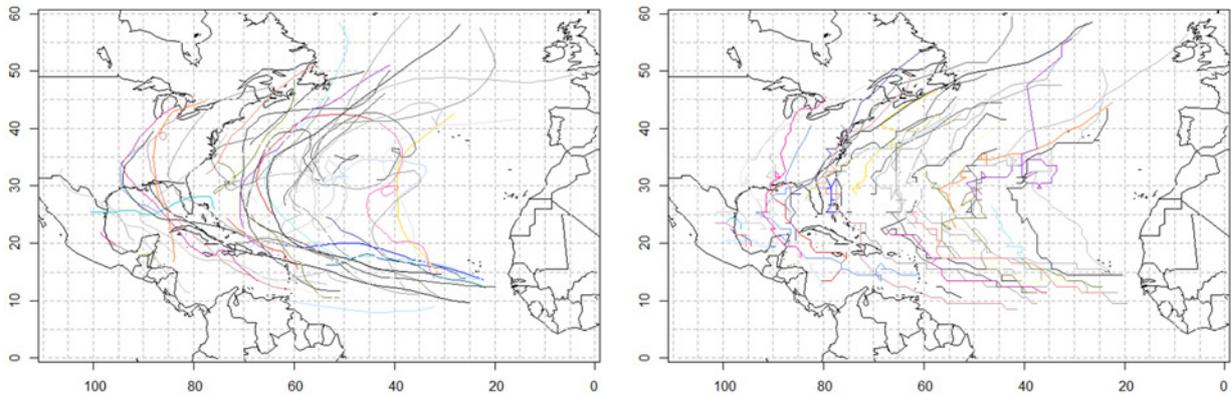


FIGURE 21 – Echantillon de 50 trajectoires empiriques (à gauche) et simulées (à droite)

Nous remarquons que les simulations captent très bien l'aspect parabolique des trajectoires observé empiriquement. En revanche, les courbes présentent des changements de directions beaucoup plus brusques avec une courbure moindre. Ceci est dû à la résolution choisie pour l'étude (quadrillage de $1^\circ \times 1^\circ$) plus faible que celle de la base HURDAT ($0, 1^\circ \times 0, 1^\circ$).

4.5.2 Confrontation avec les données empiriques

Nous nous proposons ici de confronter les trajectoires empiriques avec les simulations en effectuant quelques vérifications indispensables.

4.5.2.1 Durée des tempêtes

La durée des tempêtes simulées peut être déterminée en comptant simplement le nombre de points formant chaque trajectoires puisque l'intervalle de temps entre 2 points d'une trajectoire donnée correspond à une durée de 6 heures. Dans la figure suivante nous avons représenté les boîtes à moustaches du nombre de points formant chaque trajectoires dans le cas empirique et dans le cas des simulations :

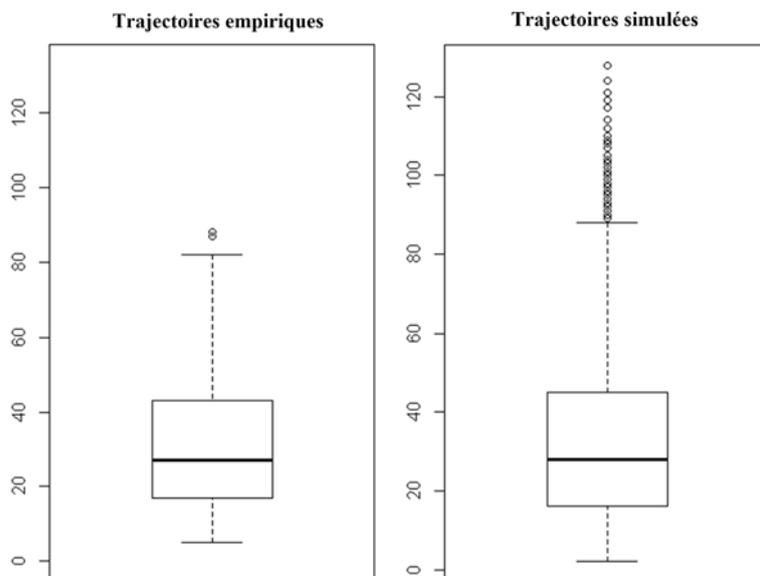


FIGURE 22 – Boîtes à moustache des nombres de points formant les trajectoires

Les statistiques descriptives associées à ces boîte à moustache sont les suivantes :

	Min	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Max
Durée empirique	5,00	17,00	27,00	31,02	43,00	88,00
Durée simulée	2,00	16,00	28,00	31,98	45,00	128,00

Nous remarquons que les simulations ont tendance à créer des durées de phénomènes cyclonique long. Cela concerne 1010 trajectoires soit environ 0,82% du nombre total généré. Dans la pratique, le record de longévité d'une trajectoire étant de 109 points (soit 27,25 jours, voir section 2.2.4 de la première partie) nous avons décidé d'utiliser ce nombre comme borne supérieure.

Nous nous proposons alors de montrer l'égalité des deux distributions de durée par un test de KOLMOGOROV-SMIRNOV dont nous rappelons brièvement le principe.

Principe du test de Kolmogorov-Smirnov : Soit (x_1, \dots, x_n) et (y_1, \dots, y_m) deux échantillons issus de la réalisations de deux variables aléatoires X et Y respectivement. On suppose que les lois de X et Y sont continues de fonctions de répartition théoriques F et G respectivement. On définit aussi leurs fonctions de répartition empiriques par :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{x_i \leq x\}}$$

et :

$$G_m(x) = \frac{1}{m} \sum_{i=1}^m \mathbb{1}_{\{y_i \leq x\}}$$

On souhaite tester l'hypothèse nulle suivante :

$$H_0 : F = G$$

Pour cela on utilise le résultat suivant :

Proposition 1. *Pour tout $c > 0$ on a la convergence suivante :*

$$\lim_{n,m \rightarrow +\infty} \mathbb{P} \left[\sup_x |F_n(x) - G_m(x)| > c \sqrt{\frac{n+m}{nm}} \right] = \alpha(c) = 2 \sum_{k=1}^{+\infty} (-1)^{k+1} \exp(-2k^2 c^2)$$

Par exemple, en prenant $c = 1,36$ on obtient $\alpha(c) = 0,05$ ce qui permet de construire un test asymptotique de niveau 5% pour la normalité. Dans la pratique, on calcule l'écart maximale entre les deux fonctions de répartition empiriques F_n et G_n en valeur absolue et on vérifie que cet écart reste inférieur à c/\sqrt{n} , ratio déterminé selon le niveau de test souhaité. Dans le cas contraire on rejettera H_0 . On trouve dans la littérature des tables donnant l'écart à ne pas dépasser en fonction de la taille des échantillons (n et m) et du niveau α du test. Ce test peut aussi être utilisé pour vérifier l'adéquation d'un échantillon à une loi théorique.

L'élaboration de ce test pour montrer l'égalité des distributions des durées empiriques et simulées donne le résultat suivant au niveau 5% :

Statistique	Seuil	p-valeur
0,0619	0,0709	0,128

La p-valeur, c'est-à-dire la probabilité de rejeter à tort l'hypothèse nulle, est de 12,8% > 5% donc nous validons le test.

4.5.2.2 Densité en latitude et longitude

En terme de fréquence d'impact, nous pouvons vérifier que les simulations sont en cohérence avec les données de la base HURDAT. Dans la figure ci-après nous avons représenté la densité des points d'impacts empiriques et celle des points issus des trajectoires simulées.

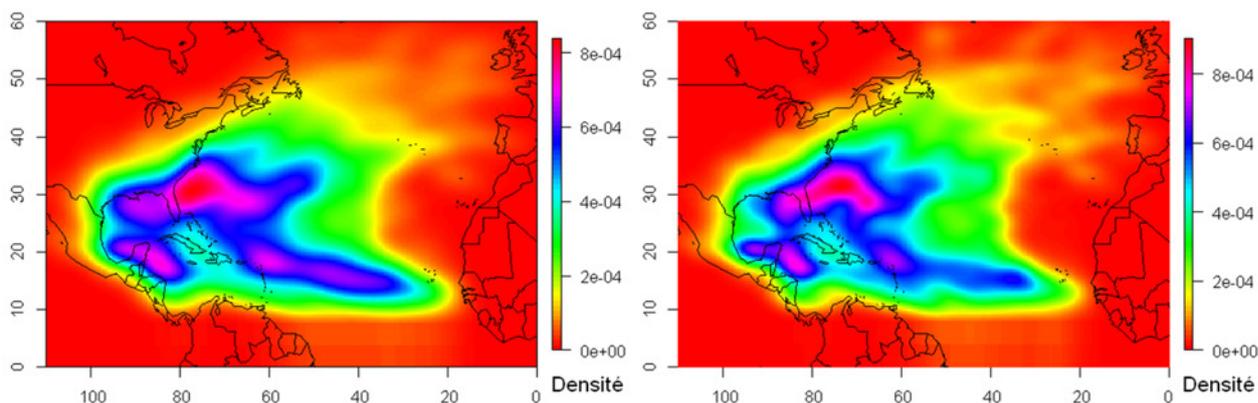


FIGURE 23 – Densités des impacts empiriques (à gauche) et simulés (à droite)

Nous observons une très grande similitude entre les densités, montrant que la chaîne de Markov respecte bien l'allure des trajectoires observées historiquement. Nous retrouvons notamment une forte concentration dans une zone comprise entre 10° et 20° de latitude à l'ouest de l'Afrique en raison du nombre important de tempêtes qui y naissent. Les caraïbes, le golfe du Mexique et

la côte Est américaine sont également particulièrement exposées.

Pour compléter ces observations, nous avons calculé les périodes de retour des tempêtes pour les zones d'intérêt de notre étude, à savoir : la Guadeloupe, la Martinique, Saint-Martin et Saint-Barthélémy. Pour cela, nous avons simplement fait le rapport entre la période considérée (30 ans pour les données empiriques et 10 000 fois un an pour les simulations) et le nombre d'évènements ayant traversé chacune des localisations $1^\circ \times 1^\circ$ associée aux quatre îles. La comparaison avec les données empiriques est là aussi probante comme le montre le tableau ci-après.

Ile	Localisation	Période de retour empirique	Période de retour simulée
Guadeloupe	1822	2,5 ans	2,3 ans
Martinique	1602	5 ans	5,2 ans
Saint-Martin	2044	2,7 ans	2,8 ans
Saint-Barthélémy	1933	3,7 ans	3,8 ans

Remarque : Il n'est pas suffisant de compter le nombre de passages par l'état de la chaîne de Markov associé à la localisation $1^\circ \times 1^\circ$. En effet, le caractère discret de la chaîne (avec un intervalle de 6 heures entre deux états) fait qu'une tempête peut très bien traverser la case associée à une île en passant, par exemple, par deux états correspondant aux cases adjacentes. En pratique, une méthode a été implémentée afin de comptabiliser ces passages en faisant l'hypothèse que la tempête se déplace de manière rectiligne entre deux localisations.

Afin d'illustrer le tableau ci-dessus, nous avons tracé l'ensemble des trajectoires traversant chacune des quatre îles assurées par GFA Caraïbes dans le graphe ci-dessous.

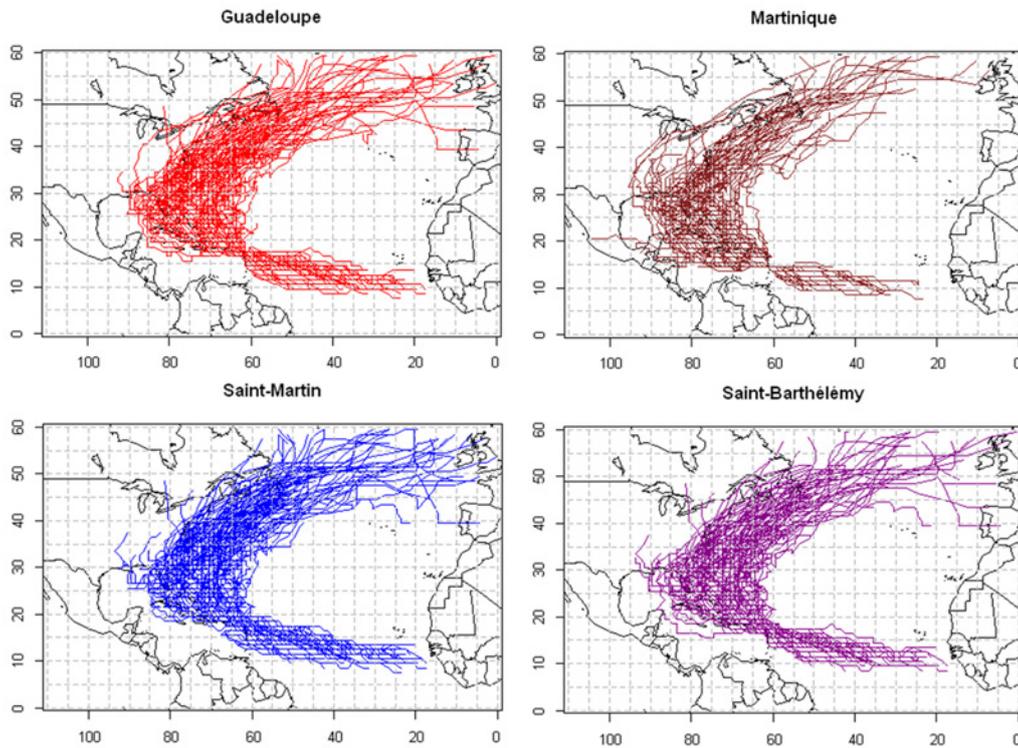


FIGURE 24 – Tempêtes simulées traversant les îles assurées

Nous observons que les tempêtes traversant les Antilles françaises prennent majoritairement leur source aux abords de la côte est du continent africain : on parle de cyclones cap-verdiens.

Chapitre 5

Détermination des caractéristiques des tempêtes le long des trajectoires simulées

Une fois que l'on est capable de simuler les trajectoires des tempêtes tropicales au cours de plusieurs saisons, l'étape suivante est de déterminer les différentes caractéristiques d'intérêt pour l'étude telles que définies en section 2.2.3 de la première partie, à savoir :

- La pression au centre de la tempête (P)
- La vitesse maximale atteinte par les vents (V_{max})
- Le rayon de vitesse maximale (R_{max})
- Le gradient de vitesse des vents

Dans ce chapitre, nous nous pencherons d'abord sur la modélisation de la pression centrale par l'intermédiaire de deux modèles selon que la tempête se trouve dans l'océan ou dans les terres. Dans le premier cas, nous utiliserons un modèle de régression basé sur la notion d'intensité relative prenant aussi la température à la surface de l'océan comme variable explicative. Dans le deuxième cas, nous considérerons que le déficit de pression se désintègre de façon exponentiel avec le temps, la tempête perdant rapidement son énergie hors de l'océan. La pression centrale étant connue, nous en déduirons la vitesse maximale des vents ainsi que le rayon associé grâce à deux modèles de régression.

Puis, un modèle de gradient de vitesse permettra d'obtenir la vitesse du vent à une distance donnée du centre de la tempête. A l'instar du chapitre précédent, ce sont les données de la base HURDAT au cours la période 1982 – 2011 qui seront à la base de chaque modélisation.

Après validation des différents modèles, leur association aux trajectoires simulées nous permettra d'obtenir la vitesse maximale à laquelle les vents atteignent les différentes îles d'intérêt pour une tempête donnée.

5.1 Méthodologie

La caractéristique clef régissant la puissance d'une tempête tropicale est le déficit de pression qu'elle possède en son centre par rapport à la pression atmosphérique, pression que nous supposerons égale à 1013 hPa. En principe, plus ce déficit est important est plus la tempête est dévastatrice. Ainsi, à l'instar de ce que propose WANG & ROSOWSKY dans [16], nous nous concentrerons sur la modélisation de la pression centrale le long des trajectoires pour ensuite en déduire les autres paramètres qui caractérisent le phénomène.

Il est important de noter que nous ne pouvons pas modéliser la pression de la même façon selon que la tempête se trouve dans l'océan ou à l'intérieur des terres. Pour illustrer cela on peut par exemple observer l'évolution de la pression au centre du cyclone Jeanne survenu en septembre 2004 :

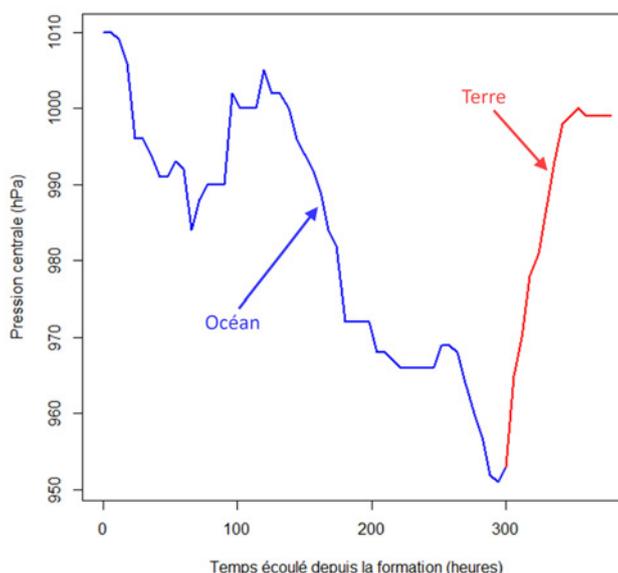


FIGURE 25 – Evolution de la pression au centre du cyclone Jeanne (2004)

Nous observons que la pression chute assez rapidement lorsque la tempête parcourt l'océan (courbe bleue) jusqu'à ce qu'elle vienne percuter la côte (on parle d'atterrissage). A partir de ce moment la pression remonte très rapidement jusqu'à atteindre un niveau beaucoup plus proche de la pression atmosphérique. Il faut tout de même noter que la pression centrale ne remonte pas uniquement lorsqu'une tempête atteint les terres. En effet, parmi les paramètres importants qui permettent à une tempête de gagner en énergie on trouve notamment la température à la surface de l'océan : l'énergie augmente dans les eaux chaudes tandis qu'elle diminue dans les eaux froides. On a donc un phénomène similaire mais en général moins important lorsque la trajectoire se trouve dans des eaux de plus en plus froides.

Cette observation a conduit à considérer deux types de modélisations différents selon que la tempête traverse l'océan, avec un modèle basé sur la l'intensité relative, ou les terres, et on parle alors de modèle de désintégration (communément appelé *decay model*). En outre, si une tempête retourne dans l'océan après avoir franchi des terres nous reprendrons le modèle valable dans l'océan. Cependant, ce dernier nécessitant 3 pas de 6 heures pour l'initialisation (comme nous le verrons dans la section suivante) il ne sera repris que si la tempête se trouve dans l'océan depuis ce même nombre de pas. Nous résumons cette méthodologie dans la figure suivante :

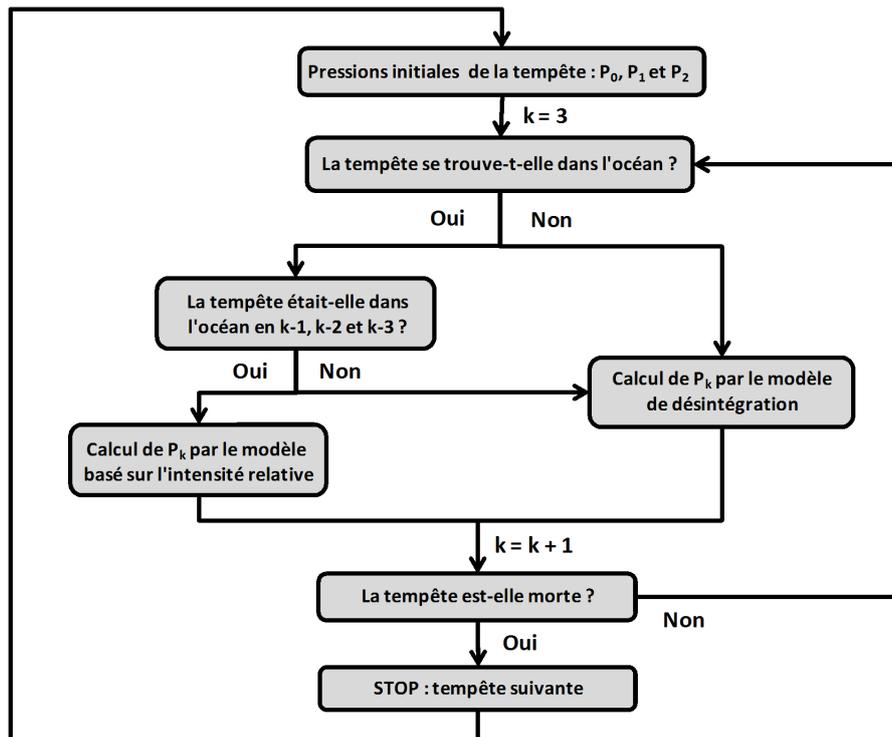


FIGURE 26 – Algorithme de modélisation de la pression centrale le long d’une trajectoire

La différenciation entre océan et terre se fera avec une précision de 1° de carré en latitude et longitude comme le montre la figure ci-après et dans laquelle les cercles bleus représentent les cases 1° × 1° considérées comme faisant partie de l’océan.

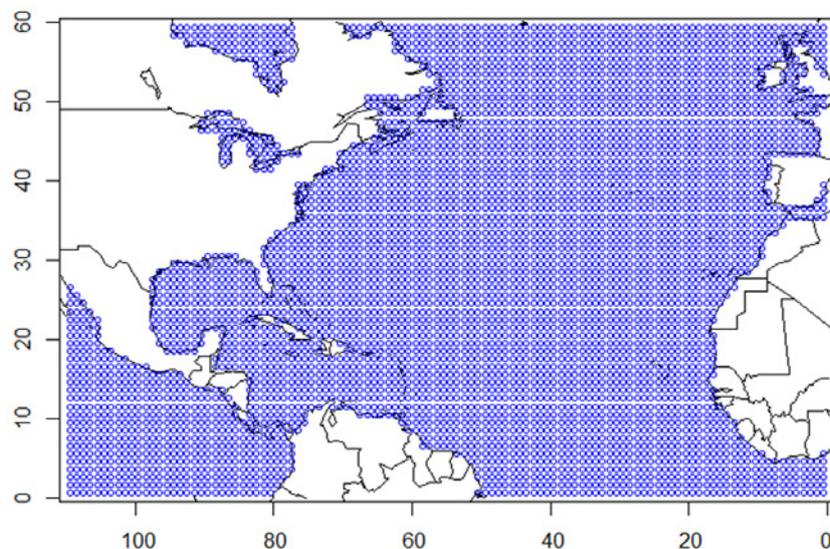


FIGURE 27 – Filtre océan-terre utilisé

Dans la suite, nous explicitons les deux modélisations retenues ainsi que le passage à la vitesse maximale, au rayon maximal et au gradient de vitesse.

5.2 Modèle de pression dans l'océan

Dans l'océan, le modèle retenu est basé sur un modèle linéaire de régression multiple dont nous rappelons quelques bases théoriques en se limitant à celles qui seront utilisées pour le développement du modèle.

5.2.1 Rappels sur les modèles de régression multiples

5.2.1.1 Définition et objectifs

L'objectif d'un modèle de régression multiple est de modéliser une variable aléatoire réelle quantitative Y , appelée variable à expliquer, en fonction de plusieurs variables réelles déterministes X_1, X_2, \dots, X_p , appelée variables explicatives.

Supposons que l'on observe simultanément n observations pour chaque variable X_j et que Y est réalisées n fois de manière indépendante. Pour la i -ème observation, les valeurs observées des variables seront alors notées $x_{1,i}, x_{2,i}, \dots, x_{p,i}$ et y_i , alors que Y_i sera la composante aléatoire de la variable Y .

Définition 3 (Modèle linéaire multiple). *On appelle modèle linéaire multiple un modèle statistique consistant à relier Y_i avec $x_{1,i}, x_{2,i}, \dots, x_{p,i}$ par la relation suivante :*

$$Y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{j,i} + \varepsilon_i, \quad i = 1, \dots, n \quad (1)$$

où :

- les β_j ($j = 1, \dots, p$) sont les paramètres du modèle à estimer
- ε_i est un terme d'erreur supposé être une variable aléatoire réelle

Si on pose $X = \begin{pmatrix} 1 & x_{1,1} & \cdots & x_{p,1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{1,n} & \cdots & x_{p,n} \end{pmatrix}$, $Y = {}^t(Y_1, \dots, Y_n)$, $\beta = {}^t(\beta_0, \dots, \beta_n)$ et $\varepsilon = {}^t(\varepsilon_1, \dots, \varepsilon_n)$ (où t désigne la transposée) on peut écrire (1) sous la forme matricielle suivante :

$$Y = X\beta + \varepsilon \quad (2)$$

En terme de notations, nous désignerons respectivement par Σ_X et I_n la matrice de variance-covariance d'un vecteur aléatoire X quelconque et la matrice identité d'ordre n .

5.2.1.2 Hypothèses

Nous ferons les hypothèses classiques suivantes :

$H1$: $\mathbb{E}[\varepsilon] = 0 \Rightarrow \mathbb{E}[Y] = X\beta$: les erreurs sont nulles en moyenne

$H2$: $\mathbb{E}[\varepsilon^t \varepsilon] = \Sigma_\varepsilon = \sigma^2 I_n \Rightarrow \Sigma_Y = \sigma^2 I_n$: absence d'autocorrélation des résidus et variance constante

$H3$: ε est de loi normale multivariée d'ordre n notée $\mathcal{N}_n(0, \sigma^2 I_n)$

$H4$: $\text{Rang}(X) = p + 1 < n$: le nombre d'observations est supérieur au nombre de paramètres

L'hypothèse $H2$ est souvent appelée hypothèse d'homoscédasticité du modèle, c'est-à-dire qu'il y a uniformité de la variance dans l'erreur de régression. L'absence d'autocorrélation traduit elle le caractère vraiment aléatoire des résidus qui ne doivent pas présenter de structure particulière. Plusieurs causes peuvent expliquer une autocorrélation : oubli d'une variable explicative dans le modèle, liaison entre les variables non linéaire, etc.

5.2.1.3 Estimation des paramètres

Les deux méthodes les plus utilisées afin d'estimer les paramètres $(\beta_0, \dots, \beta_p)$ sont les moindres carrés ordinaires et le maximum de vraisemblance. Vu l'hypothèse de normalité nous choisissons d'explicitier la deuxième approche.

Soit $y = {}^t(y_1, \dots, y_n)$ l'échantillon de valeurs observées pour la variable Y . L'hypothèse $H3$ implique que $Y \sim \mathcal{N}_n(X\beta, \sigma^2 I_n)$ dont la fonction de vraisemblance est donnée par :

$$\mathcal{V}(\beta, \sigma^2, y) = (2\pi)^{-\frac{n}{2}} [\det(\sigma^2 I_n)]^{-\frac{1}{2}} \exp \left[-\frac{{}^t(y - X\beta)(\sigma^2 I_n)^{-1}(y - X\beta)}{2} \right]$$

d'où la fonction de log-vraisemblance associée :

$$\mathcal{L}(\beta, \sigma^2, y) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} {}^t(y - X\beta)(y - X\beta)$$

On cherche alors $\hat{\beta}$ et $\hat{\sigma}^2$ estimateurs de β et σ^2 tels que :

$$(\hat{\beta}, \hat{\sigma}^2) = \arg \max_{\beta, \sigma^2} \mathcal{L}(\beta, \sigma^2, y) \quad (3)$$

ce qui revient à résoudre les équations de vraisemblance du problème données par :

$$\begin{cases} \left. \frac{\partial \mathcal{L}(\beta, \sigma^2, y)}{\partial \beta} \right|_{\beta=\hat{\beta}} = 0 \\ \left. \frac{\partial \mathcal{L}(\beta, \sigma^2, y)}{\partial \sigma^2} \right|_{\sigma^2=\hat{\sigma}^2} = 0 \end{cases}$$

et qui mènent à :

$$\begin{cases} \hat{\beta} = ({}^t X X)^{-1} {}^t X Y \\ \hat{\sigma}^2 = \frac{{}^t(y - X\hat{\beta})(y - X\hat{\beta})}{n} \end{cases}$$

On peut montrer que la matrice hessienne de $\mathcal{L}(\beta, \sigma^2, y)$ aux points $\hat{\beta}$ et $\hat{\sigma}^2$ est semi-définie négative ce qui permet de conclure que le maximum est bien atteint en ces points.

Propriété 2. On a :

- (i) $\hat{\beta}$ est un estimateur convergent et sans biais de β identique à celui de la méthode des moindres carrés ordinaires
- (ii) $\hat{\beta}$ est l'estimateur de variance minimale parmi tous les estimateurs linéaires sans biais de β (Gauss-Markov)
- (iii) $S^2 = \frac{n}{n-2} \hat{\sigma}^2$ est un estimateur convergent et sans biais de σ^2

Dans la suite nous noterons :

- $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$: moyenne empirique de Y
- $\hat{y}_i = y_i - \left(\hat{\beta}_0 + \sum_{j=1}^p \hat{\beta}_j \right)$: valeurs ajustées ou valeurs prédites par le modèle
- $\hat{\varepsilon}_i = y_i - \hat{y}_i$: résidus de la régression

5.2.1.4 Qualité de l'ajustement

La mise en œuvre d'un modèle linéaire a pour objectif d'expliquer la variabilité de la variable Y par d'autres variables. Une mesure intuitive de la qualité du modèle est donc de mesurer la part de variance de Y expliquée par le modèle. Posons :

- $SCT = \sum_{i=1}^n (y_i - \bar{y})^2$: variation totale du modèle ou somme des carrés totale
- $SCE = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$: variation expliquée par le modèle ou somme des carrés de la régression
- $SCR = \sum_{i=1}^n \hat{\epsilon}_i^2$: variation résiduelle ou somme des carrés des résidus

On peut montrer que :

$$SCT = SCE + SCR$$

On définit alors le coefficient de détermination R^2 par :

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$$

Ce coefficient est compris entre 0 et 1, une bonne régression étant caractérisée par un R^2 proche de 1 caractéristique d'un haut pourcentage de variance expliqué par le modèle. Néanmoins, ce coefficient a tendance à augmenter lorsque le nombre de variables explicatives augmente. Pour pallier à ce défaut on définit le coefficient de détermination ajusté par :

$$R_{adj}^2 = 1 - \frac{n-1}{n-p-1} \times \frac{SCR}{SCT}$$

5.2.1.5 Test de significativité globale

Une fois que les paramètres $(\beta_0, \dots, \beta_p)$ sont estimés, on peut effectuer un test visant à vérifier que le modèle est globalement significatif, autrement dit que l'ensemble des p variables explicatives apportent pas une information utile pour la prédiction de Y . Ceci revient à tester l'hypothèse nulle suivante :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

On construit alors un test en utilisant le résultat suivant :

Proposition 2. *Sous H_0 , la statistique :*

$$F = \frac{(\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n \hat{\epsilon}_i^2) / p}{(\sum_{i=1}^n \hat{\epsilon}_i^2) / (n - p - 1)} = \frac{n - p - 1}{p} \times \frac{R^2}{1 - R^2}$$

suit une loi de Fisher-Snedecor à p et $n - p - 1$ degrés de libertés notée $\mathcal{F}_{p, n-p-1}$.

Cette statistique peut s'interpréter comme l'erreur relative commise en supposant H_0 vraie. Pour un test de niveau α , la région de rejet du test est donc donnée par :

$$R = \left\{ F > f_{p, n-p-1}^{1-\alpha} \right\}$$

où $f_{p, n-p-1}^{1-\alpha}$ le quantile d'ordre $1 - \alpha$ d'une loi de Fisher-Snedecor à p et $n - p - 1$ degrés de libertés.

5.2.1.6 Validation du modèle

Une fois le modèle mis en oeuvre, on doit vérifier *a posteriori* que l'ensemble des hypothèses portant sur le terme d'erreur ε et définies en 5.2.1.2 sont vérifiées : normalité, moyenne nulle, absence d'autocorrélation et homoscédasticité. Chacune de ces hypothèses peut être validée par un test statistique approprié.

Test de normalité des résidus : On veut tester l'hypothèse suivante :

$$H_0 : \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Pour cela on utilisera le test de Kolmogorov-Smirnov explicité dans la section 4.5.2.1.

Test d'absence d'autocorrélation des résidus :

Afin de tester une éventuelle dépendance entre les erreurs issues de différentes observations on peut utiliser le test de Durbin-Watson qui détecte une autocorrélation d'ordre 1. Soit $(\varepsilon_1, \dots, \varepsilon_n)$ un échantillon de résidus, on cherche à tester si on peut écrire :

$$\varepsilon_t = \rho\varepsilon_{t-1} + \nu_t \quad \text{pour } t = 2, \dots, n$$

avec $\nu_t \sim \mathcal{N}(0, \sigma_{\nu_t}^2)$. L'hypothèse nulle du test est l'absence d'autocorrélation d'ordre 1, autrement dit :

$$H_0 : \rho = 0$$

La statistique du test s'écrit :

$$DW = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}$$

Par construction, cette statistique est comprise entre 0 et 4. Pour un niveau α et un nombre de variables explicatives donnés une table théorique donne les deux valeurs critiques du test d_1 et d_2 à comparer avec DW . Plusieurs cas sont possibles :

- Si $d_2 < DW < 4 - d_2$ on accepte H_0
- Si $0 < DW < d_1$ on rejette H_0 et $\rho > 0$
- Si $4 - d_1 < DW < 4$ on rejette H_0 et $\rho < 0$
- Si $d_1 < DW < d_2$ ou $4 - d_2 < DW < 4 - d_1$ on ne peut pas conclure

Test d'homoscédasticité des résidus : Si l'hypothèse d'homoscédasticité n'est pas satisfaite l'estimateur du maximum de vraisemblance n'est plus à variance minimale : il existe un estimateur sans biais plus précis des paramètres. Cette hypothèse peut être testée par un test de White que nous décrivons brièvement. Le principe est de régresser la variance selon :

$$\text{Var}[\varepsilon_i] = \sigma^2 + \gamma \mathbb{E}[Y_i]^2 = \sigma^2 + \gamma(\beta_0 + \beta_1 x_{i,1} + \dots + \beta_p x_{i,p})^2 + \mu_i, \quad i = 1, \dots, n$$

avec $\mu_i \sim \mathcal{N}(0, \sigma_{\mu_i}^2)$. L'hypothèse nulle correspondant à l'homoscédasticité s'écrit :

$$H_0 : \gamma = 0$$

En pratique, on régresse le carré des résidus sur les variables explicatives, leur carré et leur produit puis on calcule le coefficient de détermination R^2 de la régression. Sous H_0 , on peut montrer que nR^2 suit asymptotiquement une loi du khi-deux dont le nombre de degrés de liberté est égal au nombre de variables explicatives utilisées pour régresser ε_i^2 . On peut alors construire une région de rejet de l'hypothèse nulle.

5.2.2 Le modèle de régression de Vickery

La modélisation qui va être développée pour déterminer la pression centrale d'une tempête le long d'une trajectoire se base sur l'utilisation couplée de deux variables explicatives : l'intensité relative et la température mensuelle moyenne à la surface de l'océan. Dans la suite, nous présentons d'abord ces deux variables puis le modèle de régression retenu.

5.2.2.1 Notion d'intensité relative

Le concept d'intensité relative a d'abord été développé par EMMANUEL [6] puis reprise par DARLING [3] pour la modélisation des cyclones tropicaux. L'intensité relative est une mesure physique sans dimension définie comme le ratio entre la pression centrale de la tempête à un instant donné divisé par la pression centrale minimale potentiellement atteignable étant données les conditions thermodynamiques à cet instant. Cette notion est basée sur l'efficacité d'un système thermodynamique comme peut l'être un cyclone.

L'approche retenue pour déterminer l'intensité relative sera celle développée par DARLING explicitée dans l'annexe A. Dans cette approche, il existe une relation simple permettant de passer de l'intensité relative à la pression centrale et vice versa connaissant la température mensuelle moyenne à la surface de l'eau (voir section suivante). L'avantage de l'utilisation de l'intensité relative plutôt que des pressions centrales brutes réside dans le fait que, par construction, l'intensité relative est comprise entre 0 et 1. Ainsi, on ne risque pas d'obtenir des pressions centrales aberrantes lors de l'opération d'inversion.

Cependant, lorsque l'on convertit les pressions centrales de la base HURDAT en intensité relative il peut arriver que cette dernière dépasse la valeur 1 comme l'a justement remarqué DARLING. C'est notamment le cas lorsqu'un cyclone se déplace très rapidement dans les hautes latitudes de l'océan Atlantique. De plus, nous utilisons des températures mensuelles moyennes alors qu'il faudrait théoriquement utiliser les températures à chaque instant. Pour pallier à cela, nous bornerons l'intensité relative à l'intervalle $[0, 01, 1]$.

Nous donnons dans la figure ci-après l'évolution de l'intensité relative, de la pression centrale et de la température mensuelle moyenne à la surface de l'océan pour le cyclone Andrew survenu en 1992 (avec un pas temporel de 6 heures).

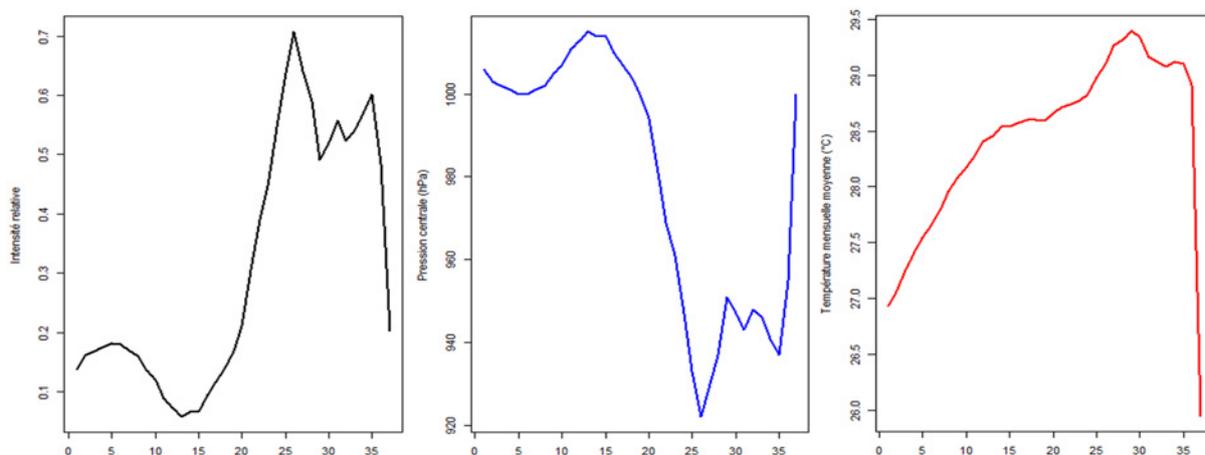


FIGURE 28 – Intensité relative, pression centrale et température mensuelle moyenne à la surface de l'océan pour Andrew (1992)

Cet exemple est parfaitement caractéristique de ce que l'on observe en règle générale, à savoir une intensité peu élevée à la formation du phénomène puis une augmentation progressive pendant la maturité due à l'énergie engendrée par la hausse de la température de l'eau et enfin une décroissance très rapide lorsque la tempête se trouve dans des eaux plus froides jusqu'à sa dissipation. Quant à la pression centrale, elle évolue de façon complètement inverse par rapport à l'intensité relative, les très basses pressions caractérisant les cyclones les plus violents.

5.2.2.2 Températures mensuelles moyennes à la surface de l'océan

Le calcul de l'intensité relative ainsi que le développement du modèle linéaire qui va être développé dans la section suivante nécessite la connaissance de la température à la surface de l'océan. Pour l'étude nous nous limiterons aux températures mensuelles moyennes que l'on peut trouver sur le site internet de l'agence américaine NOAA (*National Oceanic and Atmospheric Administration*)¹⁹. Les températures sont données pour l'ensemble des eaux du globe avec une précision de 1° de carré en latitude et longitude depuis décembre 1981. Pour fixer les idées nous avons représenté dans la figure suivante la température moyenne pour le mois d'août 2011 en respectant le filtre donné précédemment.

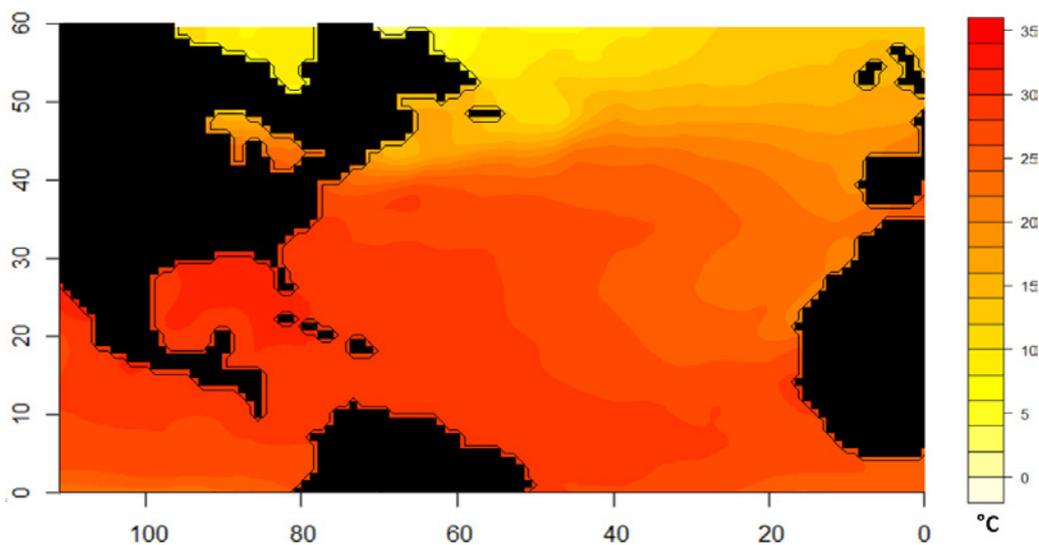


FIGURE 29 – Température moyenne à la surface de l'océan Atlantique Nord (août 2011)

Cette carte permet de comprendre l'évolution de l'intensité d'un cyclone tropical généralement constatée. Ainsi, partant des côtes africaines où la température de surface est suffisante à la formation du phénomène (plus de 28°C) la tempête gagne en intensité en se rapprochant de la mer des Caraïbes du fait de l'accroissement de la température. Les cyclones atteignent alors leur intensité maximale. Ensuite, en prenant la direction du nord où les températures sont plus froides le phénomène perd en intensité en raison de l'insuffisance de chaleur pour finalement se dissiper.

5.2.2.3 Présentation du modèle

Le modèle retenu pour modéliser la pression centrale le long des trajectoires lorsqu'une tempête se déplace dans l'océan est le modèle de régression multiple proposé par VICKERY & al. [15]. Ces derniers proposent d'utiliser les deux variables explicatives explicitées précédemment même si il faut garder à l'esprit que la pression centrale est une fonction complexe prenant en

¹⁹. <http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.oisst.v2.html>

argument d'autres paramètres tels que l'âge de la tempête ou encore la force de cisaillement.

Comme précédemment, on suppose que la durée entre deux points d'une trajectoire donnée est de 6 heures et on localise chaque point d'une trajectoire selon un quadrillage $1^\circ \times 1^\circ$. On se place en un point k quelconque ($k \in \mathbb{N}$) et on cherche à déterminer la pression centrale de la tempête au point $k + 1$ suivant. Pour cela, les auteurs proposent de calculer l'intensité relative en $k + 1$ en utilisant le modèle de régression multiple suivant :

$$\ln(I_{k+1}) = c_0 + c_1 \ln(I_k) + c_2 \ln(I_{k-1}) + c_3 \ln(I_{k-2}) + c_4(T_{k+1} - T_k) + c_5 T_{k+1} + \varepsilon$$

où :

- I_{k+1} est l'intensité relative en $k + 1$, autrement dit la variable à expliquer
- I_k , I_{k-1} et I_{k-2} sont les intensités relatives aux instants k , $k - 1$ et $k - 2$
- T_{k+1} et T_k sont les températures moyennes mensuelles à la surface de l'océan aux instants $k + 1$ et k
- c_0 , c_1 , c_2 , c_3 , c_4 et c_5 sont les coefficients de la régression qui diffèrent selon la localisation de la tempête en $k + 1$
- ε est un terme d'erreur satisfaisant les hypothèses définies en 5.2.1.2

Ce modèle rentre bien dans la catégorie des modèles de régression multiples tels que nous les avons défini précédemment avec 5 variables explicatives. Remarquons aussi que l'utilisation de ce modèle nécessite une initialisation déjà évoquée, à savoir la connaissance des intensités relatives (ou des pressions centrales) aux 3 premiers points de la trajectoire à savoir I_0 , I_1 et I_2 ainsi que des températures moyennes T_0 , T_1 et T_2 .

La particularité de ce modèle réside dans le fait que les paramètres de la régression c_0 , c_1 , c_2 , c_3 , c_4 et c_5 évoluent en même temps que la tempête se déplace. Il est donc nécessaire d'effectuer un nouveau quadrillage du bassin Atlantique, chaque case ayant son propre jeu de paramètres. Cependant on ne peut pas utiliser un quadrillage aussi fin que dans le cas de la modélisation des trajectoires car l'historique trop faible de chaque case $1^\circ \times 1^\circ$ ne permettrait pas d'effectuer la régression. En effet, une hypothèse importante du modèle de régression multiple indique que le nombre d'observations doit être strictement supérieur au nombre de paramètres à estimer (ici 6 paramètres).

Dans le cadre de l'étude, nous considérerons qu'en deçà de 10 observations le nombre de données est insuffisant pour faire la régression. Ainsi, l'analyse des 30 années d'historique de la base HURDAT a conduit à opter pour un quadrillage de résolution $2,5^\circ \times 2,5^\circ$ pour la région des Caraïbes et le golfe du Mexique, où un historique important est disponible, et un quadrillage de résolution $5^\circ \times 5^\circ$ ailleurs. A l'instar de la partie sur les chaînes de Markov, chaque localisation est repérée par le centre de la case qui lui est associée. Dans la figure ci-dessous nous explicitons la résolution choisie en fonction de la localisation.

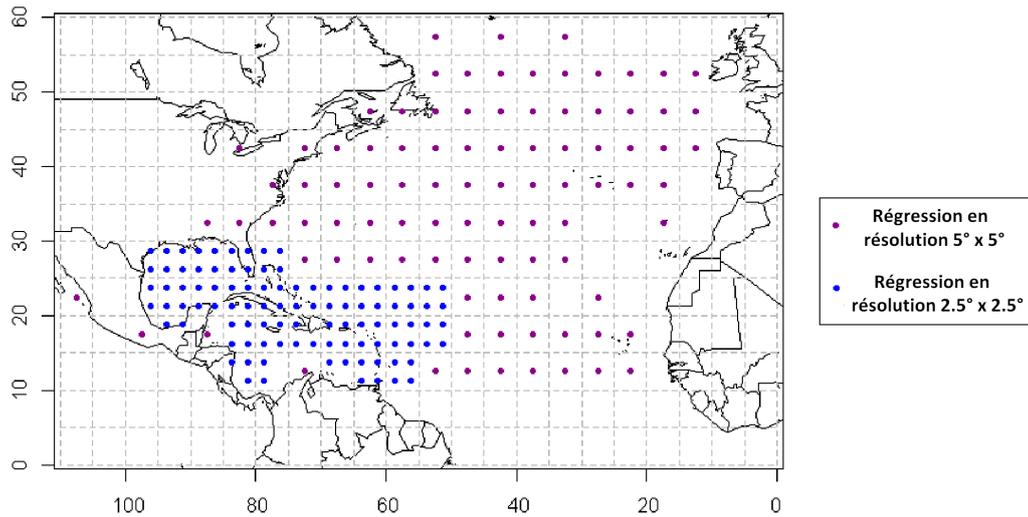


FIGURE 30 – Résolution choisie pour la détermination des coefficients de la régression

Certaines localisations se trouvant dans l’océan ont été volontairement laissées vides car la régression y est impossible par manque de données. Ces dernières seront traitées dans la section 5.2.4. Au total nous devons effectuer 93 régressions en résolution $5^\circ \times 5^\circ$ et 100 régressions en résolution $2,5^\circ \times 2,5^\circ$. Les nombres d’observations disponibles pour chacune d’entre elles sont données dans les deux graphes ci-après.

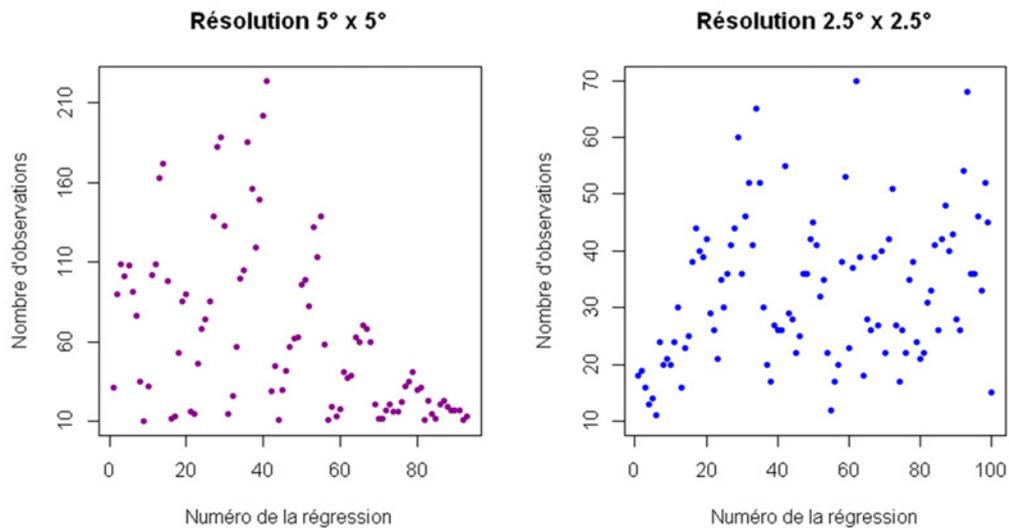


FIGURE 31 – Nombre de données disponibles pour effectuer les régressions

Une fois que les paramètres sont estimés, on peut déterminer les intensités relatives en chaque points des trajectoires hors des terres puis, en inversant la fonction donnant accès à l’intensité relative (voir annexe A) on obtient les pressions centrales le long de celles-ci. En ce qui concerne les températures mensuelles moyennes nous avons retenues celles de 2011 après avoir vérifié que ces dernières ne correspondaient pas à une année exceptionnelle. La démarche est résumée dans la figure suivante.

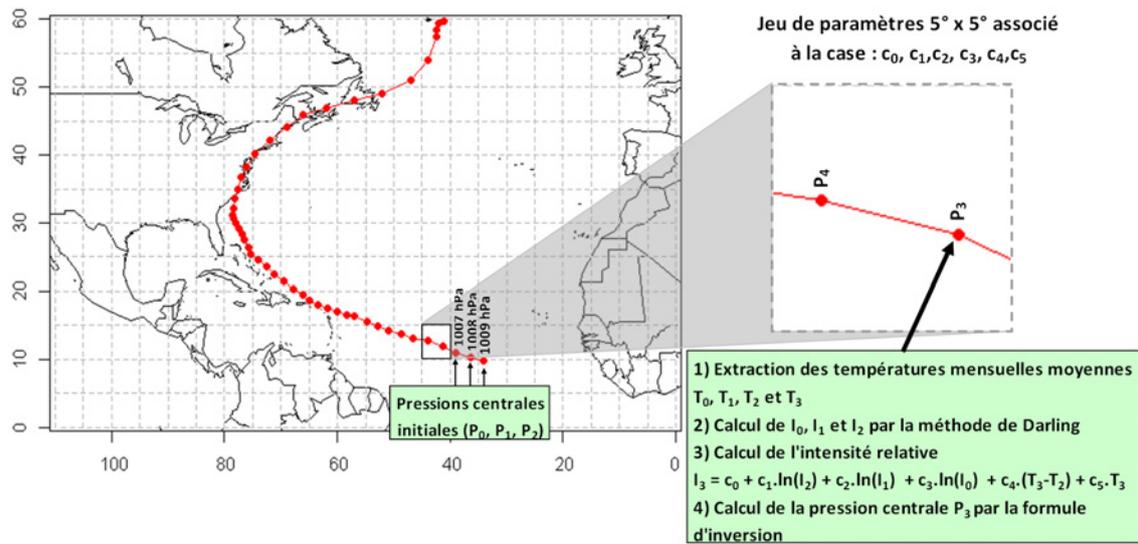


FIGURE 32 – Détermination de la pression centrale le long des trajectoires dans l'océan

5.2.3 Application du modèle et analyse des résidus

Une régression a été réalisée à partir des données de la base HURDAT pour chaque case $5^\circ \times 5^\circ$ ou $2,5^\circ \times 2,5^\circ$ selon la localisation. Les paramètres ont été estimés grâce au logiciel R *via* la méthode du maximum de vraisemblance. La validation de chaque régression passe par la vérification de la qualité de l'ajustement, de la significativité globale du modèle et des hypothèses propres au terme d'erreur ε . Pour vérifier cela nous avons utilisé les tests présentés dans la section 5.2.1 avec un niveau de 5%. Les résultats sont présentés dans les deux figures ci-après et pour les deux types de résolution. Les hachures rouges représentent les zones de rejet des tests.

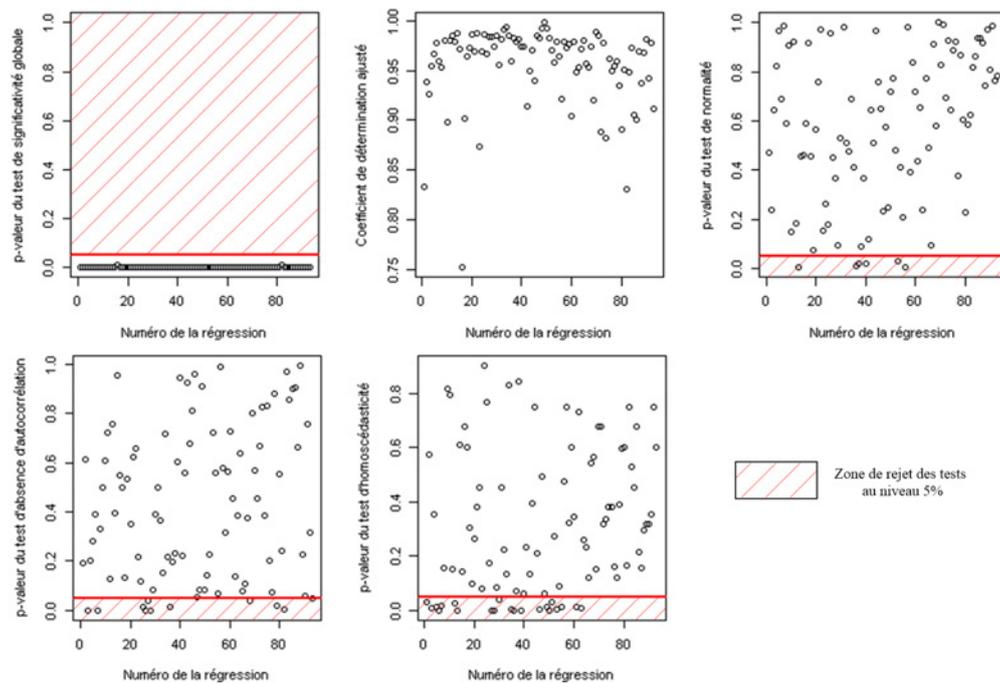


FIGURE 33 – Résultat des tests pour les régressions en résolution $5^\circ \times 5^\circ$

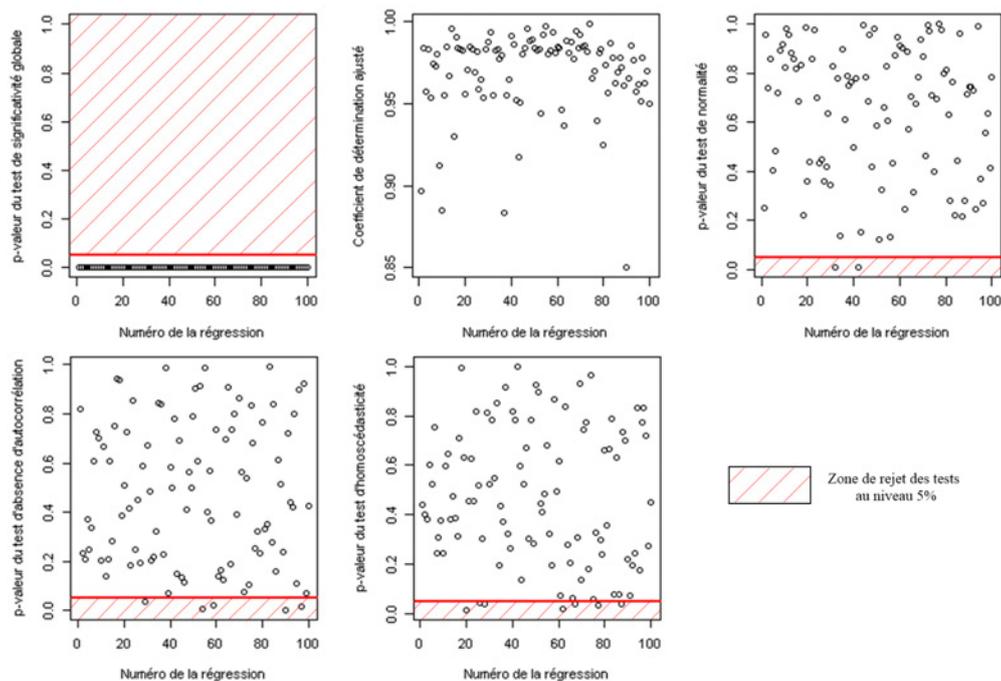


FIGURE 34 – Résultat des tests pour les régressions en résolution $2,5^\circ \times 2,5^\circ$

Nous pouvons synthétiser ces résultats par le tableau suivant qui donne les pourcentages de cases réussissant avec succès les différents tests statistiques pour les deux types de résolution :

Test	$5^\circ \times 5^\circ$	$2,5^\circ \times 2,5^\circ$
R^2 ajusté $\geq 0,75$	100%	100%
p-valeur $\leq 0,05$ pour le test de significativité global	100%	100%
p-valeur $\geq 0,05$ pour le test de normalité des résidus	90%	98%
p-valeur $\geq 0,05$ pour le test d'absence d'autocorrélation des résidus	89%	95%
p-valeur $\geq 0,05$ pour le test d'homoscédasticité des résidus	75%	93%

Au vu des résultats, nous pouvons affirmer que le modèle a toujours un sens puisque le test de significativité est systématiquement positif. L'absence d'autocorrélation est en générale constatée et dans le cas contraire cela s'explique par l'absence d'une variable explicative puisque le modèle proposé par VICKERY ne prend pas en compte tous les phénomènes qui régissent la pression centrale de la tempête.

Dans la pratique, nous avons conservé tous les résultats obtenus puisque les ajustements sont de bonne qualité avec un R^2_{adj} qui ne descend jamais en dessous de 75%. En outre, les résultats sont très bons dans la zone en résolution $2,5^\circ \times 2,5^\circ$, zone concernée par notre étude. Concernant les termes d'erreur aléatoires associés à chaque case, nous les modéliserons par des lois normales dont les paramètres sont estimés par maximum de vraisemblance pour chacune des régressions.

5.2.4 Traitement des localisations avec peu ou pas d'historiques

Il s'agit ici de traiter le cas des localisations où l'élaboration d'une régression n'a pas été possible en raison d'un manque de données historiques. Pour cela, il a été choisi d'attribuer à ces localisations le même jeu de paramètres que la case la plus proche.

On considère une localisation i quelconque à paramétrer. On repère cette localisation par la latitude (notée Lat_i) et la longitude (notée $Long_i$) du milieu de la case qui lui est associée en

degrés. Ensuite, on détermine la distance entre ce point et toutes les autres localisations j en supposant la Terre parfaitement sphérique de rayon $R_T = 6\,400$ km. La formule utilisée est la suivante pour tout $j \neq i$:

$$d(i, j) = R_T \times \arccos \left[\sin \left(\frac{\pi \cdot Lat_i}{180} \right) \sin \left(\frac{\pi \cdot Lat_j}{180} \right) + \cos \left(\frac{\pi \cdot Lat_i}{180} \right) \cos \left(\frac{\pi \cdot Lat_j}{180} \right) \cdot \cos \left(\frac{\pi \cdot (Long_i - Long_j)}{180} \right) \right]$$

Cette distance est appelée distance orthodromique et correspond à la plus courte distance entre deux points d'une sphère repérés par leur latitude et leur longitude. La démonstration de la formule est proposée en annexe B.

La localisation la plus proche de i (notée $loc(i)$) correspond alors à :

$$loc(i) = \arg \min_j d(i, j)$$

Dans la pratique, il peut arriver de trouver plusieurs localisations minimisant la distance. Dans ce cas, chaque paramètre de la case i est calculé comme étant la moyenne des coefficients des localisations les plus proches.

En ce qui concerne la distribution du terme d'erreur aléatoire, nous supposons qu'elle correspond à la moyenne des distributions gaussiennes associées aux localisations adjacentes minimisant la distance avec i . Les hypothèses sous-jacentes étant l'indépendance des termes d'erreurs et la stabilité de la loi normale par moyenne. On rappelle que si on note $(\varepsilon_1, \dots, \varepsilon_n)$ ces erreurs avec $\varepsilon_1 \sim \mathcal{N}(0, \sigma_1^2), \dots, \varepsilon_n \sim \mathcal{N}(0, \sigma_n^2)$ on a :

$$\frac{\varepsilon_1 + \dots + \varepsilon_n}{n} \sim \mathcal{N} \left(0, \frac{\sigma_1^2 + \dots + \sigma_n^2}{n^2} \right)$$

5.2.5 Confrontation avec les données empiriques

A partir du modèle de VICKERY appliqué aux données empiriques (sans le terme d'erreur aléatoire) nous avons calculé la moyenne des pressions centrales passant par chaque localisation en résolution $1^\circ \times 1^\circ$. On peut alors confronter ces résultats aux moyennes observées empiriquement.

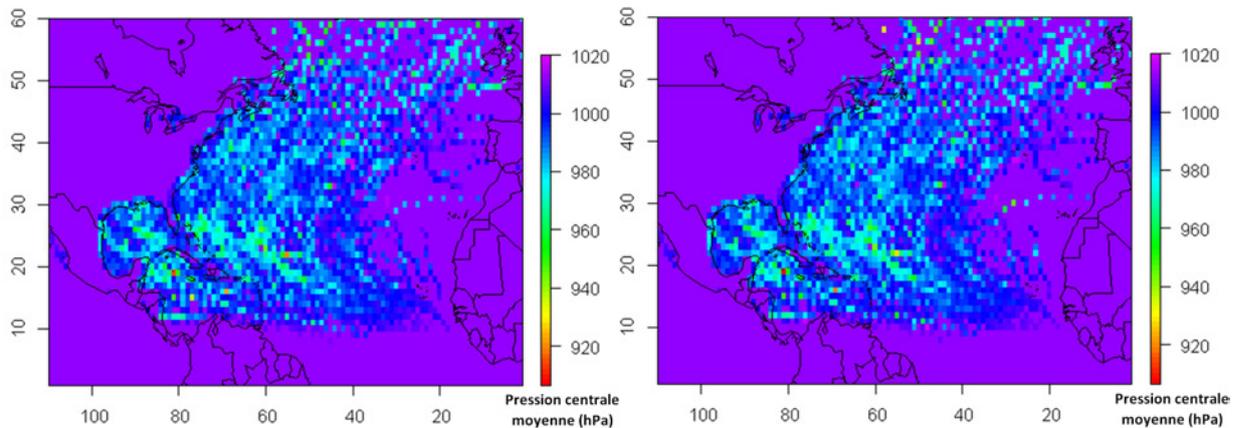


FIGURE 35 – Pressions centrales moyennes empiriques (à gauche) et modélisées (à droite)

Nous observons une grande similitude entre les deux figures avec des pressions centrales moyennes élevées à proximité de la côte Ouest africaine et les pressions centrales les plus faibles dans les Caraïbes et le golfe du Mexique. Visuellement, la régression semble donc être très bien adaptée pour modéliser la pression centrale des tempêtes lors de la traversée d'étendues d'eau.

On peut aussi comparer les histogrammes des pressions centrales issues de la base HURDAT et de celles issues du modèle de VICKERY afin de vérifier que la cohérence des fréquences en terme d'intensité. Ces histogrammes sont représentés ci-dessous ainsi que leur densités par lissage.

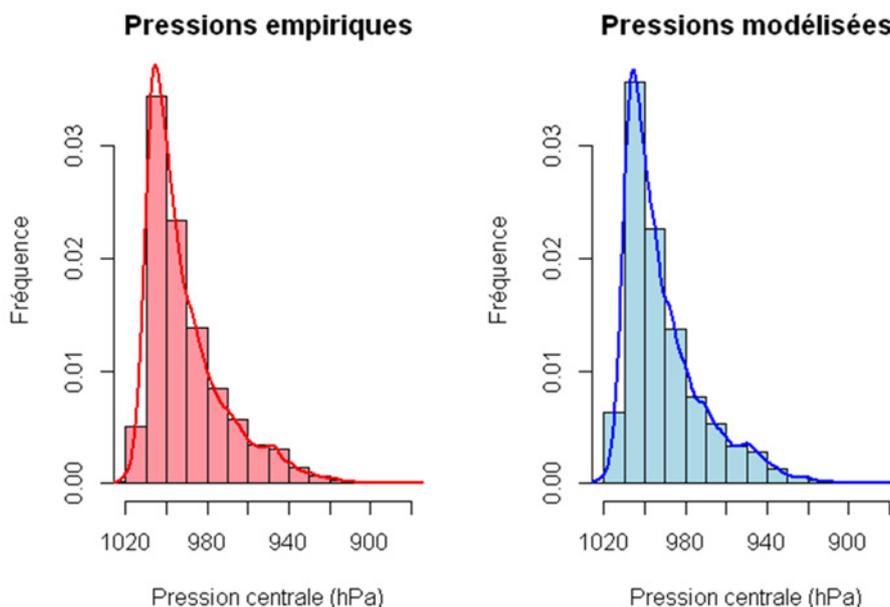


FIGURE 36 – Histogramme des pressions centrales

Les statistiques descriptives associées sont les suivantes :

	Min	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Max
Pression empirique	882	983	997	991	1005	1024
Pression modélisée	863,8	982,9	997,2	991	1005	1021

Nous pouvons aussi tracer leur graphe quantile-quantile (Q-Q plot) :

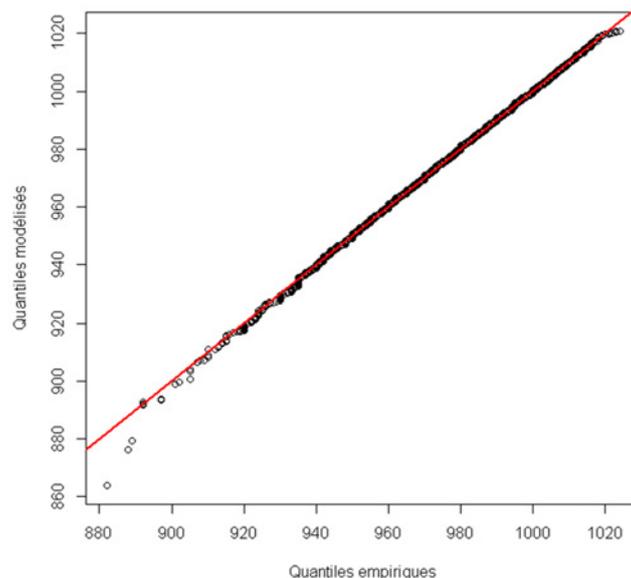


FIGURE 37 – Q-Q plot entre pressions empiriques et pressions modélisées

Là encore nous observons que la modélisation reste en accord avec les observations sans sortir de valeurs aberrantes. Notons tout de même que la pression centrale minimale obtenue par la régression est de 863,8 hPa contre 882 hPa empiriquement ce qui laisse penser que le modèle a tendance à faire chuter la pression un peu plus rapidement que dans la réalité ce qui est prudent.

5.3 Modèle de pression dans les terres

Le modèle utilisé pour déterminer la pression centrale lors d'une traversée d'étendue terrestre (aussi appelé atterrissage) est basé sur une décroissance exponentielle du déficit de pression généré par la tempête tropicale.

5.3.1 Modèle de désintégration exponentielle avec le temps

Lorsqu'une tempête tropicale traverse les terres les conditions de sa formation ne sont plus réunies et en particulier la nécessité de se trouver dans des eaux chaudes. Ainsi, la dissipation s'accélère et le phénomène peut disparaître dans les 24 heures. Cette perte d'énergie se caractérise par une hausse importante de la pression centrale qui se rapproche alors de la pression atmosphérique : le déficit de pression chute jusqu'à atteindre une valeur nulle. Dans la figure suivante nous donnons l'évolution du déficit de pression pour le cyclone Jeanne au cours du temps après son entrée dans les terres en prenant une pression atmosphérique de 1013 hPa.

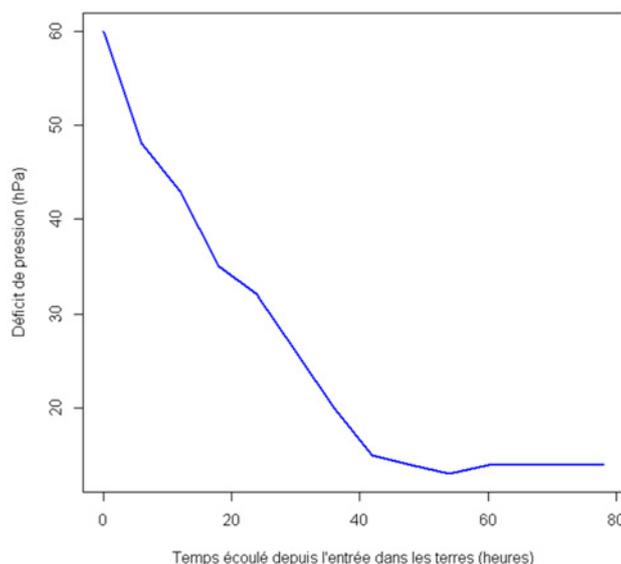


FIGURE 38 – Décroissance du déficit de pression au centre de Jeanne (2004)

Plusieurs modèles de désintégrations ont été proposés pour modéliser cette chute. Dans le cadre de ce mémoire nous retiendrons le modèle le plus simple proposé par KAPLAN & DEMARIA [12] basé sur une décroissance exponentielle du déficit de pression au cours du temps. Formellement, en prenant $t = 0$ l'instant d'entrée dans les terres de la tempête, le modèle s'écrit à toute date t :

$$P_{atm}(t) - P(t) = [P_{atm}(0) - P(0)] \times \exp(-at)$$

soit :

$$\Delta P(t) = \Delta P(0) \times \exp(-at)$$

où :

- $P_{atm}(t)$ est la pression atmosphérique en t que nous considérerons comme constante égale à 1013 hPa
- $P(t)$ est la pression au centre de la tempête en t
- $\Delta P(t) = 1013 - P(t)$ est le déficit de pression et t
- a est la constante de désintégration

La constante de désintégration a caractérise la vitesse de la décroissance et elle est différente selon le site où est observée l'entrée dans les terres. Pour simplifier, nous considérerons que a est une variable aléatoire réelle et positive mais constante le long de la portion de trajectoire pénétrant dans les terres. Si une tempête retourne dans l'océan puis traverse à nouveau une zone terrestre une nouvelle constante de désintégration est simulée *via* la distribution de probabilité.

Ce modèle peut être transposé à notre étude en discrétisant l'espace des temps avec un pas de 6 heures ($t = 0, 6, 12, \dots$). Ensuite, en répertoriant tous les morceaux de trajectoires historiques traversant des étendues terrestres dans la base HURDAT entre 1982 et 2011 on peut déterminer pour chacun d'entre eux la constante de désintégration associée par une régression exponentielle

entre le déficit et le temps écoulé depuis l'atterrissage sur les terres.

Dans la pratique, les morceaux de trajectoires possédant moins de 3 points ont été exclus afin d'effectuer les régressions sur au moins 4 points. Nous avons aussi exclu les portions ayant des pressions centrales supérieures à 1013 hPa pour respecter le domaine de définition de la fonction exponentielle ainsi que les atterrissages atypiques. Enfin, seules les régressions expliquant au moins 75% de variance ($R_{adj}^2 \geq 75\%$) ont été retenues. Au final, un échantillon de 64 constantes de désintégration a pu être obtenu sur un total de 90 atterrissages présentant plus de 4 observations. La figure ci-dessous présente le nombre d'observations ayant permis d'effectuer ces 64 régressions.

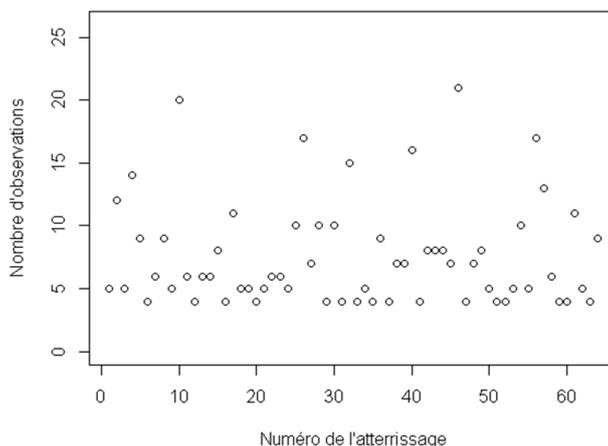


FIGURE 39 – Nombre d'observations à disposition

Les représentations graphiques des 64 régressions ainsi que leur coefficient de détermination ajusté associé sont présentées en annexe C.

5.3.2 Ajustement d'une distribution sur la constante de désintégration

L'histogramme de l'échantillon obtenue ainsi que sa boîte à moustache sont représentés ci-dessous :

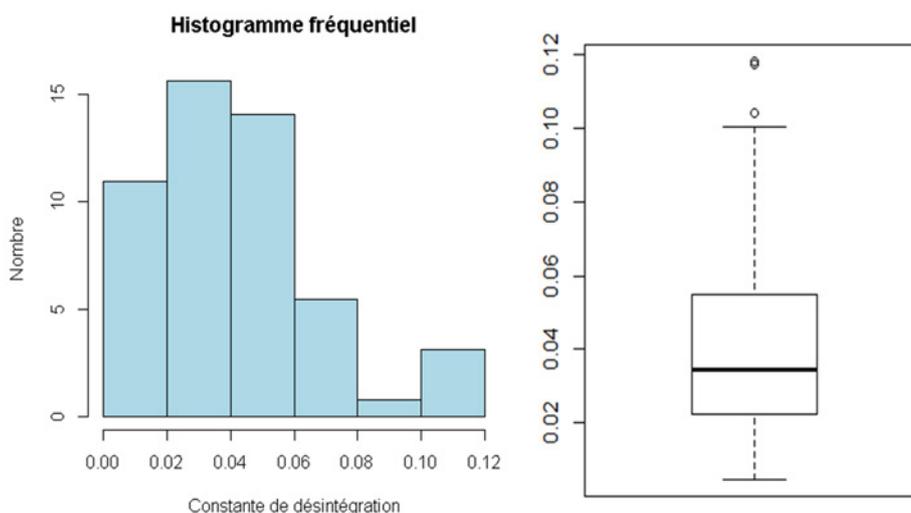


FIGURE 40 – Histogramme et boîte à moustache des constantes de désintégration

Nous nous proposons de tester l'adéquation de l'échantillon à une loi gamma et à une loi log-normale. Pour fixer les idées, précisons que WANG & ROSOWSKY indique dans leur étude que pour la région du Texas la constante de désintégration suit une loi log-normale de moyenne 0,04 et d'écart-type 0,032. Ce résultat est à rapprocher aux données de notre échantillon dont la moyenne est de 0,042 et l'écart-type de 0,026.

Pour l'étude, nous retiendrons la loi donnant le meilleur maximum de vraisemblance puis nous effectuerons un test d'adéquation pour valider la distribution. Auparavant, nous rappelons rapidement la définition et les estimateurs du maximum de vraisemblance de ces deux lois.

Soit (a_1, \dots, a_n) l'échantillon de constantes de désintégration obtenu avec $n = 64$ la taille de l'échantillon.

- Si a suit une loi log-normale de paramètres μ et σ^2 (notée $\mathcal{LN}(\mu, \sigma^2)$) sa densité s'écrit :

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

Les estimateurs du maximum de vraisemblance de μ et σ^2 sont donnés par :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln(a_i) \quad \text{et} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln(a_i) - \hat{\mu})^2$$

- Si a suit une loi gamma de paramètres α et β (notée $\mathcal{G}(\alpha, \beta)$) sa densité s'écrit :

$$\forall x \in \mathbb{R}^+, \quad f_a(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x)$$

où $\alpha \mapsto \Gamma(\alpha)$ est la fonction gamma d'Euler. Il n'existe pas de formules fermées pour les estimateurs du maximum de vraisemblance de α et β mais ces derniers vérifient :

$$\ln(\hat{\alpha}) - \frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} - \ln\left(\frac{1}{n} \sum_{i=1}^n a_i\right) + \frac{1}{n} \sum_{i=1}^n \ln(a_i) = 0 \quad \text{et} \quad \hat{\beta} = \frac{n\hat{\alpha}}{\sum_{i=1}^n a_i}$$

Ces équations peuvent être résolues en utilisant une méthode de Newton avec pour initialisation les estimateurs des moments suivants :

$$\hat{\alpha} = \frac{\hat{\beta}}{n} \sum_{i=1}^n a_i \quad \text{et} \quad \hat{\beta} = \frac{\frac{1}{n} \sum_{i=1}^n a_i}{\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i\right)^2}$$

Nous donnons dans le tableau suivant la log-vraisemblance obtenue *via* le logiciel R pour ces deux distributions ainsi que les estimateurs des paramètres.

Distribution	Log-vraisemblance	Paramètres estimés
$\mathcal{LN}(\mu, \sigma^2)$	151,96	$\hat{\mu} = -3,37$ et $\hat{\sigma}^2 = 0,429$
$\mathcal{G}(\alpha, \beta)$	152,98	$\hat{\alpha} = 2,71$ et $\hat{\beta} = 64,97$

La distribution gamma ayant la meilleure log-vraisemblance c'est celle que nous allons tester. L'hypothèse nulle du test est donc la suivante :

$$H_0 : a \sim \mathcal{G}(\hat{\alpha}, \hat{\beta})$$

Nous pouvons utiliser le test de Kolmogorov-Smirnov présenté précédemment afin de tester H_0 . Le logiciel R donne alors les résultats suivants pour un test de niveau 5% :

Statistique	Seuil critique	p-valeur
0,0634	0,1669	0,9446

La p-valeur très largement supérieure à 5% indique que nous ne devons pas rejeter H_0 : la constante de désintégration suit une loi gamma de paramètres $\alpha = 2.71$ et $\beta = 64.97$.

Le graphe quantile-quantile entre cette distribution et les données de l'échantillon est donné dans la figure suivante.

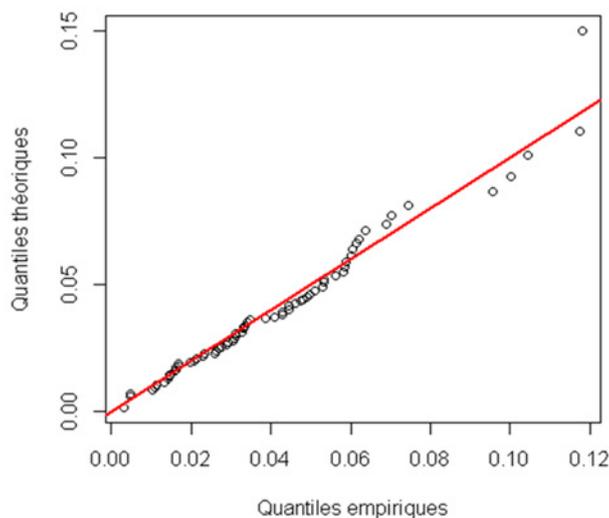


FIGURE 41 – Q-Q plot de la constante de désintégration

5.3.3 Confrontation avec les données empiriques

La figure suivante confronte les pressions centrales moyennes par localisation en résolution $1^\circ \times 1^\circ$ obtenue *via* la base HURDAT avec celles obtenues *via* l'association du modèle de VICKERY (avec le terme d'erreur gaussien) et du modèle de désintégration.

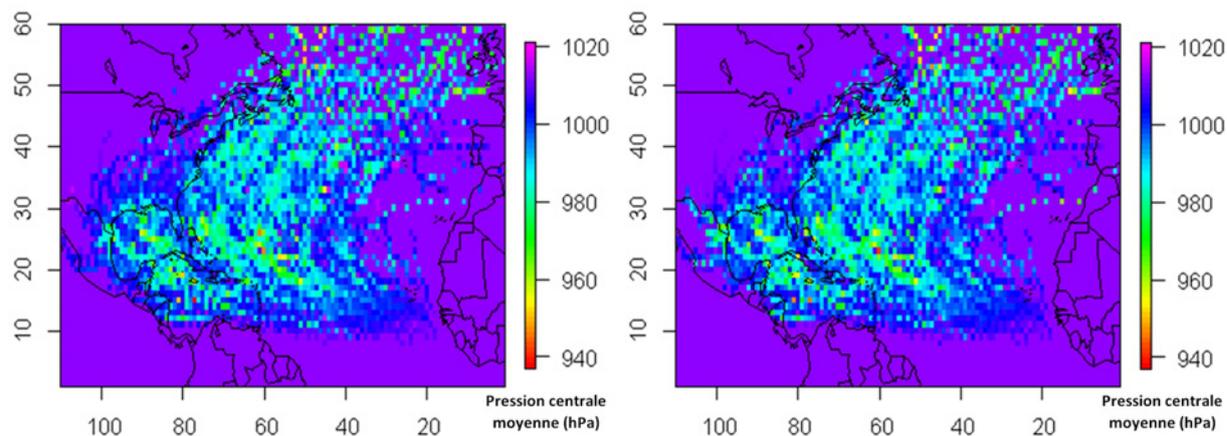


FIGURE 42 – Pressions centrales moyennes empiriques (à gauche) et modélisées (à droite)

Les pressions centrales issues de l'atterrissage des cyclones sont proches des données historiques même si elles ont tendance à être sous estimées par rapport aux données historiques.

Notons que ceci est prudent puisque des pressions centrales plus faibles correspondent à des vitesses de vents plus importantes comme nous allons le voir dans la section suivante.

5.4 Détermination de v_{max} , R_{max} et d'un gradient de vitesse des vents

La pression le long des trajectoires étant modélisée, on peut déterminer les autres caractéristiques clés d'une tempête tropicale c'est-à-dire la vitesse maximale des vents ainsi que le rayon de vitesse maximale des vents pour finalement en déduire un gradient de vitesse.

5.4.1 Vitesse maximale des vents

Dans de nombreux articles la vitesse maximale des vents (v_{max}) en chaque point de la trajectoire est directement déduite de la pression centrale (P). Cela se vérifie graphiquement à partir des données de la base HURDAT en traçant l'ensemble des vitesses maximales en fonction des pressions centrales pour la période 1982 – 2011 ce qui représente 11536 points.

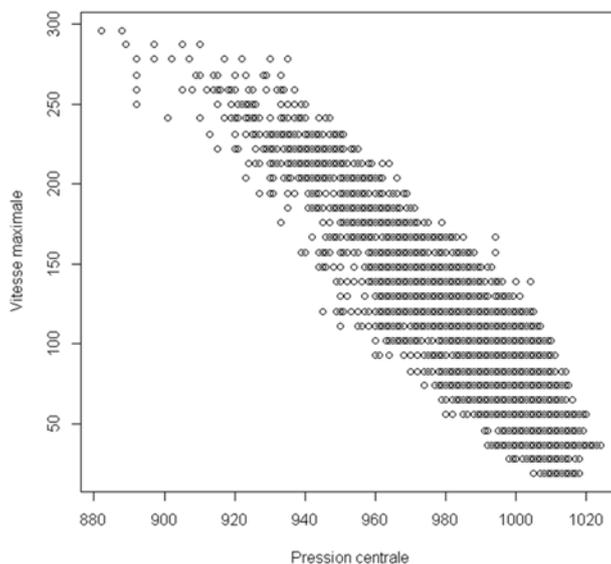


FIGURE 43 – Vitesse maximale des vents en fonction de la pression centrale (HURDAT)

La table de correspondance entre pression centrale et vitesse maximale la plus connue est celle DVORAK [4] dont le meilleur ajustement est donné par HOLLAND [10] dans sa révision des modèles vitesse-pression. Formellement nous avons :

$$v_{max} = 3,92 \times (1015 - P)^{0,644}$$

Le principale problème de cette approche est que l'on ne peut pas déterminer la vitesse maximale pour les hautes pressions (> 1015 hPa). Ainsi, dans le cadre de ce mémoire le modèle utilisé est un modèle de régression prenant comme variable explicative le déficit de pression défini par $\Delta P = 1013 - P$ et le carré de ce déficit. Ce choix est justifié par l'allure parabolique de la fonction reliant v_{max} et P dans la figure 43. Mathématiquement cela s'écrit :

$$v_{max} = a_0 + a_1 \Delta P + a_2 \Delta P^2 + \varepsilon$$

où :

- a_0 , a_1 et a_2 sont les constantes de la régression
- ε est un terme d'erreur avec $\mathbb{E}[\varepsilon] = 0$ et $\text{Var}[\varepsilon] = \sigma_\varepsilon^2$

L'estimation des paramètres par maximum de vraisemblance réalisé à partir du logiciel R mène à :

$$c_0 = 43,18 \quad c_1 = 2,79 \quad c_2 = -0,0066 \quad \sigma_\varepsilon^2 = 285,8$$

pour un coefficient de détermination ajusté de 0,87. Le tracé du graphe quantile-quantile (Q-Q plot) avec pour quantiles théoriques d'une loi normale centrée réduite est donné ci-dessous.

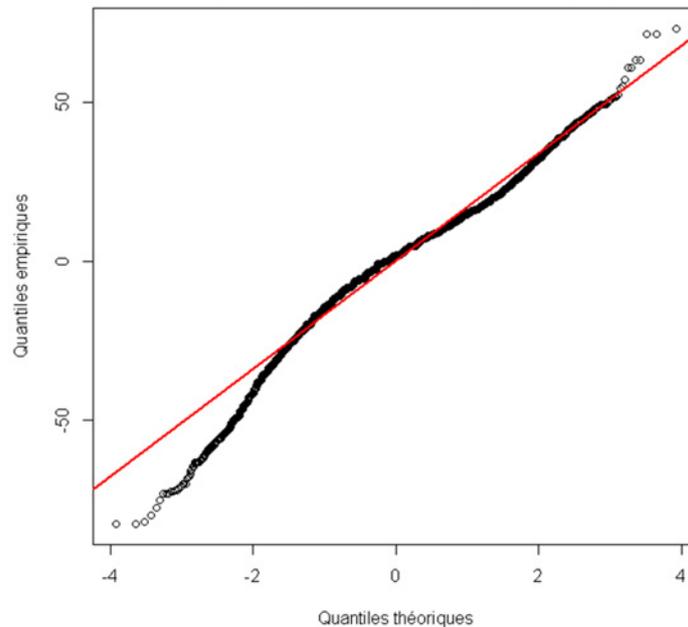


FIGURE 44 – Q-Q plot des erreurs

Au vu du graphe les extrêmes faibles sont sous estimée si on les considère gaussiens alors que le reste des résidus est bien modélisé. Cependant, nous retiendrons quand même une distribution gaussienne de moyenne 0 et de variance 285,8 même si cette hypothèse a tendance à surestimer les vitesses maximales.

Dans la figure suivante nous avons tracé la droite de régression issue du modèle de DVORAK (en noire) et celle issue de notre modèle (en bleue) le tout superposé aux données empiriques (points blancs). Nous avons aussi représenté les points issus de notre modèle avec simulation du terme d'erreur selon une loi normale (points bleus clairs).

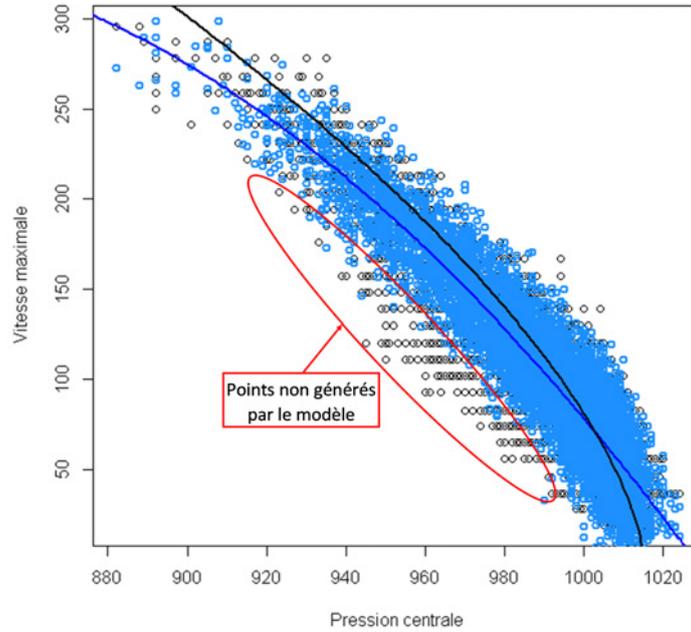


FIGURE 45 – Représentation des modèles liant v_{max} et P

Nous remarquons que le modèle de DVORAK sous-estime les vitesses maximales pour les hautes pressions (> 1005 hPa) et qu’au contraire il les sur-estime pour les faibles pressions (< 930 hPa). Ceci est un problème majeur car nous risquons d’avoir des vitesses de vents jamais observées ($>> 300$ km/h) aux très faibles pressions. Si on ajoute à cela un terme d’erreur aléatoire on va très vite obtenir des vitesses aberrantes.

A l’inverse, le modèle développé respecte bien l’aspect parabolique et les points résultant de l’ajout du terme d’erreur se superposent très bien aux données empiriques pour la partie située au dessus de la courbe. En revanche, l’observation faite *via* le Q-Q plot est confirmée puisque, pour la partie inférieure, certains points ne peuvent pas être modélisés à cause du choix de la loi normale qui modélise mal les extrêmes faibles. Si nous nous plaçons selon une approche prudente cela ne pose pas de problème majeur puisque ces points correspondent à des vitesses maximales anormalement faibles.

Enfin, avec la prise en compte du terme d’erreur il peut arriver que le modèle donne des vitesses négatives pour les très hautes pressions. Par convention nous prendrons donc le maximum entre 10 km/h est la valeur obtenue *via* le modèle. Pour la suite de l’étude, nous retiendrons le modèle quadratique plutôt que celui de DVORAK.

5.4.2 Rayon de vitesse maximale des vents

La notion de rayon de vitesse maximale n’a de sens que lorsqu’une tempête a atteint le stade de cyclone. Ainsi, seules les portions de trajectoire ayant atteint la catégorie 1 sur l’échelle de SAFFIR-SIMPSON (soit une vitesse supérieure à 118 km/h) sont concernées.

Dans la littérature, le rayon de vitesse maximale des vents (R_{max}) est une grandeur dont le logarithme est souvent exprimé en fonction de la latitude, de la longitude et du déficit de pression. Par exemple, VICKERY & AL. proposent la modélisation suivante dans [15] :

$$\ln(R_{max}) = 2,636 - 0,00005086\Delta P^2 + 0,0394899\Psi + \varepsilon$$

où :

- $\Delta P = 1013 - P$ est le déficit de pression centrale (en hPa)
- Ψ est la latitude du centre de la tempête (en degrés nord)
- ε est un terme d'erreur de loi normale centrée et d'écart-type 0,4 (en km)

Le coefficient de détermination associé est d'environ 28% ce qui est plutôt faible. Ainsi, nous nous proposons d'opter pour une autre modélisation grâce à la base HURDAT étendue. A partir de cette dernière on peut tracer le logarithme de R_{max} en fonction de diverses variables telles que la latitude, la longitude, le déficit de pression et son carré afin de déceler une éventuelle tendance. Les graphiques sont présentés dans la figure suivante.

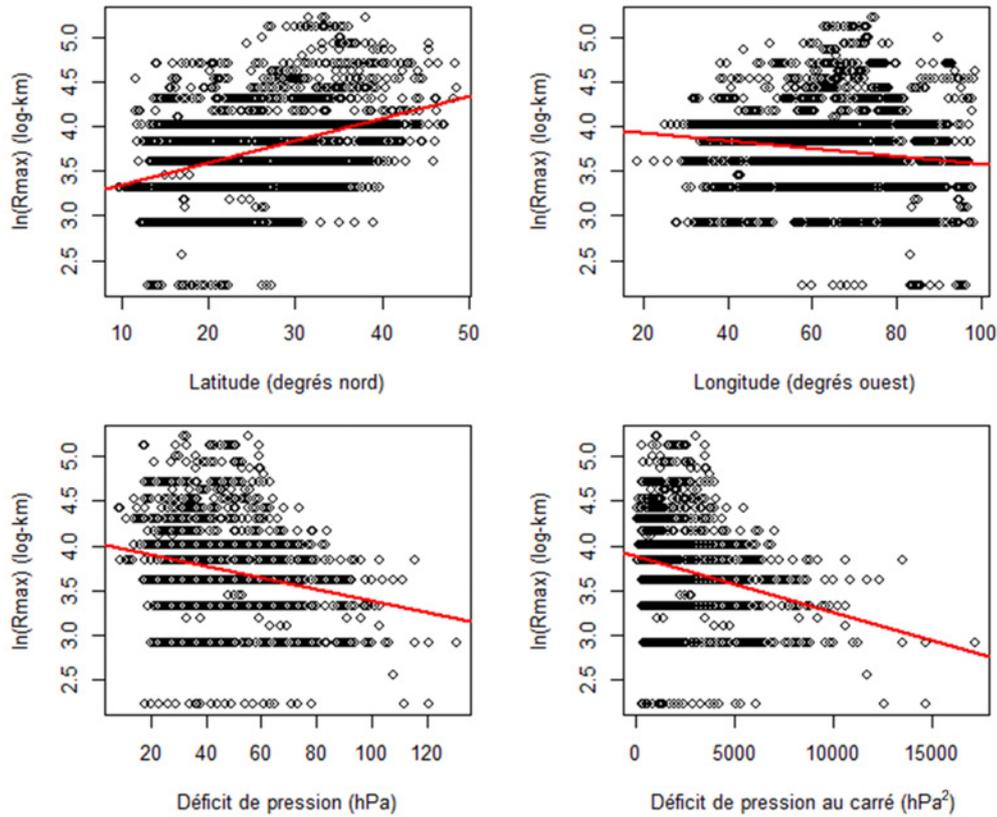


FIGURE 46 – Représentation du logarithme de R_{max} en fonction de diverses variables

Plusieurs tendances se dégagent de ces représentations. En effet, il apparaît que le rayon est d'autant plus important que le cyclone gagne en latitude alors qu'il a tendance à diminuer avec la longitude. Aussi, plus le phénomène est puissant (déficit de pression important) et plus R_{max} est faible ce qui indique que la taille de l'œil est en général plus faible pour les cyclones les plus violents. A ces tendances générales, il faut ajouter le fait que R_{max} est très stable le long d'une trajectoire. Cela apparaît clairement en traçant R_{max}^i pour un point i d'une trajectoire donnée en fonction du R_{max}^{i-1} associé au point précédent $i - 1$.

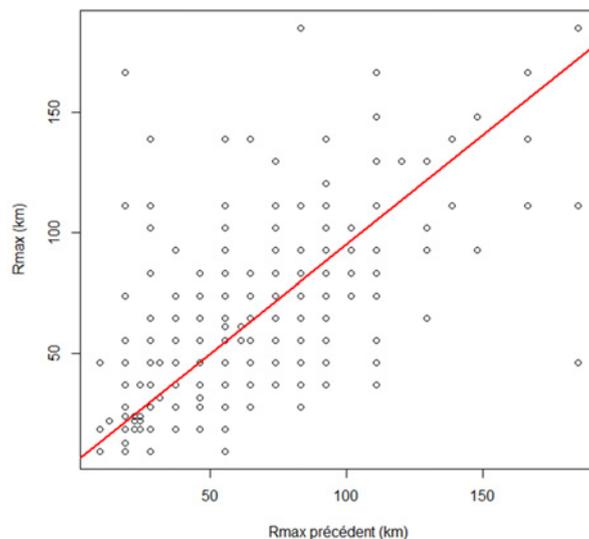


FIGURE 47 – Représentation de R_{max}^i en fonction du R_{max}^{i-1}

Dans le tableau suivant nous donnons les coefficients de détermination calculés pour chacune des tendances ci-dessus.

Variable	Latitude	Longitude	ΔP	ΔP^2	R_{max}^{i-1}
Coefficient R^2	15,1%	1,8%	5,7%	6,3%	80,4%

Au vu des résultats précédents et afin d'expliquer le maximum de variance nous proposons de modéliser R_{max} en chaque point i d'une trajectoire par le modèle de régression multiple suivant :

$$\ln(R_{max}^i) = d_0 + d_1\Psi + d_2\Lambda + d_3\Delta P + d_4\Delta P^2 + d_5 \ln(R_{max}^{i-1}) + \varepsilon$$

où :

- Ψ est la latitude (en degrés nord)
- Λ est la longitude (en degrés ouest)
- $\Delta P = 1013 - P$ est le déficit de pression (en hPa)
- R_{max}^{i-1} est le rayon de vitesse maximale en $i - 1$ (en km)
- ε est un terme d'erreur

Par convention, si la vitesse maximale des vents en i est inférieure ou égale à 118 km/h nous fixerons R_{max}^i à 0. L'estimation des coefficients a été réalisé grâce au logiciel R avec en *input* la base HURDAT étendue. Nous obtenons le modèle suivant :

$$\ln(R_{max}^i) = 0,263 + 0,0051\Psi - 0,00032\Lambda + 0,0034\Delta P - 0,000031\Delta P^2 + 0,88 \ln(R_{max}^{i-1})$$

Le coefficient de détermination ajusté associé est d'environ 80,7%. Ce pourcentage n'est pas beaucoup plus élevé que celui obtenu lors de la régression entre R_{max}^{i-1} et R_{max}^i . Néanmoins, une modélisation aussi simpliste n'aurait pas de sens d'un point de vu physique. En ce qui concerne les erreurs, ces dernières n'étant pas distribués selon une loi normale nous ne les modéliserons pas.

L'utilisation du modèle ci-avant requiert la connaissance du rayon initial. Ainsi, à chaque fois qu'une trajectoire passera du stade de tempête tropicale au stade de cyclone nous ferons un

tirage aléatoire parmi les rayons de vitesse maximale initiaux empiriques dont l'histogramme est représenté ci-après. Ensuite, le modèle sera appliqué le long du tronçon de trajectoire tant que la tempête reste au stade de cyclone.

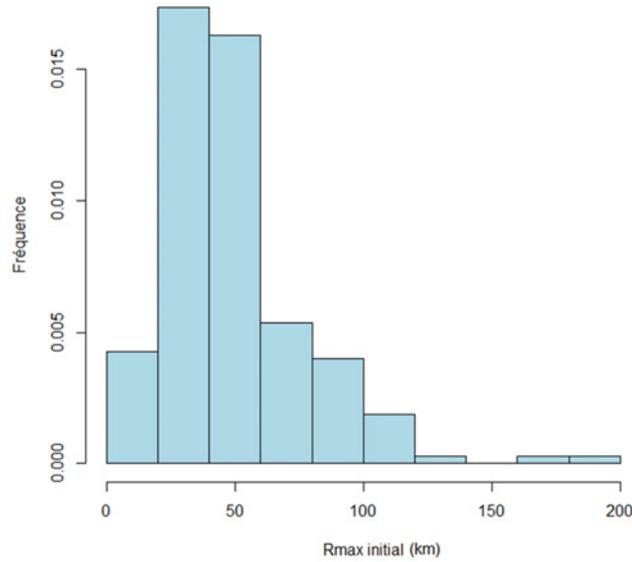


FIGURE 48 – Histogramme des R_{max} initiaux

5.4.3 Gradient de vitesse

Nous cherchons ici à modéliser la variation de la vitesse des vents en fonction de la distance à l'œil de la tempête par un modèle radial puisque nous considérons que les surfaces de vent forme des disques. De nombreux modèles paramétriques permettent de répondre à cette problématique et dans cette étude nous opterons pour celui proposé par JELESNIANSKI & AL. [11]. Ce modèle a été utilisé dans un programme américain de modélisation des cyclones appelé SLOSH. Il consiste à écrire la vitesse v en fonction de la distance r à l'œil selon :

$$v(r) = v_{max} \times \frac{2rR_{max}}{r^2 + R_{max}^2}$$

Ce gradient est une bonne approximation dès lors que le cyclone a une vitesse de déplacement proche de la stationnarité, hypothèse que l'on peut raisonnablement retenir. On remarque que la formule ne fait intervenir que deux grandeurs caractéristiques d'une tempête : R_{max} et v_{max} . Nous donnons ci-dessous l'allure du gradient de vitesse pour Katrina le 28 août 2005 à 18 heure UTC. Notons que cet instant correspond à l'instant de plus forte intensité de Katrina avec $R_{max} = 37$ km et $v_{max} = 280$ km/h.

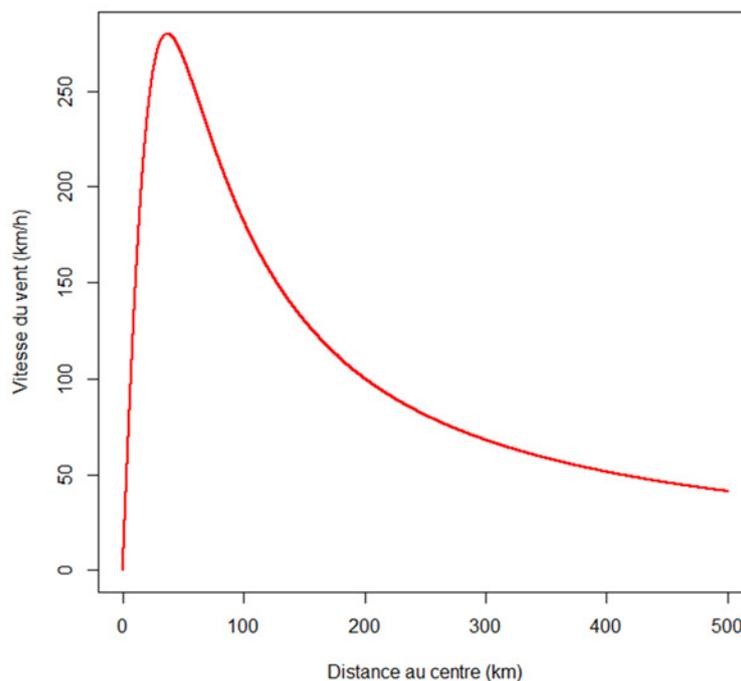


FIGURE 49 – Exemple d'un gradient de vitesse (Katrina, 2005)

L'avantage de ce modèle réside dans le fait qu'il respecte les caractéristiques d'un cyclone, à savoir une vitesse de vent quasiment nulle au centre, puis une croissance proportionnelle à r avec une vitesse maximale atteinte au mur de l'œil et enfin une décroissance en $1/r$.

Remarque : Comme nous l'avons remarqué dans la section précédente, le rayon R_{max} est fixé à 0 tant que la tempête n'a pas atteint le stade de cyclone et il n'est alors pas possible de construire un gradient de vitesse. Dans la suite cela ne sera pas préjudiciable car la vitesse maximale du vent est alors inférieure à 118 km/h, soit une vitesse d'atteinte des îles très faible ne présentant pas de danger particulier pour les bâtiments assurés.

5.5 Vitesse d'atteinte des îles

Nous avons associé les caractéristiques des tempêtes aux 10 000 trajectoires simulées à partir des différents modèles présentés précédemment. Ensuite, afin d'obtenir un échantillon de vitesse de vent atteignant chacune des quatre îles assurées nous avons procédé selon la méthodologie suivante :

- Chacune des îles assurées est supposée occuper le centre de la localisation en résolution $1^\circ \times 1^\circ$ qui lui correspond. Cette localisation est déterminée selon la même convention que celle explicitée dans la section 4.4.1. La figure suivante explicite cette convention géographique.

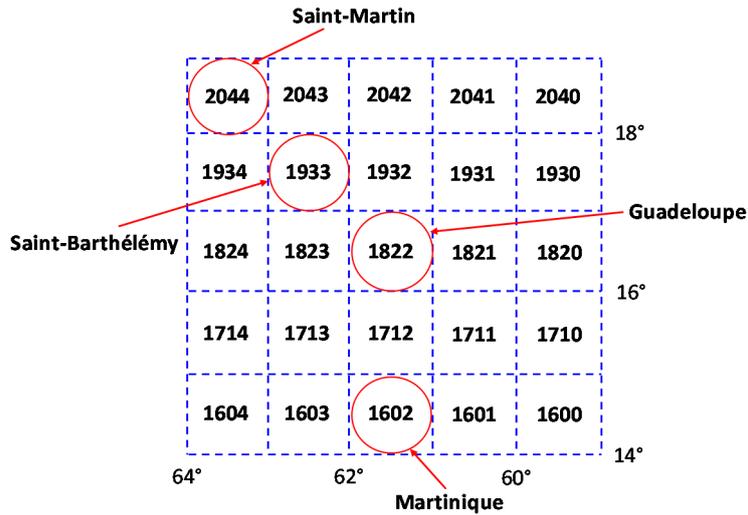


FIGURE 50 – Localisation des îles en résolution $1^\circ \times 1^\circ$

- Si un des points formant la trajectoire d'une tempête correspond à la localisation en résolution $1^\circ \times 1^\circ$ associée à une île on considère que cette dernière est touchée par la vitesse maximale des vents soutenue pendant une minute associée à ce point.
- Comme nous l'avons déjà fait remarqué précédemment, il peut arriver qu'une trajectoire traverse une île sans pour autant que l'un des points formant la trajectoire coïncide avec la localisation de cette île. On donne ci-dessous un exemple de cette situation.

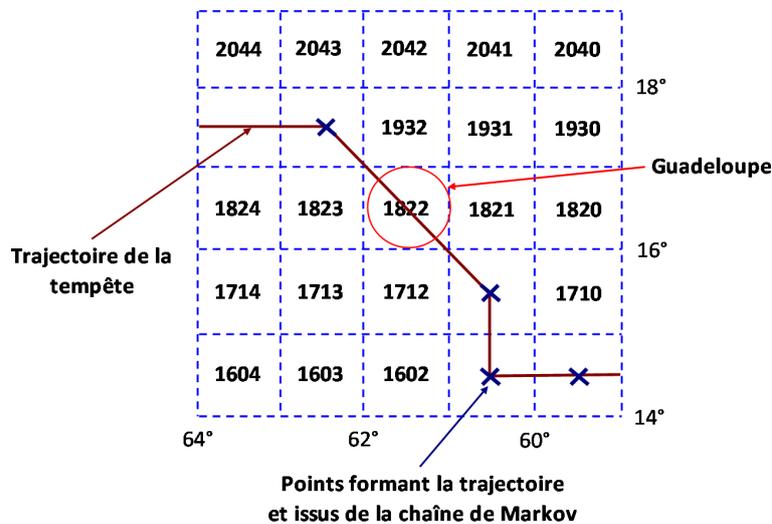


FIGURE 51 – Exemple d'une trajectoire traversant une île

Dans de tels cas, nous avons choisi de retenir la vitesse associée au point de la trajectoire correspondant à l'instant précédent directement la traversée de l'île.

- Lorsque la trajectoire d'une tempête ne traverse pas directement une île c'est le gradient de vitesse qui est utilisé. La distance entre les points de la trajectoire et l'île est calculée *via* la distance orthodromique. Une vitesse d'impact est calculée pour chacune des îles et pour l'ensemble des points de toutes les trajectoires. Remarquons que de nombreuses

vitesse vont tendre vers 0 km/h du fait de la distance très importante entre les îles et les trajectoires de certaines tempêtes.

- Pour une trajectoire donnée, c'est le maximum des vitesses d'impact qui est retenu. Ensuite, toutes les vitesses inférieures à 63 km/h sont fixées à zéro. Ce choix correspond à l'intensité d'une tempête tropicale sur l'échelle de SAFFIR-SIMPSON et il se justifie par l'absence de dommages aux bâtiments en dessous de cette catégorie. Les courbes de vulnérabilité présentées dans la suite valideront cette hypothèse.
- Enfin, les vitesses d'atteinte obtenues pour chaque île sont supposées affecter celle-ci totalement.

On donne dans la figure suivante les distributions des vitesses de vent obtenues pour chacune des quatre îles en suivant cette méthodologie.

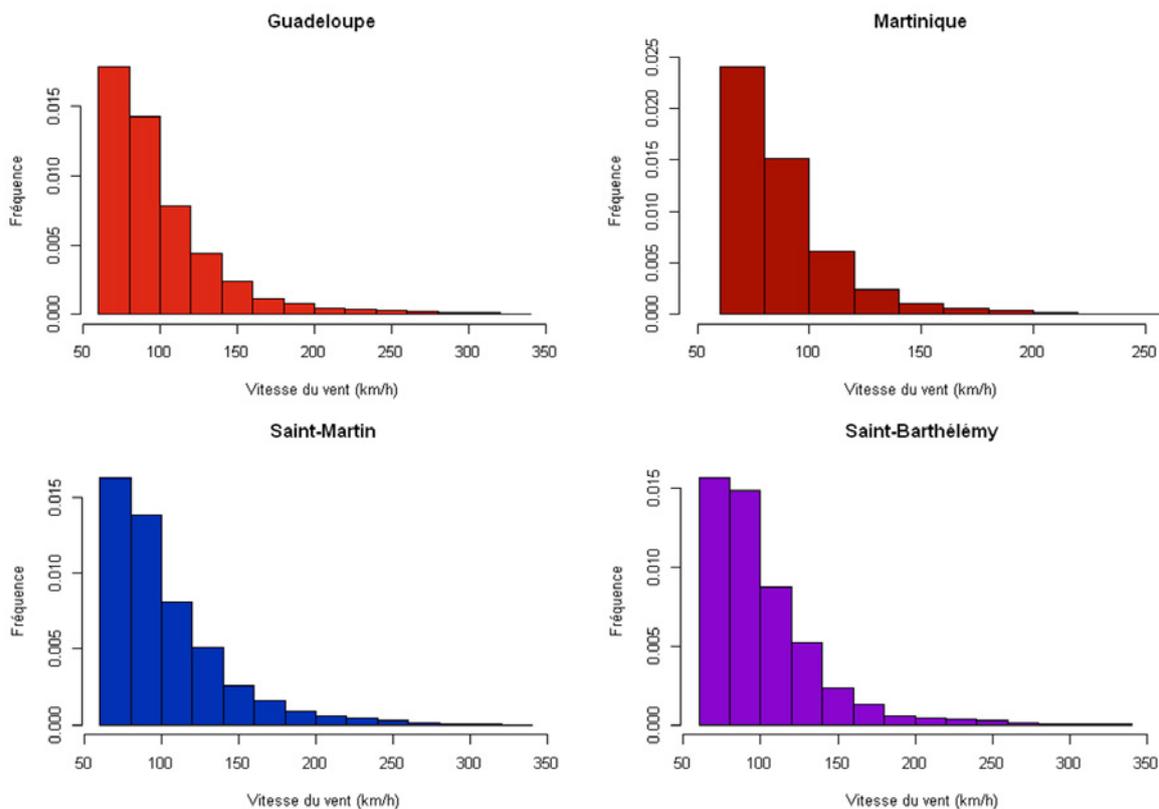


FIGURE 52 – Histogrammes des vitesses de vent simulées

On remarque des queues de distributions épaisses avec une vulnérabilité moindre pour la Martinique. Cela se vérifie en observant la base HURDAT puisque durant la période 1982 – 2011 la vitesse maximale ayant atteint la Martinique est de 165 km/h (ouragan Dean, 2007) alors que celle ayant atteint la Guadeloupe est de 225 km/h (ouragan Hugo, 1989). En remontant plus loin dans la base de donnée le constat reste le même : la Guadeloupe et les îles du nord (Saint-Martin et Saint-Barthélemy) sont touchées par des cyclones plus intenses que la Martinique.

Troisième partie

Simulation du coût pour l'assureur

Chapitre 6

Modélisation des coûts bruts de réassurance

L'objectif de ce chapitre est de faire le lien entre les intensités des tempêtes en terme de vitesse de vents et les dommages causés aux bâtiments assurés. Pour se faire, il existe deux principaux modèles :

- Les **modèles qualitatifs** qui décrivent les conséquences des vents extrêmes *a posteriori* par observation visuelle des dégâts, autrement dit par jugements d'experts. L'échelle de SAFFIR-SIMPSON en est un exemple puisqu'elle donne les dommages potentiels causés par les ouragans selon leur catégorie. Nous pouvons aussi citer l'échelle de TORRO ou encore l'échelle de FUJITA concernant les tornades.
- Les **modèles quantitatifs** qui associent la vitesse du vent à un taux de dommages en fonction du type de bâtiment. Dans la pratique, on s'intéresse aux dommages au sens économique du terme définis comme le coût de la réparation du bâtiment divisé par son coût de reconstruction total. En associant chaque taux de dommages à une vitesse de vent on crée alors ce que l'on appelle des courbes de vulnérabilité. Il existe deux façons de construire de telles courbes. La première consiste à effectuer une régression sur les données relatives aux dommages observés afin de construire des courbes de vulnérabilité empiriques. La deuxième consiste à étudier la résistance des structures pour en déduire des courbes de vulnérabilité théoriques. Dans le cadre de ce mémoire nous utiliserons ce deuxième type de courbe.

Dans la suite, nous présentons le portefeuille de bâtiments assurés par GFA Caraïbes puis nous explicitons le modèle de vulnérabilité choisi. Ensuite, nous utilisons les scénarios tempêtes simulés dans la partie précédente afin d'obtenir la distribution des coûts bruts de réassurance pour GFA Caraïbes. Enfin, nous calculons le SCR de tempête brut que nous comparons avec le résultat de la formule standard.

6.1 Lien entre vitesse du vent et taux de dommage

6.1.1 Présentation du portefeuille concerné

GFA Caraïbes assurent de nombreux type de bâtiments allant de la simple résidence pavillonnaire au complexe industriel en passant par les structures commerciales. En outre, l'assureur ne couvre pas seulement les structures mais aussi les contenus, ces derniers pouvant être exposés aux vents violents dès lors que des ouvertures apparaissent dans la toiture, les fenêtres ou le bâti. En terme de coût, la compagnie mesure son exposition maximale en sommant la totalité des sommes assurées en tempêtes/ouragans/cyclones. La répartition de ces sommes pour chacune des îles est donnée dans le tableau suivant :

En k€	Bâtiments résidentiels		Bâtiments commerciaux / industriels/ gouvernementaux		Total
	Structure	Contenu	Structure	Contenu	
Guadeloupe	357 500	35 800	671 000	74 600	1 138 900
Martinique	222 000	24 600	955 000	95 500	1 297 100
Saint-Martin	24 000	2 700	50 100	5 600	82 400
Saint-Barthélemy	10 800	1 200	22 400	2 500	36 900

FIGURE 53 – Sommes assurées par GFA Caraïbes

Au total, GFA Caraïbes est exposée à hauteur de plus de 2,5 milliards d'euros. Les bâtiments commerciaux, industriels et gouvernementaux ont été volontairement agrégés pour la suite. Notons que dans la catégorie "bâtiments résidentiels" on trouve aussi bien des pavillons individuels que des appartements de plusieurs étages.

6.1.2 Modèle de vulnérabilité

6.1.2.1 Présentation

L'étude de la vulnérabilité des bâtiments face aux vents est une étape délicate en raison notamment de l'incertitude liée aux taux dommages. Cette incertitude provient de nombreux facteurs tels que le relief de la zone sinistrée, les types de bâtiments endommagés ou encore l'espacement entre les bâtiments. Aussi, pour simplifier, nous considérerons que chaque île forme une seule zone sinistrée impactée par la même vitesse de vent sans distinction des quartiers. Ainsi, chaque bâtiment subira la même intensité de tempête.

Le modèle de vulnérabilité choisi est tiré d'une étude effectuée par EARLE pour le gouvernement des Iles Vierges Britanniques [8]. Ces dernières se situe dans l'arc antillais à l'instar des îles couvertes par GFA Caraïbes. Le modèle est de type quantitatif et utilise des courbes de vulnérabilité théoriques. Le rapport présente la construction de nombreuses courbes en fonction de deux grandes classes de bâtiments : les bâtiments résidentiels et les bâtiments industriels/commerciaux/gouvernementaux. Pour chacune de ces classes, sept courbes de vulnérabilité concernant la structure des bâtiments et sept courbes de vulnérabilités concernant le contenu des bâtiments sont proposées selon la qualité de la construction. Les classes sont explicitées ci-dessous par qualité de résistance décroissante.

	Bâtiments résidentiels	Bâtiments commerciaux / industriels/ gouvernementaux
Classe 1	Bâti en maçonnerie, fenêtres et portes protégées, toit à base de béton	Bâti en maçonnerie, fenêtres et portes protégées, toit à base de béton
Classe 2	Bâti en maçonnerie, fenêtres avec jalousies métalliques, toit à base de béton	Bâti en maçonnerie, fenêtres avec jalousies métalliques, toit à base de béton
Classe 3	Bâti en maçonnerie, fenêtres et portes en verre, toit à base de béton	Bâti en maçonnerie, fenêtres et portes en verre, toit à base de béton
Classe 4	Bâti en maçonnerie, fenêtres et portes protégées, autres matériaux de toiture	Bâti en maçonnerie, fenêtres et portes protégées, autres matériaux de toiture
Classe 5	Bâti en maçonnerie, fenêtres avec jalousies métalliques, autres matériaux de toiture	Bâti en maçonnerie, fenêtres avec jalousies métalliques, autres matériaux de toiture
Classe 6	Bâti en maçonnerie, fenêtres et portes à base de verre, autres matériaux de toiture	Bâti en maçonnerie, fenêtres et portes à base de verre, autres matériaux de toiture
Classe 7	Bâtiment à structure en bois	Bâtiment à structure métallique

FIGURE 54 – Classes proposées en fonction de la qualité des bâtiments

En toute rigueur, il faudrait répertorier tous les bâtiments et les classer par classes homogènes selon la qualité de leur structure afin de leur appliquer la bonne courbe de vulnérabilité. Dans la pratique, nous n'avons pas accès à ces informations il est de toute façon très difficile de réaliser une telle étude. Néanmoins, nous savons que les bâtiments en bois ne sont pas acceptés ainsi que les bâtiments à structure métallique. Ainsi, nous avons choisi de nous concentrer sur une classe intermédiaire, à savoir la troisième classe toujours par principe de prudence. Toutefois, afin d'avoir une idée de la sensibilité des coûts due à un changement de courbe, nous avons également testé les deux premières classes. Ces dernières ont la particularité de correspondre à des bâtiments possédant des protections contre les vents extrêmes : présence de jalousies métalliques (classe 2), c'est-à-dire de volets mobiles composés de lames métalliques orientables permettant une protection accrue des fenêtres notamment contre les projectiles (voir figure 55), et présence de protections aux portes et fenêtres (classe 1) correspondant bien souvent à des barricades en bois ou en métal. Ces protections sont souvent provisoires et consécutives à une alerte cyclonique de la part des autorités, il est donc difficile de quantifier le pourcentage de bâtiments concernés.



FIGURE 55 – Fenêtres munies de jalousies métalliques

6.1.2.2 Courbes de vulnérabilité

Les courbes de vulnérabilité associées à la classe 3 sont au nombre de 4 : deux courbes concernant les bâtiments résidentiels (structure et contenu) et deux courbes concernant les bâtiments

industriels, commerciaux et gouvernementaux (structure et contenu). Ces dernières ne sont pas fournies sous la forme d'une équation explicite donnant le taux de dommages en fonction de la vitesse des vents, mais sous la forme d'une série de 15 points. Ces points sont représentés dans la figure ci-après. Les vitesses de vent sont les vitesses moyennes soutenues pendant une minute.

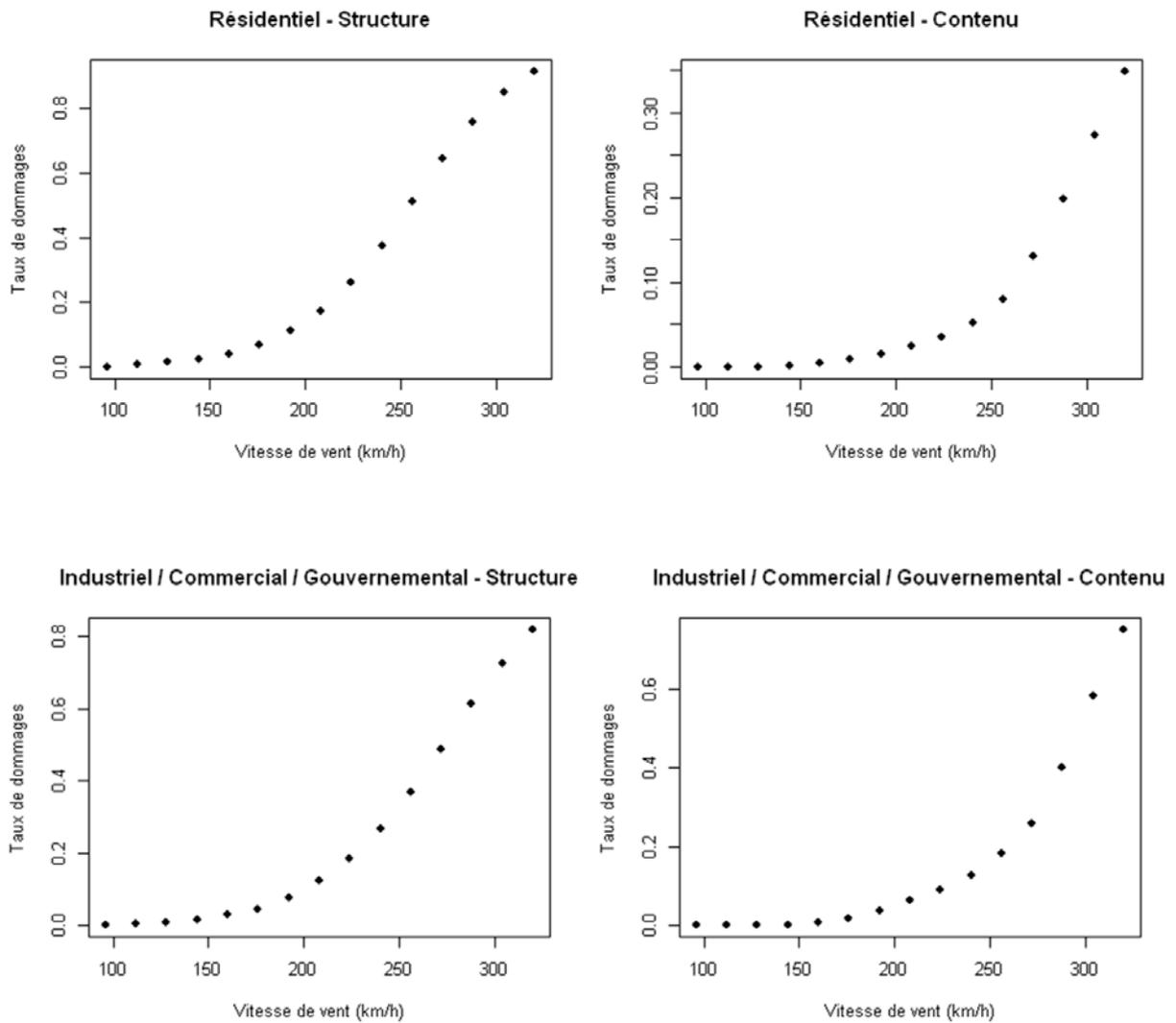


FIGURE 56 – Courbes de vulnérabilité retenues

Ces quatre courbes ont toute la même allure avec un taux de dommages nul jusqu'à un premier seuil de vitesse de vent. Puis, les dommages sont strictement croissant avec la vitesse jusqu'à un second seuil correspondant à une destruction totale du bâtiment.

En l'état ces courbes ne sont pas exploitables et il est nécessaire de leur ajuster une équation. Pour cela nous avons opté pour des interpolations polynomiales allant jusqu'au degrés 6. Les équations trouvées sont données dans l'annexe D. Les courbes ajustées sont représentées ci-dessous (en rouge) en superposition des points théoriques.

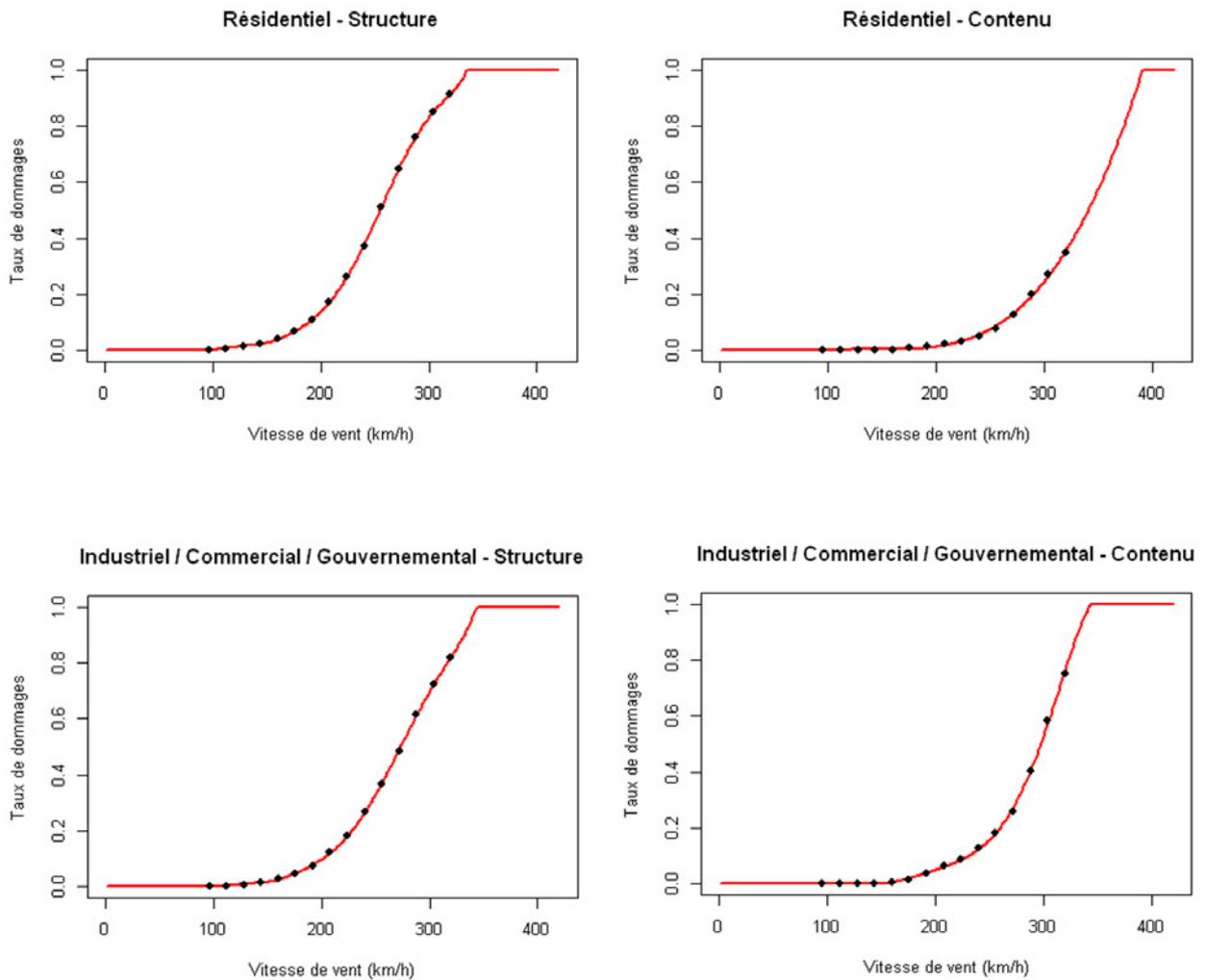
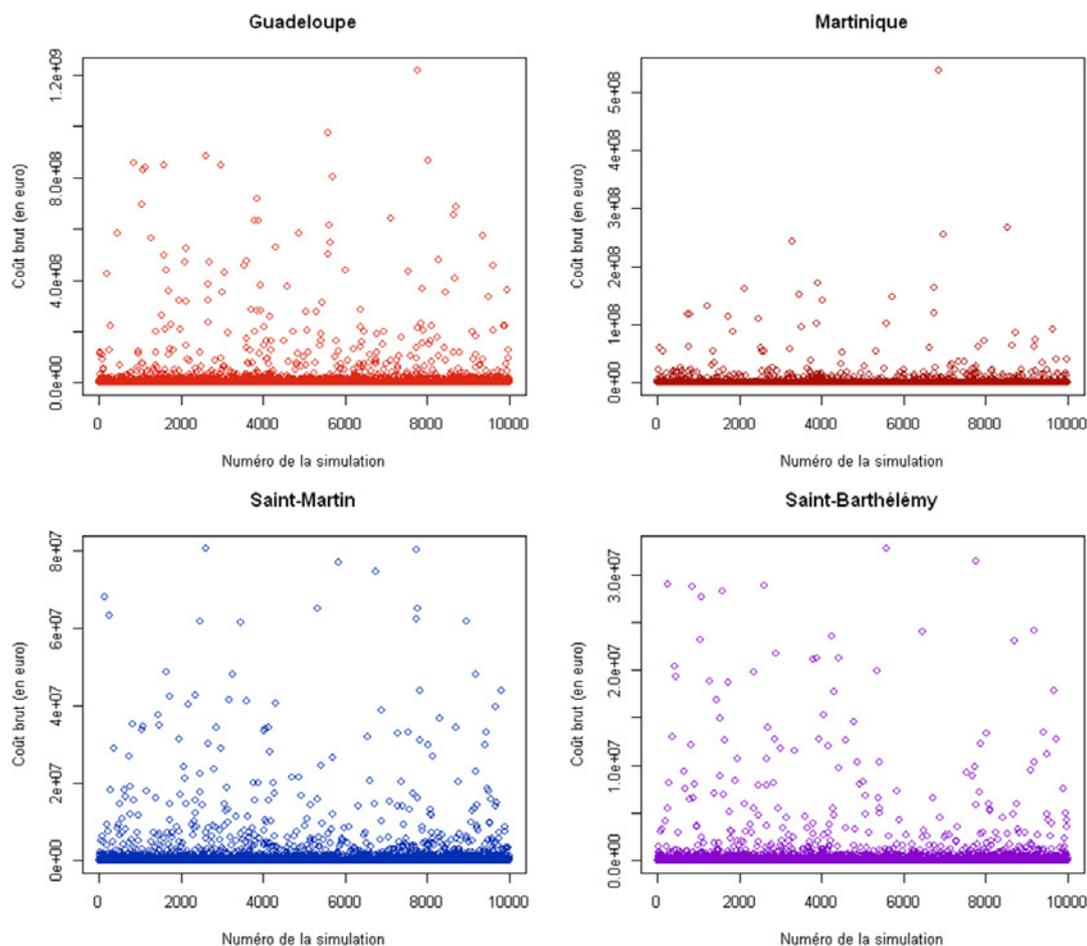


FIGURE 57 – Courbes de vulnérabilité ajustées

6.1.3 Distribution des coûts bruts de réassurance

Nous avons converti chaque vitesses générées dans la section 5.5 en terme de taux de dommages grâce aux courbes précédentes. Puis, en multipliant les taux obtenus par les sommes assurées correspondantes un coût brut de réassurance a été obtenue pour chaque simulation et pour chacune des îles assurées. Nous rappelons qu'une simulation correspond à une année cyclonique. Les coûts sont représentés dans la figure ci-après.



De nombreux coûts sont nuls en raison de l'absence de tempêtes pour certaines années ou de vents trop faibles pour créer des dommages matériels. Par ailleurs, nous remarquons des points extrêmes très importants correspondant à des années d'intensité cyclonique exceptionnelle. Ce sont essentiellement ces points qui vont contribuer au SCR de tempêtes.

6.2 Calcul du SCR de tempêtes brut de réassurance

6.2.1 Utilisation de l'outil de simulation de tempêtes

L'exigence de fonds propres liée au risque de tempêtes pour GFA Caraïbes correspond simplement au quantile à 99,5% de la distribution des coûts bruts totaux annuels. Remarquons simplement que ce SCR ne correspond pas à la somme des SCR de chacune des îles puisque la corrélation entre les sinistres n'est pas parfaite. Nous donnons ci-dessous les coûts totaux bruts générés ainsi que l'évolution du SCR de tempêtes brut en fonction du nombre de simulation réalisées.

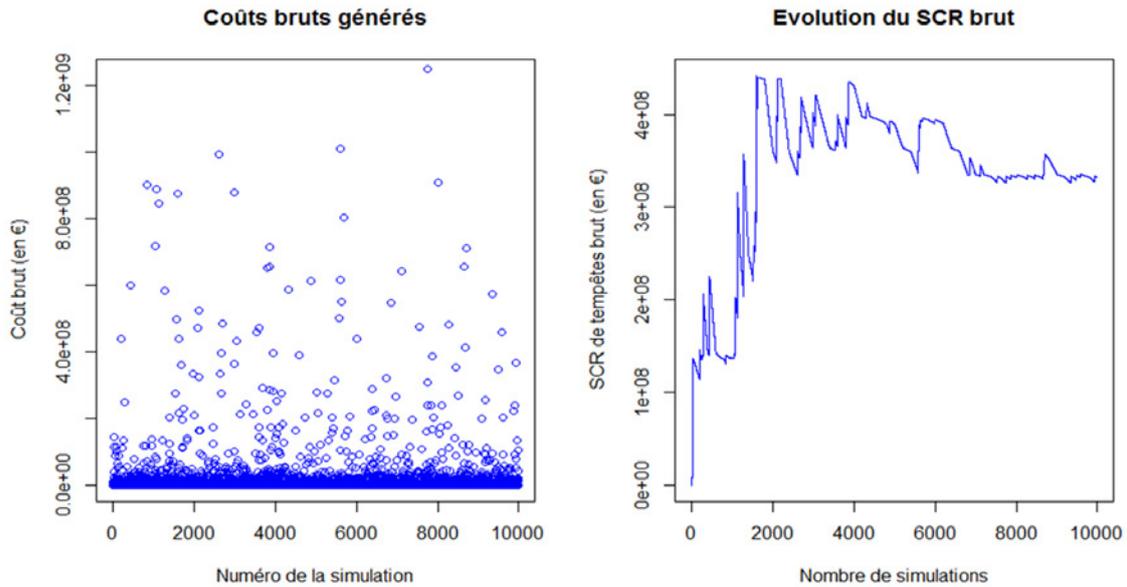


FIGURE 58 – Résultats bruts issus des simulations

Nous observons une convergence satisfaisante du SCR et il semble donc que le nombre de simulations effectuées soit suffisant. Ainsi, nous retiendrons le résultat final suivant qui correspond au quantile trouvé après 10 000 simulations :

SCR brut de réassurance (en k€)	
Classe 3	332 381

Ce résultat correspond à environ 24 années de chiffre d'affaires dommages aux biens ce qui est insupportable pour la plupart des compagnies, *a fortiori* pour une petite structure telle que GFA Caraïbes. Ainsi, la cession d'une grande partie du risque à la réassurance apparaît comme essentielle et fera l'objet du chapitre suivant.

A titre de comparaison, nous donnons dans le tableau ci-après les résultats obtenus en utilisant les courbes associées aux classes 1 et 2. A l'instar des courbes de classe 3 leurs équations sont données en annexe D.

	SCR brut de réassurance (en k€)	Comparaison classe 3
Classe 1	293 181	-11,8%
Classe 2	319 583	-3,9%

Nous observons un gain maximal en capital réglementaire de l'ordre de 39,2 millions d'euros soit 11,8 points de moins par rapport au résultat obtenu *via* la classe 3. En considérant les courbes de la classe numéro 2 l'écart est encore moins important ce qui conforte notre choix de départ (classe 3) guidé par le principe de prudence sur lequel s'appuie Solvabilité II.

6.2.2 Comparaison avec la formule standard

Actuellement, les calculs des différents SCR par la formule standard sont en cours de calibrage par l'intermédiaire de ce que l'on nomme communément les études quantitatives d'impact (QIS). Depuis le lancement de Solvabilité 2 cinq études ont été lancées, la dernière en date étant le QIS5. Les méthodes de calcul de chaque SCR sont explicitées dans les spécifications techniques fournies par l'EIOPA (*European Insurance and Occupational Pensions Authority*) [5]

et on trouve les éléments de calcul dans des fichiers Excel annexes ²⁰.

Concernant le SCR de tempêtes brut de GFA Caraïbes, le calcul s'effectue en deux étapes. La première étape consiste à appliquer un pourcentage déterministe appelé facteur de risque au montant des sommes assurées en assurance dommages incendie afin d'obtenir le risque de base pour chacune des îles. Ce pourcentage varie selon la localisation et correspond théoriquement à une perte provoquée par un événement de période de retour bicentenaire. Les facteurs à appliquer sont donnés dans le tableau suivant. Précisons que le facteur de risque n'est pas précisé pour Saint-Barthélemy, néanmoins la localisation géographique de cette dernière étant très proche de Saint-Martin nous retenons le même taux.

	Guadeloupe	Martinique	Saint-Martin	Saint-Barthélemy
Facteur de risque	2,74%	3,19%	5,16%	5,16%

Nous observons dans ce tableau un facteur de risque plus élevé pour la Martinique que pour la Guadeloupe ce qui est en contradiction avec les résultats obtenus précédemment.

Pour la suite des calculs, un facteur égal à 1,2 est appliqué au risque de base ce qui permet d'obtenir un SCR brut pour chaque île. Nous reviendrons sur ce facteur dans la partie sur le calcul du SCR net. La seconde étape consiste à agréger les SCR intermédiaires en supposant que les risques associés sont parfaitement dépendants (corrélation de 100%), formellement cela revient à sommer les SCR bruts de la façon suivante :

$$SCR_{GFAC} = SCR_{Guadeloupe} + SCR_{Martinique} + SCR_{St\ Martin} + SCR_{St\ Barthélemy}$$

En appliquant cette méthodologie nous obtenons les résultats présentés dans le tableau ci-dessous. Précisons que les sommes assurées sont plus importante que celles présentées dans la section 6.1 par prudence vis-à-vis de la définition assez flou des sommes assurées formulée dans les spécifications techniques. Ainsi, nous considérons ici la totalité des sommes assurées incendie alors qu'en modèle interne, où la vision du risque est plus fine, nous retenons les sommes assurées en tempêtes, ouragans et cyclones.

	Sommes assurées incendie (en k€)	SCR brut (en k€)
Guadeloupe	1 487 120	48 896
Martinique	1 663 800	63 690
Saint-Martin	110 980	6 872
Saint-Barthélemy	49 750	3 080
Total	3 311 650	122 539

Comparativement au résultat obtenu par l'outil de modélisation interne, ce capital réglementaire est environ 2,7 fois plus faible. Cependant ce résultat est à relativiser notamment en raison de la simplicité du modèle de vulnérabilité utilisé et de la méthodologie prudente retenue pour les vitesses de vent atteignant les îles (voir section 5.5 de la deuxième partie).

²⁰. Paramètres de calcul du risque de catastrophes non-vie : <https://eiopa.europa.eu/consultations/qis/quantitative-impact-study-5/technical-specifications/index.html>

Chapitre 7

Introduction de la réassurance

Le chapitre précédent montre que les coûts bruts de réassurance générés par les tempêtes tropicales peuvent s'avérer extrêmement élevés et en particulier lorsque l'on va chercher des quantiles extrêmes. Ainsi, GFA Caraïbes possèdent une couverture de réassurance très importante afin de faire face sinistres engendrés par les ouragans. Outre la réassurance classique, la particularité de ces territoires réside dans le fait qu'il existe un régime spécial permettant aux assureurs de se réassurer de manière illimitée auprès de l'Etat français par l'intermédiaire d'un organisme appelé la Caisse Centrale de Réassurance (CCR).

Ce chapitre commence par reprendre brièvement les principes des traités de réassurance classiques pour ensuite expliciter le régime de la CCR. Puis, nous présentons le programme de réassurance de GFA Caraïbes que nous appliquons aux coût bruts précédents afin d'obtenir le SCR de tempêtes net de réassurance. Nous appliquons ensuite une méthode de rééchantillonnage afin d'obtenir un intervalle de confiance pour ce capital réglementaire. Enfin, nous comparons le résultat avec celui de la formule standard.

7.1 Généralités sur la réassurance

7.1.1 Principe de la réassurance

D'après le code des assurances, la réassurance est l'activité d'un organisme qui consiste à accepter des risques d'assurance cédés par une entreprise d'assurance. On a alors coutume de dire que le réassureur est l'assureur de l'assureur, le premier étant qualifié de cessionnaire est le second de cédant.

Pour une compagnie d'assurance telle que GFA Caraïbes, les principaux objectifs d'un recours à la réassurance sont :

- Faire face aux pics de sinistralité exceptionnels de type cyclones tropicaux et ainsi réduire significativement la probabilité de ruine de la compagnie
- Lisser les ratios sinistres sur primes annuels et donc stabiliser le résultat
- Augmenter la solvabilité de l'entreprise

Il existe deux formes de réassurance : la réassurance facultative pour laquelle le transfert des risques s'effectue au cas par cas et le traité de réassurance qui couvre un ensemble de risques sur un portefeuille. Dans la suite nous nous limitons à la présentation des trois traités de base qui font partie du programme de GFA Caraïbes.

7.1.2 Les traités de réassurance classiques

Les traités de réassurance se divisent en deux catégories :

- **Les traités proportionnels** pour lesquels l'assureur cède au réassureur un pourcentage des primes encaissées en échange de la prise en charge de ce même pourcentage du montant de tout sinistre survenant. Le traité proportionnel le plus simple est le *quote-part* dans lequel le réassureur partage une proportion équivalente des primes et des sinistres du portefeuille de la cédante. Formellement, si on note X le montant du sinistre et d le pourcentage qui est pris en charge par le réassureur, ce dernier intervient à hauteur de :

$$R = d \times X$$

Une illustration de ce traité est proposée dans la figure suivante à travers l'exemple d'un quote-part avec $d = 30\%$.

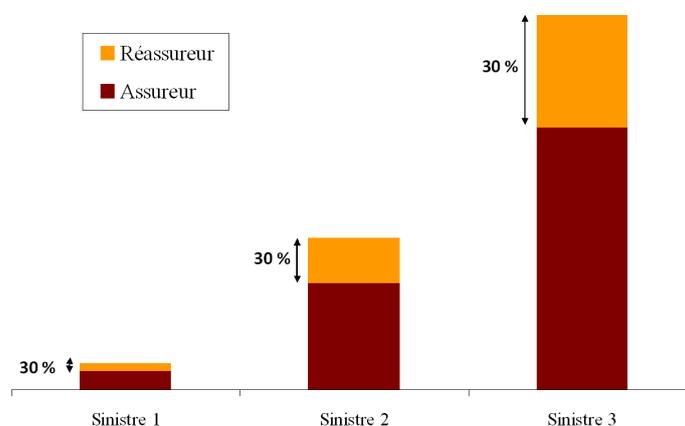


FIGURE 59 – Illustration du traité quote-part

Ce type de traité permet de diminuer l'engagement de l'assureur mais ne permet pas un écrêtement, il se contente d'abaisser les pics des sinistres. Ainsi, son principal inconvénient est la non couverture des sinistres extrêmes.

- **Les traités non-proportionnels** pour lesquels il n'y a pas de relation de proportionnalité entre la prime cédée et la participation du réassureur dans les sinistres. Dans ce type de traité, on définit une priorité et une portée. Le réassureur prend alors à sa charge la compensation financière relative au sinistre survenu jusqu'à un montant qui excède la priorité et dans la limite de la portée. Les deux traités non-proportionnels les plus classiques sont l'excédent de sinistres et l'excédent de pertes annuelles.

- *L'excédent de sinistre* peut couvrir les sinistres police par police ou bien les sinistres agrégés résultant d'un même évènement comme par exemple le passage d'un ouragan. Traditionnellement, on note ce type de traité $a XS b$ où a est la portée et b la priorité. Pour un sinistre de montant X la charge qui revient au réassureur se calcule alors comme suit :

$$R = \min [\max (X - b ; 0) ; a]$$

Il est courant de voir des montages avec plusieurs traités en excédent de sinistre mis bout à bout. Chaque réassureur intervient alors sur la tranche pour laquelle il s'est engagé. Nous donnons l'exemple d'un montage avec deux traités en excédent de sinistre ci-dessous.

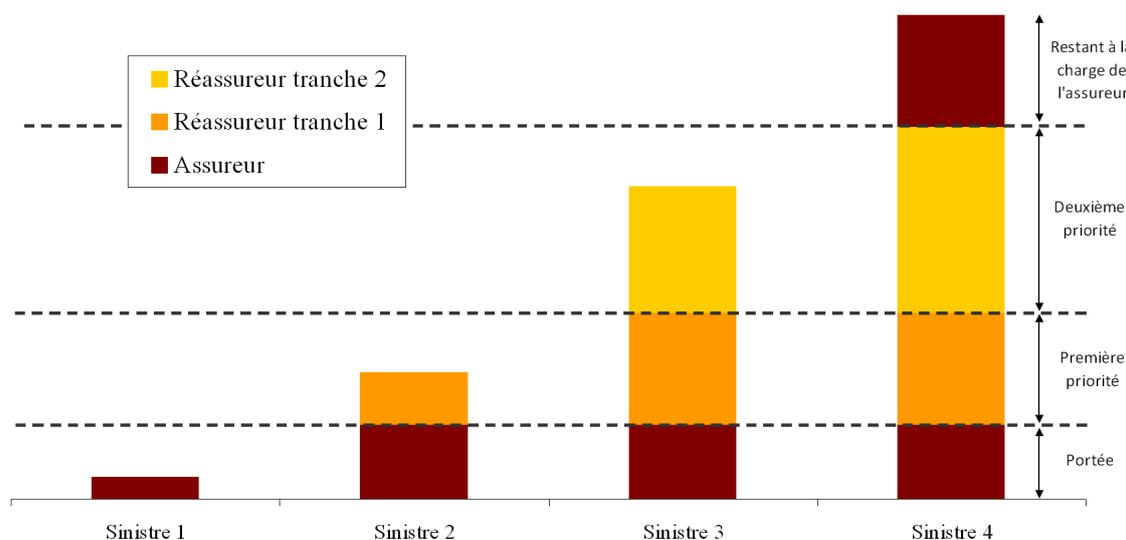


FIGURE 60 – Illustration du traité en excédent de sinistre

Ce traité de réassurance permet d'écrêter le profil des montants de sinistres. Cependant, il ne protège pas contre le risque de fréquence, c'est-à-dire contre un nombre inhabituel de petits sinistres n'excédant pas la priorité. C'est l'excédent de perte annuel qui protège contre ce risque. Par ailleurs, le traité prévoit en général un certain nombre de reconstitutions de la garantie durant la période de couverture. En contrepartie, la cédante verse une prime supplémentaire souvent exprimée en pourcentage de la prime initiale et au prorata de la portée à reconstituer. On parle alors de prime de reconstitution.

- L'*excédent de pertes annuelles* (ou *stop loss*) repose sur le même principe que l'*excédent de sinistres mis à part* qu'il protège le ratio sinistres sur primes annuel (noté S/P). Classiquement, on note cette couverture x SL y avec x la portée et y la priorité exprimées en pourcentage de la prime P perçue par l'assureur au titre du risque réassuré. Formellement, si on note S le montant des sinistres de la cédante durant la période de couverture du traité de réassurance, la charge du *stop loss* pour le cessionnaire sera :

$$R = P \times \min \left[\max \left(\frac{S}{P} - y; 0 \right); x \right]$$

Avec ce type de couverture l'assureur cède une partie de son résultat technique déficitaire au réassureur et se protège ainsi contre une sinistralité trop importante au cours de l'année.

7.1.3 Le cas particulier de la Caisse Centrale de Réassurance (CCR)

Dans le cadre du régime CAT-NAT (catastrophe naturelle) mis en place par l'Etat français, les assureurs ont la possibilité de se réassurer auprès d'un organisme dont l'Etat est actionnaire à 100% dénommé Caisse Centrale de Réassurance (CCR) moyennant une prime annuelle. Précisons que cette réassurance n'est pas obligatoire et que les assureurs peuvent lui préférer un cessionnaire privé. Cependant, très peu de compagnies optent pour cette option puisque selon une publication de l'APREF (Association des Professionnels de la Réassurance En France) le pourcentage de primes cédées à la CCR au titre de la garantie catastrophes naturelles s'élève à 95% (voir [1]).

L'intervention de la CCR en tant que réassureur est conditionnée par la reconnaissance de l'état de catastrophe naturelle par un arrêté interministériel. La procédure de reconnaissance de cet état est la suivante :

1. Les maires des communes concernées par l'évènement naturel émettent une demande auprès du préfet du département
2. Le préfet établit alors un dossier départemental visant à démontrer la nature et l'intensité anormale de l'évènement. Parmi les documents le constituant, on peut notamment citer :
 - des rapports techniques établis par des jugements d'experts
 - la liste des communes touchées
 - les rapports d'activité des forces d'intervention (police, sapeur-pompiers, etc.)
3. Le dossier départemental est transmis à une commission interministérielle dans le but d'être analysé
4. La commission interministérielle émet un avis sur l'état ou l'absence de CAT-NAT
5. En cas de reconnaissance de catastrophe naturelle par la commission interministérielle, un arrêté interministériel est publié au Journal Officiel de la République Française (JORF) .

En ce qui concerne le risque cyclonique, l'état de catastrophe naturelle est systématiquement déclaré dès lors que la vitesse des vents dépasse 145 km/h sur 10 minutes ou 215 km/h en rafales. L'organisme en charge de mesurer cette vitesse est Météo France. Il n'existe pas de seuil correspondant à une mesure sur 1 minute alors que c'est celle que nous utilisons dans ce mémoire. Cependant, comme nous avons pu le voir dans la section 2.2.3 de la première partie, en appliquant une majoration de l'ordre de 12% à une vitesse moyenne sur 10 minutes nous pouvons

passer à une mesure moyenne sur 1 minute. Ainsi, nous supposons que l'état de catastrophe naturelle est déclaré à partir d'une vitesse de vent moyenne de 160 km/h sur une minute.

Après dépassement du seuil précédent la CCR intervient en tant que réassureur. Le schéma de réassurance proposé par cette dernière se compose d'un quote-part à 50% et d'une garantie en excédent de pertes annuelles avec un plafond illimité. Cela implique que l'Etat apporte sa garantie illimitée en cas d'arrêt CAT-NAT ce qui fait que la CCR est très difficilement concurrentiable sur le marché de la réassurance privée.

7.2 Le programme de réassurance de GFA Caraïbes

Le programme de réassurance de GFA Caraïbes est différent selon que la tempête tropicale est classée en catastrophe naturelle ou en simple tempête. Dans le premier cas, l'assureur antillais ayant souscrit au programme de réassurance de la CCR cette dernière intervient par l'intermédiaire de son traité en quote-part à 50% et de son traité en excédent de perte annuel de portée illimitée. En ce qui concerne la priorité, elle correspond à 3200% de l'assiette de primes conservée par GFA Caraïbes après le paiement de la prime relative au quote-part. Cela correspond à un montant d'environ 21 millions d'euros. Le montant de sinistre restant à la charge de la compagnie après cession à la CCR est ensuite couvert par un réassureur privé *via* un traité en excédent de sinistre. La priorité de ce traité est de 2 millions d'euros et la portée de 23 millions d'euros. Ce traité peut être reconstitué trois fois moyennant une prime de reconstitution égale à 50% de la prime initiale, soit un montant de 1,4 millions d'euros, et *prorata capita*. Dans le second cas, la CCR n'intervient pas et le réassureur privé demeure le seul cessionnaire. Nous synthétisons ce programme de réassurance dans la figure ci-après à travers deux scénarios de sinistres différents et en supposant la survenance d'un seul sinistre dans l'année.

Scénario 1 (montants en millions d'euros)					
Sinistre	Cessions à la CCR			Cessions à la réassurance privée	
	Quote part 50%	Stop-loss +∞ SL 21	à charge GFA Caraïbes après CCR	Excédent de sinistre 23 XS 2	à charge GFA Caraïbes après cessions
Tempête classée CAT-NAT					
Montant brut : 100	50	29	21	19	2
	Total à charge CCR : 79				

Scénario 2 (montants en millions d'euros)					
Sinistre	Cessions à la CCR			Cessions à la réassurance privée	
	Quote part 50%	Stop-loss +∞ SL 21	à charge GFA Caraïbes après CCR	Excédent de sinistre 23 XS 2	à charge GFA Caraïbes après cessions
Tempête non CAT-NAT					
Montant brut : 26	0	0	26	23	3
	Total à charge CCR : 0				

FIGURE 61 – Fonctionnement du programme de réassurance de GFA Caraïbes

Ces deux exemples montrent que ce n'est pas forcément le sinistre brut de réassurance le plus important qui crée la charge nette la plus coûteuse pour la compagnie. Ainsi, le SCR de tempête net de réassurance est alimenté par deux principaux phénomènes : les années où le nombre de tempêtes CAT-NAT est anormalement élevé (cumul de petit coûts, dans la limite de la portée du *stop-loss* de la CCR) et les années où une (ou plusieurs) tempête flirte avec le seuil de vent requis pour être qualifiée de CAT-NAT (160 km/h en moyenne sur une minute) mais sans le dépasser.

Remarque : Dans le cas où une année contient plusieurs évènements le montant restant à la charge de la compagnie après intervention de la CCR est ventilé entre les différents évènements. Cette ventilation est effectuée au prorata des montants bruts associés à chaque évènements.

7.3 Calcul du SCR de tempêtes net de réassurance

7.3.1 Utilisation de l'outil de simulation de tempêtes

En reprenant les 10 000 simulations précédentes auxquels nous avons appliqué le programme de réassurance de GFA Caraïbes, une distribution des coûts nets de réassurance a pu être obtenu en retenant la classe de vulnérabilité numéro 3. Dans la figure suivante nous présentons l'évolution du SCR net de réassurance calculé à partir de cette distribution en fonction du nombre de simulations avant et après la prise en compte de la prime de reconstitution.

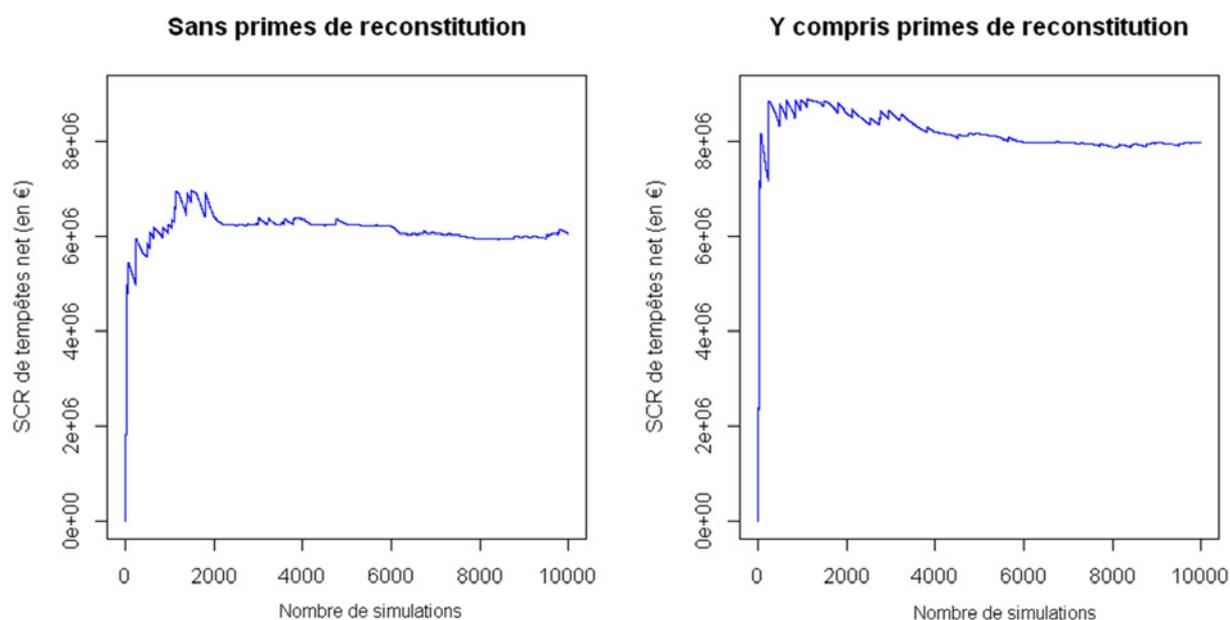


FIGURE 62 – Evolution du SCR net en fonction du nombre de simulations

A l'instar du SCR brut nous observons que, graphiquement, le capital réglementaire net de réassurance converge dans les deux cas. La prime de reconstitution faisant partie intégrante des coûts liés au risque cyclonique il faut l'intégrer dans le SCR net. Ainsi, numériquement nous retenons le résultat suivant :

SCR net de réassurance (en k€)	
Classe 3	7 960

La réassurance a donc permis de réduire fortement le coût en capital lié au risque tempête. Ce capital réglementaire net de réassurance correspond à environ 58% du chiffre d'affaires annuel

dommage aux biens, perte qui reste importante mais qui est absorbable par la compagnie.

En utilisant les courbes de vulnérabilité des classes 1 et 2 nous obtenons les résultats suivants (primes de reconstitution incluses) :

	SCR net de réassurance (en k€)	Comparaison classe 3
Classe 1	6 062	-23,8%
Classe 2	6 847	-14%

Les écarts par rapport au résultat obtenu *via* les courbes de la classe numéro 3 sont plus élevés dans le cas du calcul du SCR net que dans celui du calcul du SCR brut sans pour autant être très significatifs. Cela est dû au fait que les courbes des classes 1 et 2 démarrent plus tard (c'est-à-dire après 161 km/h et 144 km/h respectivement en ce qui concerne les courbes de structure) ce qui exclu la plupart des sinistres qui ne seraient pas couverts par la CCR (< 160 km/h). Ainsi, les tempêtes extrêmes au sens de la législation CAT-NAT constituent alors l'immense majorité des sinistres et bénéficient de fait de la couverture de réassurance proposée par la CCR. Ce n'est pas le cas des courbes de la classe 3 qui démarrent à 96 km/h laissant toutes les tempêtes de moyenne intensité provoquer des dommages certes assez faibles mais non couverts par la réassurance de l'Etat français.

7.3.2 Détermination d'un intervalle de confiance par rééchantillonnage

Un moyen simple d'obtenir un intervalle de confiance pour le SCR net précédent (calculé *via* la classe de bâtiments numéro 3) est de tirer partie de la puissance des outils de simulation modernes en utilisant une méthode de rééchantillonnage (ou *bootstrap* en anglais). Cette méthode consiste à tirer aléatoirement et avec remise des valeurs de l'échantillon des 10 000 coûts nets précédents pour produire un nombre N de nouveaux échantillons. Pour chacun de ces échantillons, nous calculons le quantile à 99,5% afin d'obtenir un échantillon Q de N quantiles. Puis, en déterminant les quantiles d'ordre $\alpha/2$ et d'ordre $1 - \alpha/2$ (respectivement notés $q_{\alpha/2}$ et $q_{1-\alpha/2}$) de la distribution associée à l'échantillon Q nous obtenons un intervalle de confiance IC_{α} de niveau $1 - \alpha$ pour le SCR net de réassurance par :

$$IC_{1-\alpha} = [q_{\alpha}; q_{1-\alpha/2}]$$

Nous avons choisi de créer un nombre $N = 10\,000$ nouveaux échantillons par *bootstrap* ce qui nous a permis d'obtenir l'intervalle à 95% (soit $\alpha = 5\%$) suivant en k€ :

$$IC_{95\%} = [7\,460; 8\,140]$$

ou encore en pourcentage du SCR net moyen (7 960 k€) :

$$IC_{95\%}/SCR_{net} = [93,7\%; 102,3\%]$$

Rappelons tout de même qu'un des principal inconvénient du rééchantillonnage réside dans le fait que les échantillons créés ne contiennent pas de valeurs supérieures à la valeur maximale observée dans l'échantillon de départ. Cela peut être problématique puisque nous nous intéressons à un quantile élevé. Cette remarque est d'autant plus valable que l'échantillon de départ est petit, ce qui n'est pas notre cas puisque nous travaillons sur 10 000 valeurs.

7.3.3 Comparaison avec la formule standard

Nous reprenons la définition du risque de base explicité dans la section 6.2.2, c'est-à-dire le montant auquel on applique le facteur de 1,2. Nous donnons ci-dessous la valeur de ce risque pour chacune des îles.

	Facteur de risque	Facteur de base (en k€)
Guadeloupe	2,74%	40 747
Martinique	3,19%	53 075
Saint-Martin	5,16%	5 727
Saint-Barthélemy	5,16%	2 567

Pour chacun de ces risques de base deux scénarios de deux évènements sont à étudier, le plus défavorable en terme de coût constituant le SCR net de la compagnie. Le premier scénario consiste à répartir le facteur de 1,2 selon un coefficient 0,4 et un coefficient 0,8 à appliquer au risque de base de chacune des îles. Les deux montants obtenus correspondent en fait à la survenance de deux évènements distincts dans l'année. Ensuite, le programme de réassurance de la compagnie est appliqué à chacun de ces deux montants bruts. Le second scénario est similaire mais en répartissant le facteur 1,2 selon un coefficient 1 et un coefficient 0,2. Ces deux scénarios permettent de tester les programmes de réassurance de la compagnie : le premier teste l'impact d'une sur-fréquence de sinistres de moyenne amplitude et le second teste l'impact d'un sinistre d'amplitude très importante.

Puisque les facteurs de risques correspondent à des sinistres exceptionnels, nous pouvons supposer que chacun des évènements constitue une catastrophe naturelle au sens du régime CAT-NAT. Dans les tableaux ci-dessous nous présentons les résultats obtenus.

Scénario 1 (facteurs 0,4 et 0,8)

En k€	Coûts bruts		A charge GFA Caraïbes après cessions CCR		Répartition des charges post CCR		A charge GFA Caraïbes après cessions à la réassurance privée	
	Evènement 1	Evènement 2	Après QP	Après SL	Evènement 1	Evènement 2	Evènement 1	Evènement 2
Guadeloupe	16 299	32 598	61 270	21 000	7 000	14 000	2 000	2 000
Martinique	21 230	42 460						
Saint-Martin	2 291	4 581						
Saint-Barthélemy	1 027	2 054						
Total	122 539							

Scénario 2 (facteurs 0,2 et 1)

En k€	Coûts bruts		A charge GFA Caraïbes après cessions CCR		Répartition des charges post CCR		A charge GFA Caraïbes après cessions à la réassurance privée	
	Evènement 1	Evènement 2	Après QP	Après SL	Evènement 1	Evènement 2	Evènement 1	Evènement 2
Guadeloupe	8 149	40 747	61 270	21 000	3 500	17 500	2 000	2 000
Martinique	10 615	53 075						
Saint-Martin	1 145	5 727						
Saint-Barthélemy	513	2 567						
Total	122 539							

FIGURE 63 – Calcul du SCR net par la formule standard

Les deux scénarios donnent des résultats identiques. De plus, nous remarquons que le traité en excédent de sinistres est consommé deux fois. Ainsi, deux primes de reconstitutions sont à ajouter aux coûts nets soit 2,8 millions d'euros. Finalement nous obtenons le SCR net suivant :

SCR net de réassurance (en k€)
6 800

Ce résultat est similaire à celui résultant de notre outil de modélisation via l'utilisation des courbes de la classe numéro 2. En revanche, il est inférieur au SCR issu de la vision prudente retenu dans cette étude (classe 3) et supérieur à celui de la vision la moins coûteuse (classe 1).

Chapitre 8

Incorporation de l'outil dans le modèle interne

Nous nous proposons d'achever ce mémoire par l'élaboration d'un outil de modélisation simplifié du risque de tempête pour GFA Caraïbes à partir des résultats obtenus dans les chapitres précédents. Cette outil permettra d'éviter la lourdeur des simulations et d'incorporer le calcul du SCR de tempête brut dans le logiciel de modélisation interne du Groupe Generali.

La méthode repose sur une approche fréquence-coût classique. Dans un premier temps, nous ajustons une loi de probabilité sur le nombre d'évènements impactant la compagnie dans son ensemble à horizon un an. Puis, dans un second temps nous ajustons une loi sur les coûts des sinistres générés. Enfin, nous mettons en œuvre l'outil afin de le confronter au simulateur de tempêtes développé précédemment pour validation de ses résultats.

Remarque préalable : dans toute la suite nous parlerons de "simulateur complet" pour désigner l'outil de modélisation développé dans les chapitres précédents alors que nous parlerons d'"outil de simulation simplifié" pour désigner l'outil développé dans ce chapitre.

8.1 Loi du nombre de tempête impactant GFA Caraïbes à horizon un an

Nous considérons l'échantillon contenant le nombre de tempêtes annuelles impactant GFA Caraïbes. Précisons que lorsqu'une tempête donnée traverse plusieurs îles au cours de sa trajectoire cette dernière ne compte que pour une occurrence. Nous obtenons l'histogramme fréquentiel empirique suivant :

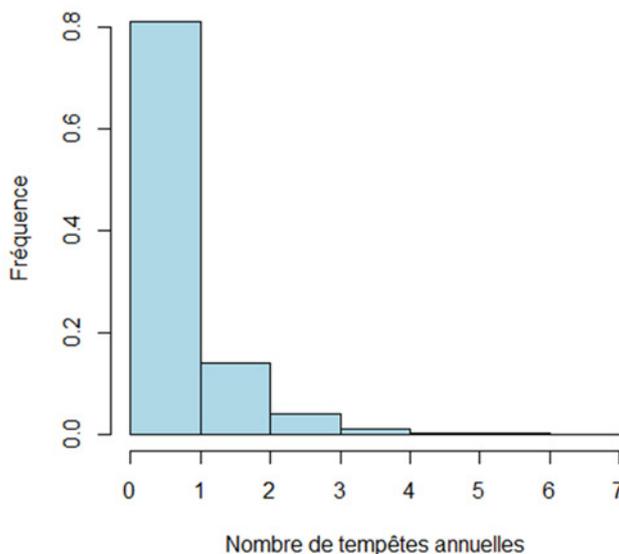


FIGURE 64 – Histogramme fréquentiel du nombre de tempêtes impactant GFA Caraïbes

Les statistiques descriptives associées sont données ci-après :

	Min	1 ^{er} quartile	Moyenne	Variance	3 ^{ème} quartile	Max
Nombre de tempêtes	0	0	0,79	0,85	1	7

Nous donnons dans le tableau suivant la log-vraisemblance ainsi que les estimateurs des paramètres obtenus à partir du logiciel R pour la loi de Poisson et la loi binomiale négative :

Distribution	Log-vraisemblance	Paramètres estimés
$\mathcal{P}(\lambda)$	-11856,74	$\hat{\lambda} = 0,79$
$\mathcal{BN}(m, p)$	-11844,62	$\hat{m} = 10,9$ et $\hat{p} = 0,93$

Au vu des résultats, nous retenons la loi binomiale négative puisque cette dernière maximise la log-vraisemblance. Ce choix est cohérent avec le fait que la variance de l'échantillon soit supérieure à la moyenne. Pour confirmer cela nous avons effectué un test du chi-deux avec 6 classes. Le résultat du test au niveau de 5% est présenté dans le tableau suivant :

Statistique	Seuil critique	p-valeur
1,1245	11,0705	0,9519

La p-valeur très largement supérieure à 5% indique que nous pouvons accepter l'hypothèse selon laquelle l'échantillon suit une loi binomiale négative. Graphiquement l'adéquation semble également de très bonne qualité comme le montre la figure ci-après. Cette dernière comporte l'histogramme l'échantillon ainsi que l'histogramme théorique d'une $\mathcal{BN}(10, 9; 0, 93)$

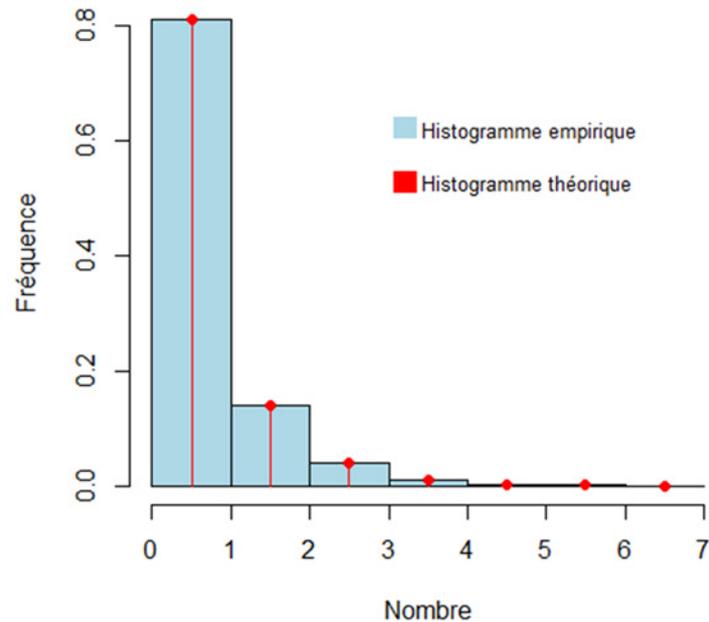


FIGURE 65 – Adéquation graphique

8.2 Loi des coûts bruts générés par les tempêtes

Nous cherchons ici à ajuster une loi de probabilité sur les coûts provoqués par une tempête générée par le simulateur complet. Notons tout d'abord qu'une tempête ne provoque pas systématiquement des dommages puisque les courbes de vulnérabilité utilisées précédemment sont croissantes à partir d'un certain seuil. Ainsi, une manière de modéliser cette particularité consiste à utiliser une loi de Bernoulli dont le paramètre correspond au taux de tempêtes provoquant un montant de dommage strictement positif. Soit B la variable aléatoire prenant la valeur 0 si aucun dommage n'est provoqué et 1 sinon. Empiriquement nous trouvons :

$$\begin{cases} \mathbb{P}[B = 0] = 0,5511 \\ \mathbb{P}[B = 1] = 0,4489 \end{cases}$$

Sachant que le coût est non nul ($B = 1$) il faut alors modéliser le coût provoqué par la tempête. Nous donnons-ci dessous l'allure des coûts ainsi que l'histogramme fréquentiel associé.

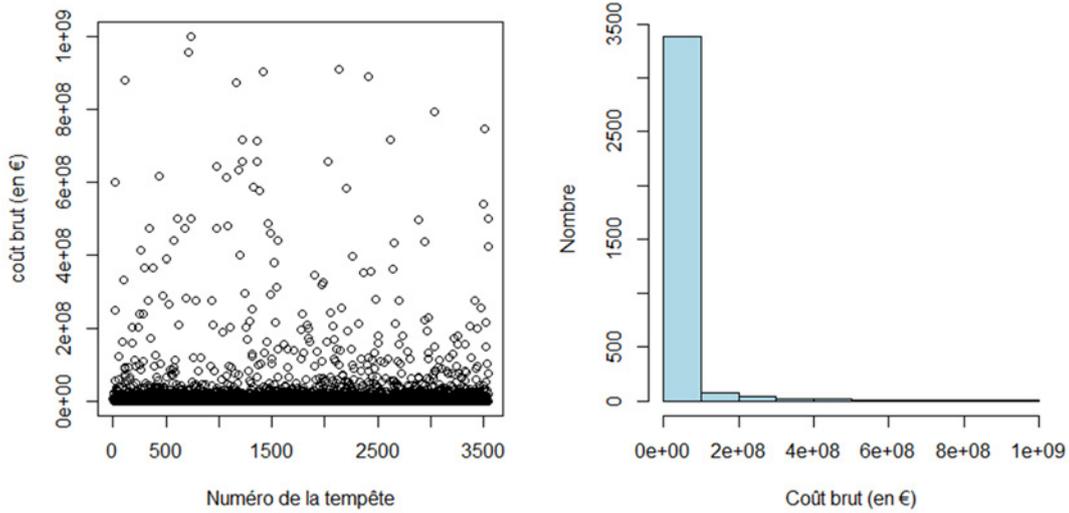


FIGURE 66 – Coûts non nuls et histogramme fréquentiel associé

A priori il semble qu'une loi de Pareto généralisée soit adéquate pour modéliser les excès au delà d'un certain seuil. Nous rappelons que la fonction de répartition de cette loi est donnée par :

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi} & \text{si } \xi \neq 0 \\ 1 - \exp(-x/\beta) & \text{si } \xi = 0 \end{cases}$$

où $\beta > 0$, et $x \geq 0$ si $\xi \geq 0$ et $0 \leq x \leq -\beta/\xi$ si $\xi < 0$. Le cas $\xi = 0$ correspond simplement au cas de la distribution exponentielle alors que le cas $\xi < 0$ correspond au cas d'une distribution de Pareto de type II.

Le seuil peut être déterminé en traçant ce que l'on appelle la *mean excess function* définit pour une variable aléatoire X positive par :

$$\forall u > 0, \quad e(u) = \mathbb{E}[X - u | X > u]$$

Etant donné un échantillon de la variable X indépendant et identiquement distribué (X_1, \dots, X_n) l'estimateur empirique de la fonction précédente est donné par :

$$\forall u > 0, \quad \hat{e}(u) = \frac{\sum_{i=1}^n (X_i - u) \mathbb{1}_{\{X_i > u\}}}{\sum_{i=1}^n \mathbb{1}_{\{X_i > u\}}}$$

L'intérêt d'une telle fonction réside dans le fait qu'elle est linéaire dans le cas de la loi de Pareto généralisée. Plus précisément nous avons :

$$e(u) = \frac{\beta}{1 - \xi} + \frac{\xi}{1 - \xi} u$$

avec $0 \leq u < +\infty$ si $0 \leq \xi < 1$ et $0 \leq u \leq -\beta/\xi$ si $\xi < 0$.

Le tracé de la *mean excess function* dans le cas de notre échantillon donne le résultat graphique suivant :

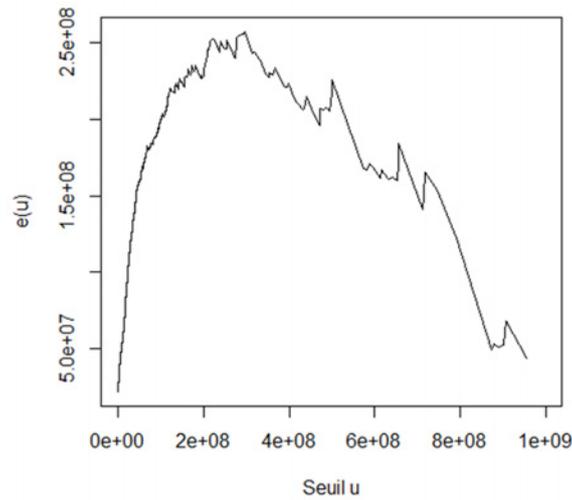


FIGURE 67 – Mean excess plot des coûts bruts

Il apparaît une tendance linéaire à pente négative à partir du seuil $S_1 = 250$ M€ ce qui fait penser à une loi de Pareto de type II. Cette loi est un cas particulier de la loi de Pareto généralisée présentée précédemment. La détermination des paramètres par maximum de vraisemblance donne les résultats suivants :

Seuil (k€)	Paramètres
250 000	$\hat{\xi} = -0,1959$ et $\hat{\beta} = 2,934 \cdot 10^8$

Pour valider la distribution nous avons effectué un test de Kolmogorov-Smirnov dont le résultat est donné dans le tableau ci-après. Nous donnons aussi la représentation du graphe quantile-quantile à titre de test graphique.

Statistique	Seuil	p-valeur
0,0611	0,165	0,962

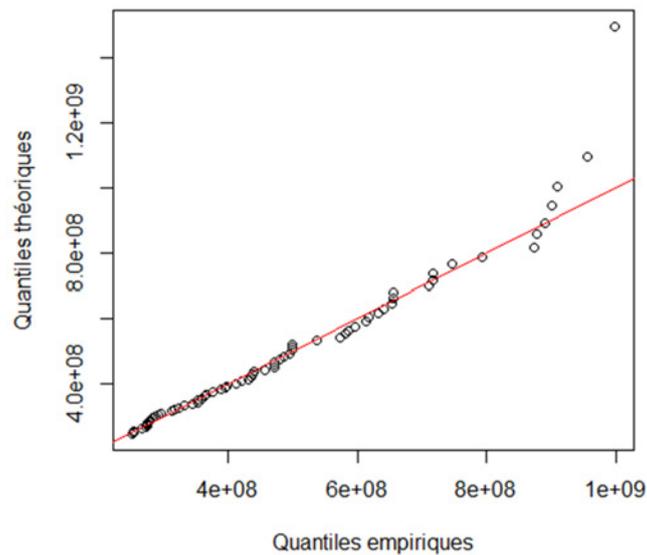


FIGURE 68 – Q-Q plot des excès au delà de 250 M€

La p-value très largement supérieure à 5% ainsi que le graphe quantile-quantile nous indiquent que ne devons pas rejeter l'hypothèse selon laquelle les excès au delà de S_1 suivent une loi de Pareto généralisées.

Il s'agit alors de modéliser les coûts inférieurs au seuil S_1 . Si l'on trace de nouveau la *mean excess function* pour ces coûts nous retrouvons une allure similaire à la figure 67 avec une tendance linéaire à pente négative au delà du seuil $S_2 = 60$ M€. En outre, les excès sont encore modélisables par une Pareto généralisée. En fait, nous montrons en annexe E que ce raisonnement s'applique 8 fois de suite *via* 8 seuils différents que nous désignerons par (S_1, S_2, \dots, S_8) . Dans la suite nous noterons C la variable aléatoire désignant les coûts bruts avec :

$$C = S + E$$

où E représente la variable aléatoire désignant les excès au delà du seuil S également aléatoire.

Nous donnons ci-dessous la valeur de chacun des seuils S_j ainsi que la probabilité de retenir chacun d'entre-eux, autrement dit la probabilité que le coût brut C soit supérieur à S_j et inférieur à S_{j+1} (en fixant S_9 à $+\infty$) :

j	1	2	3	4	5	6	7	8
S_j (en M€)	250	60	28	6	2,2	0,67	0,08	0
$\mathbb{P}[S = S_j]$	1,92%	4,63%	5,25%	28,20%	14,86%	17,40%	19,83%	7,91%

Finalement, nous pouvons caractériser la distribution des excès E par la fonction de répartition conditionnelle suivante :

$$\forall x > 0, \quad \mathbb{P}[E \leq x | B = i, S = S_j] = \begin{cases} 0 & \text{si } B = 0 \\ G_{\xi_j, \beta_j}(x) & \text{si } B = 1 \text{ et } S = S_j \end{cases}$$

Notation : dans la suite nous noterons P_s le vecteur de probabilité contenant les $\mathbb{P}[S = S_j]$.

8.3 Mise en œuvre de l'outil

A partir des deux distributions précédentes nous sommes en mesure d'effectuer une simulation de Monte-Carlo afin d'obtenir un échantillon de coûts bruts pour comparaison avec les coûts issus du simulateur. La procédure de simulation d'un coût annuel brut est donnée dans l'algorithme ci-dessous.

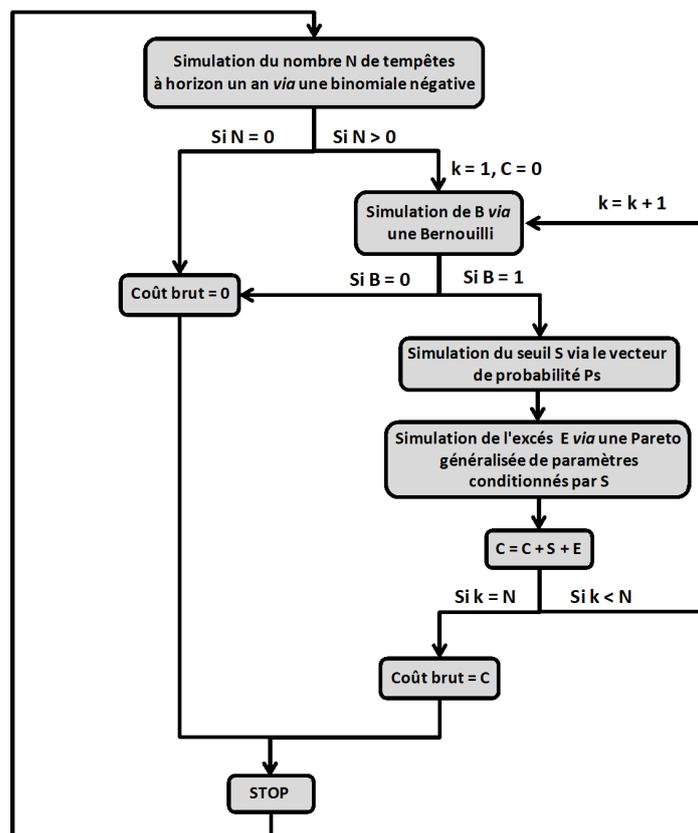


FIGURE 69 – Algorithme de simulation du coût brut annuel

Nous avons effectué 100 000 simulations *via* l'algorithme précédent. Un test graphique d'adéquation entre l'échantillon généré et l'échantillon initial obtenu grâce à l'outil de modélisation complet peut être effectué au moyen d'un Q-Q plot :

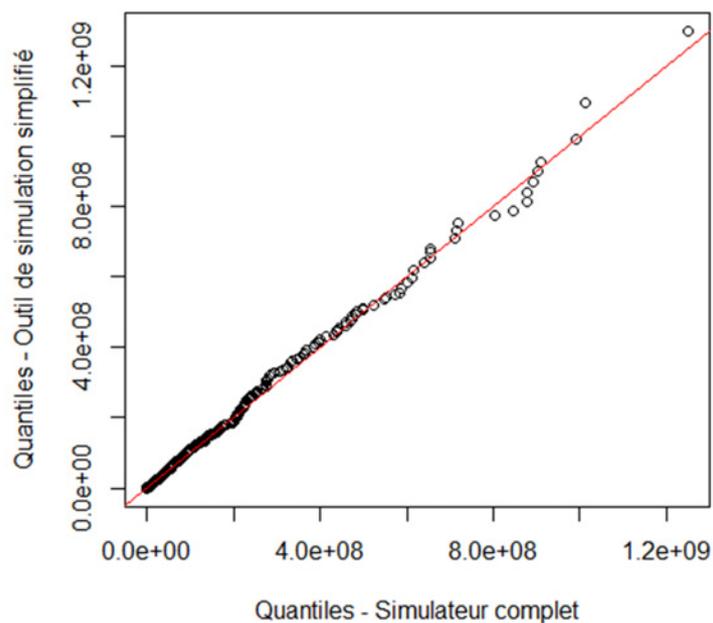


FIGURE 70 – Q-Q plot des deux échantillons

Graphiquement, l'adéquation semble donc de bonne qualité. Un test de Kolmogorov-Smirnov entre les deux échantillons nous conforte dans notre choix puisque nous obtenons une p-valeur de 5,9%. L'évolution du quantile à 99,5% issu de l'échantillon généré en fonction du nombre de simulations est donné ci-après en comparaison avec ce que nous avons obtenu par le simulateur complet.

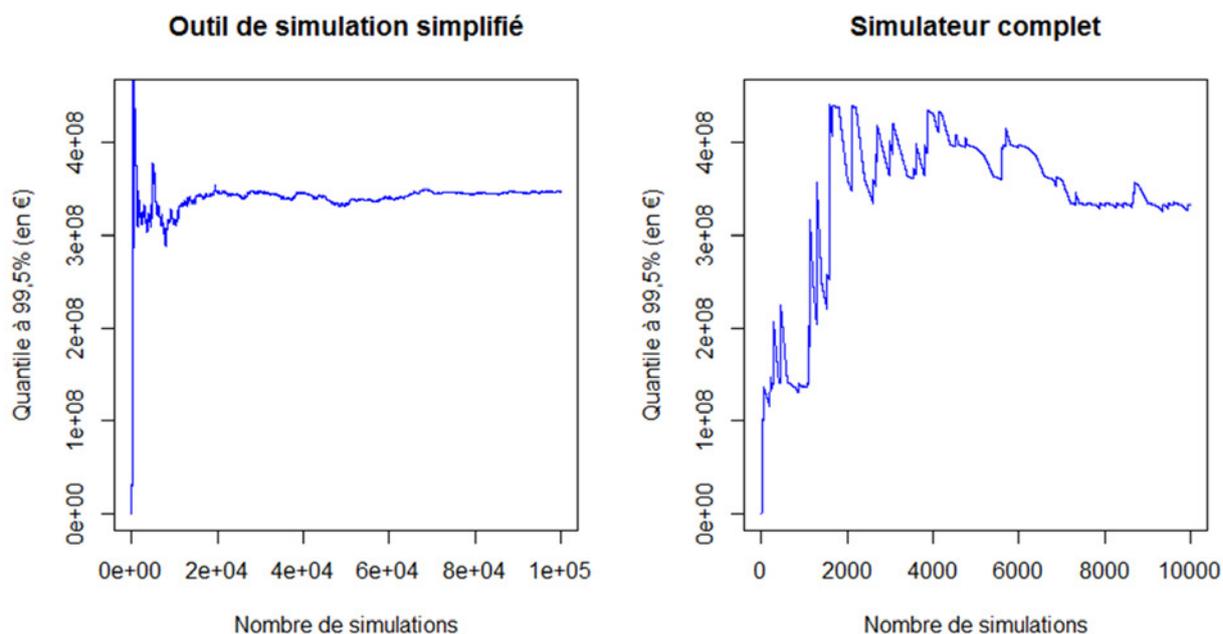


FIGURE 71 – Evolution des quantiles à 99,5%

La convergence est un peu moins rapide dans le cas de l'outil simplifié mais l'obtention de l'échantillon nécessite un temps de calcul très inférieur (de l'ordre de quelques minutes alors qu'il est de plusieurs dizaines d'heures dans le cas du simulateur complet). Pour finir, nous présentons ci-dessous les statistiques descriptives principales des échantillons générés selon les deux méthodes :

	Min	Médiane	Moyenne	Quantile à 99,5%	Max
Outil simplifié (en k€)	0	0	7 763	336 952	1 326 834
Simulateur complet (en k€)	0	0	7 803	332 381	1 250 260
Ecart absolu	0	0	0,52%	1,37%	6,12%

Nous obtenons donc des résultats similaires par les deux méthodes ce qui vient une nouvelle fois valider le modèle développé tout au long de cet ultime chapitre.

Conclusion

La modélisation du risque de tempêtes est un domaine en plein essor qui va bien au-delà des sciences actuarielles compte tenu notamment de l'enjeu humain qu'elle implique. La principale difficulté rencontrée par les modèles réside dans la complexité et la multitude des éléments qui rentrent en jeu dans la formation des ouragans alors que les scientifiques n'ont pas encore percé tous les secrets du phénomène. Ainsi, nous avons opté dans ce mémoire pour une approche statistique qui, avec un minimum de connaissances sur l'aspect thermodynamique des tempêtes tropicales, permet de généraliser ce que l'on a observé au cours des 30 dernières années aux tempêtes futures.

Le développement d'un tel modèle trouve son application dans la nouvelle politique de gestion des risques prônée par la récente directive européenne Solvabilité 2 concernant les sociétés d'assurance avec notamment le nouveau mode de calcul du capital réglementaire risque par risque (SCR).

Le modèle qui a été développé couvre le bassin Atlantique Nord et permet la simulation de saisons cycloniques dont le nombre annuel de tempêtes est distribué selon une loi binomiale négative avec pour finalité le calcul du SCR de tempêtes. Les trajectoires de ces dernières ont été simulées par une chaîne de Markov avec des résultats très probants puis, deux modèles de pressions ont été utilisés le long de celles-ci. Le premier modèle, dit de régression multiple, a été appliqué dans les étendues d'eaux avec un pourcentage de variance expliquée très satisfaisant. Le second, applicable lors de l'atterrissage du cyclone dans les terres, est un modèle de décroissance exponentielle du déficit de pression dont la constante de désintégration est distribuée selon une loi gamma. Deux régressions ont ensuite été développées pour relier les pressions aux vitesses des vents et aux rayons de vitesse maximale puis, la distribution des vitesses atteignant les îles couvertes a été déduite *via* un gradient de vitesse.

Les dommages provoqués par les tempêtes simulées ont été obtenus par l'intermédiaire d'un ensemble de courbes de vulnérabilité issues d'une étude menée dans les îles Vierges britanniques. Ces dernières nous ont permis d'associer à chaque vitesse de vent un pourcentage de destruction des structures et des contenus de deux catégories de bâtiments : les bâtiments résidentiels et les bâtiments industriels, commerciaux et gouvernementaux. Pour chacune de ces deux catégories, nous avons fait l'hypothèse que tous les bâtiments avaient une structure similaire en justifiant ce choix par l'absence d'informations précises concernant la typologie des bâtiments assurés ainsi que par des tests de sensibilité. Ensuite, l'application des taux obtenus aux sommes assurées par la compagnie nous a permis d'établir une distribution des coûts annuels générés par les saisons cycloniques. Un capital réglementaire brut pour le risque de tempête en a été déduit et il s'avère être 2,7 fois supérieur au résultat obtenu *via* la formule standard. Enfin, la prise en compte de la réassurance réduit très fortement le SCR brut et aboutit à un résultat très proche de la formule standard. Cela s'explique essentiellement par la garantie illimitée apportée par la CCR lors de la survenance de tempêtes extrêmes.

Notre étude s'est achevée par la mise en œuvre d'un modèle simplifié construit au moyen de distributions simples et facilement implémentables dans un outil informatique. Cependant, la relative simplicité du modèle final ne doit pas masquer le fait que l'étude repose sur des hypothèses fortes notamment en ce qui concerne la partie dommage. Une quantification des taux de destruction associés aux tempêtes historiques ainsi qu'une étude approfondie du portefeuille détenu par la compagnie permettrait d'obtenir des courbes de vulnérabilité plus fiables et la prise en compte d'un paramètre de volatilité. Pour conclure, précisons que cette étude peut être transposée à d'autres bassins cycloniques comme par exemple le bassin Indien dans lequel se trouve Prudence Créole une autre filiale de Generali France.

Pistes d'amélioration

Ce mémoire pose les grands principes d'une modélisation statistique du risque cyclonique dans le bassin Atlantique nord et pourra à l'avenir faire l'objet d'améliorations. Nous donnons ci-dessous quelques pistes de réflexions concernant les grands axes de l'étude :

- **Modélisation des trajectoires** : Un des inconvénients de l'utilisation d'une chaîne de Markov réside dans le fait que nous avons dû procéder à un quadrillage du bassin Atlantique selon une résolution $1^\circ \times 1^\circ$. Ce choix implique une absence d'information précise concernant la latitude et la longitude exacte de la tempête lors de son déplacement. Cela implique aussi de considérer qu'une île occupe entièrement la case qui lui est associée. Pour pallier à cela des modèles de régression multiple prenant comme variables explicatives la direction ainsi que la vitesse du cyclone ont été développés récemment par VICKERY. A l'instar du modèle de pression développé dans le mémoire, ces modèles ont un paramétrage qui varie en fonction du lieu géographique (paramétrage selon un quadrillage $5^\circ \times 5^\circ$ par exemple).
- **Caractéristiques des tempêtes le long des trajectoires** : La modélisation du gradient de vitesse pourrait être amélioré en optant pour ce que l'on appelle une modélisation par vortex de Rankine. Ce modèle est issu d'une équation aux dérivées partielles de mécanique des fluides. Il consiste à écrire la vitesse v du vent à une distance r du centre du cyclone comme suit :

$$v_1(r) = \begin{cases} v_{max} \left(\frac{r}{R_{max}} \right)^B & \text{si } r < R_{max} \\ v_{max} \left(\frac{R_{max}}{r} \right)^B & \text{si } r \geq R_{max} \end{cases}$$

avec B un paramètre de forme ($\simeq 0,5$), v_{max} la vitesse au mur de l'œil et R_{max} le rayon de vitesse maximale. Ce modèle peut être ajusté pour prendre en compte l'asymétrie du cyclone due à sa vitesse de déplacement v_d :

$$v_2(r) = \sqrt{(v_1^2(r) \cos(\theta_r))^2 + \left[V \sin(\theta_r) + \frac{v_d r R_{max}}{(R_{max}^2 + r^2)} \right]^2}$$

avec θ_r la direction du cyclone par rapport à une référence. Ce modèle suppose de connaître la vitesse de déplacement et la direction du cyclone à chaque date et serait parfaitement compatible avec le modèle de trajectoire décrit dans le point précédent.

- **Modèle de vulnérabilité** : Un jeu de courbe de vulnérabilité empirique peut être construit en effectuant une expertise *post* tempête du taux de destruction d'un échantillon de bâtiments représentatif du portefeuille détenu par l'assureur. A ce titre, nous pouvons citer le travail du chercheur allemand HENEKA [9] qui propose la forme suivante pour modéliser les taux de destruction en fonction de la vitesse de vent v :

$$tx(v) = \left(\frac{v - v_{min}}{v_{max} - v_{min}} \right)^\alpha$$

où α est un paramètre à déterminer, v_{min} est la vitesse à partir de laquelle les bâtiments commencent à subir des dommages et v_{max} la vitesse pour laquelle le bâtiment est complètement détruit. Ces deux vitesses sont supposées corrélées à 100% :

$$v_{max} = v_{min} + \Delta v$$

L'incertitude concernant le taux de destruction est introduit par le fait que v_{min} soit supposée gaussienne. Bien sûr le développement d'un tel modèle suppose d'avoir à disposition un historique conséquent.

Annexes

A Calcul de l'intensité relative par la méthode de Darling

On cherche à construire un indicateur de l'intensité d'une tempête à un instant donné sachant sa pression centrale et la température à la surface de l'eau. L'approche de DARLING [3] est basée sur les calculs d'EMMANUEL [6] en omettant les termes d'ordre deux. La théorie d'EMMANUEL s'appuie sur l'idée que, sous les conditions atmosphériques et thermiques dans lesquelles se trouve le cyclone, sa pression centrale ne peut descendre en dessous d'une certaine valeur. En comparant cette valeur à la pression centrale régnant réellement dans le cyclone on obtient un indicateur théoriquement compris entre 0 et 1. Dans ce mémoire nous nous limiterons simplement à donner les étapes de calcul de l'intensité relative qui reposent en grande partie sur des concepts de thermodynamique.

On suppose que la température à la surface de l'étendue d'eau traversée par le cyclone vaut T_s exprimée en Kelvin (K), que la température de la couche supérieure de la troposphère vaut $T_0 = 203$ K et que la pression centrale vaut P . Nous prendrons une pression atmosphérique dans des conditions normales égale à $P_{atm} = 1013$ hPa.

On commence par déterminer l'efficacité du cyclone par :

$$\epsilon = \frac{T_s - T_0}{T_s}$$

puis on calcule la pression de vapeur saturante :

$$e_s = 6,112 \times \exp\left(\frac{17,67(T_s - 273)}{T_s - 29,5}\right)$$

et la chaleur latente de vaporisation :

$$L_v = 2,5 \times 10^6 - 2320(T_s - 273)$$

On fixe l'humidité relative de l'air à $Rh = 0,75$. La pression partielle de l'air sec ambiant en surface est alors donnée par :

$$p_{da} = P_{atm} - (Rh \times e_s)$$

On définit deux valeurs A et B par les relations suivantes :

$$\begin{cases} A = \frac{\epsilon L_v T_s}{(1 - \epsilon) R_v T_s p_{da}} \\ B = Rh \times \left[1 + \frac{e_s \ln(Rh)}{A p_{da}} \right] \end{cases}$$

et on résout l'équation d'inconnue x suivante :

$$\ln(x) = -A \left(\frac{1}{x} - B \right)$$

dont la solution est donnée par :

$$x = \exp[LW(-A \exp(-A \times B)) + A \times B]$$

où $LW : u \mapsto LW(u)$ est la fonction W de Lambert c'est-à-dire la réciproque de la fonction $w \mapsto we^w$. On est alors en mesure de déterminer la valeur minimale atteignable par la pression centrale (de l'air sec) du cyclone pour les conditions thermiques dans lequel il se trouve par :

$$p_{dc} = x p_{da}$$

Finalement, on définit l'intensité relative associée à la pression centrale P pour une température à la surface de l'eau T_s par :

$$IR = \frac{p_{da} - (P - e_s)}{p_{da} - p_{dc}}$$

Pour des raisons évoquées dans le mémoire nous retiendrons plutôt la définition suivante :

$$IR = \min \left[1; \max \left(0, 01; \frac{p_{da} - (P - e_s)}{p_{da} - p_{dc}} \right) \right]$$

Réciproquement, pour passer de l'intensité relative à la pression centrale du cyclone nous inverserons simplement cette formule en utilisant les différents éléments thermodynamiques calculés précédemment :

$$P = p_{da} + e_s - IR \times (p_{da} - p_{dc})$$

B Détermination de la distance orthodromique entre deux points

On considère la Terre comme étant une sphère parfaite de rayon R_T et on se place dans le cadre de la figure suivante.

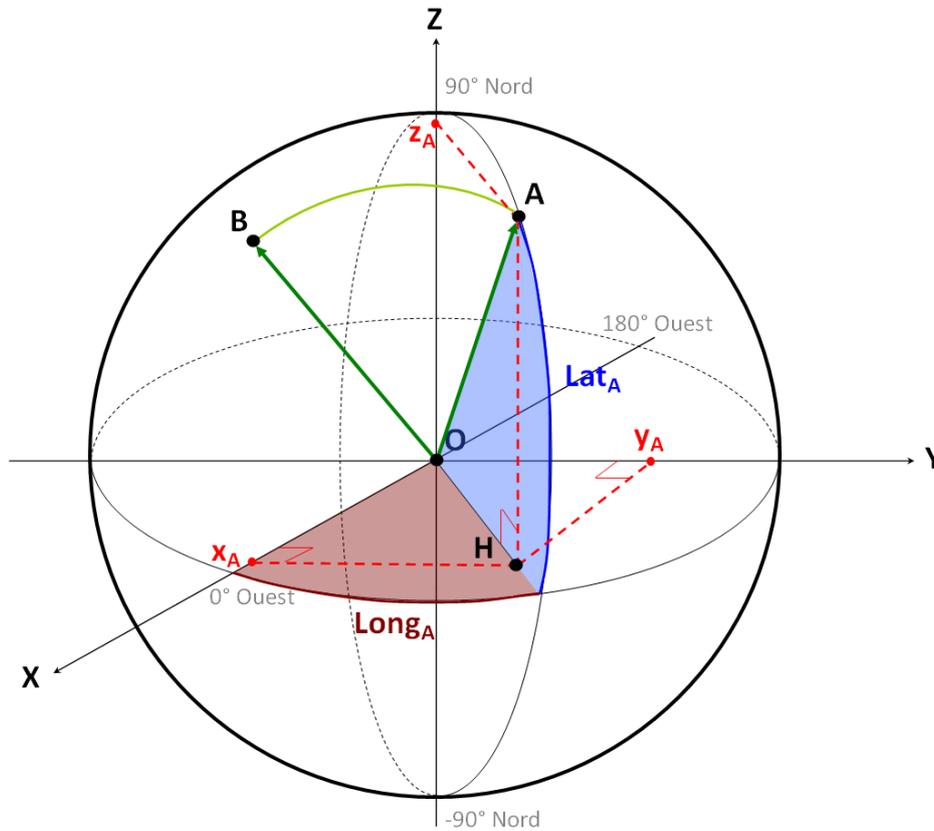


FIGURE 72 – Distance entre deux points d'une sphère

On cherche à déterminer la distance orthodromique entre le point A et le point B repérés par leur latitude et leur longitude en radians. Cette distance est le plus court chemin de surface entre les deux points, c'est-à-dire la longueur de l'arc de cercle qui passe par A et B. Compte tenu des hypothèses on a :

$$- \|\vec{OA}\| = \|\vec{OB}\| = R_T$$

$$- \widehat{AB} = R_T \times \widehat{AOB}$$

L'angle \widehat{AOB} est l'inconnu du problème. Pour le déterminer, l'idée est d'écrire le produit scalaire entre \vec{OA} et \vec{OB} de deux façons différentes et d'égaliser les expressions. On a d'une part :

$$\vec{OA} \cdot \vec{OB} = \|\vec{OA}\| \times \|\vec{OB}\| \times \cos(\widehat{AOB}) = R_T^2 \cos(\widehat{AOB})$$

et d'autre part :

$$\vec{OA} \cdot \vec{OB} = x_A x_B + y_A y_B + z_A z_B$$

avec :

$$\begin{cases} x_A = \|\overrightarrow{OH}\| \times \cos(Long_A) = R_T \cos(Lat_A) \cos(Long_A) \\ y_A = \|\overrightarrow{OH}\| \times \sin(Long_A) = R_T \cos(Lat_A) \sin(Long_A) \\ z_A = R_T \sin(Lat_A) \end{cases}$$

et de même :

$$\begin{cases} x_B = \|\overrightarrow{OH}\| \times \cos(Long_B) = R_T \cos(Lat_B) \cos(Long_B) \\ y_B = \|\overrightarrow{OH}\| \times \sin(Long_B) = R_T \cos(Lat_B) \sin(Long_B) \\ z_B = R_T \sin(Lat_B) \end{cases}$$

Ainsi, en égalant les deux expressions du produit scalaire il vient :

$$R_T^2 \cos(\widehat{AOB}) = R_T^2 \left[\cos(Lat_A) \cos(Lat_B) (\cos(Long_A) \cos(Long_B) + \sin(Long_A) \sin(Long_B)) \right. \\ \left. + \sin(Lat_A) \sin(Lat_B) \right]$$

Finalement, en simplifiant par R_T^2 et en utilisant la formule de trigonométrie donnant le cosinus d'une différence on obtient :

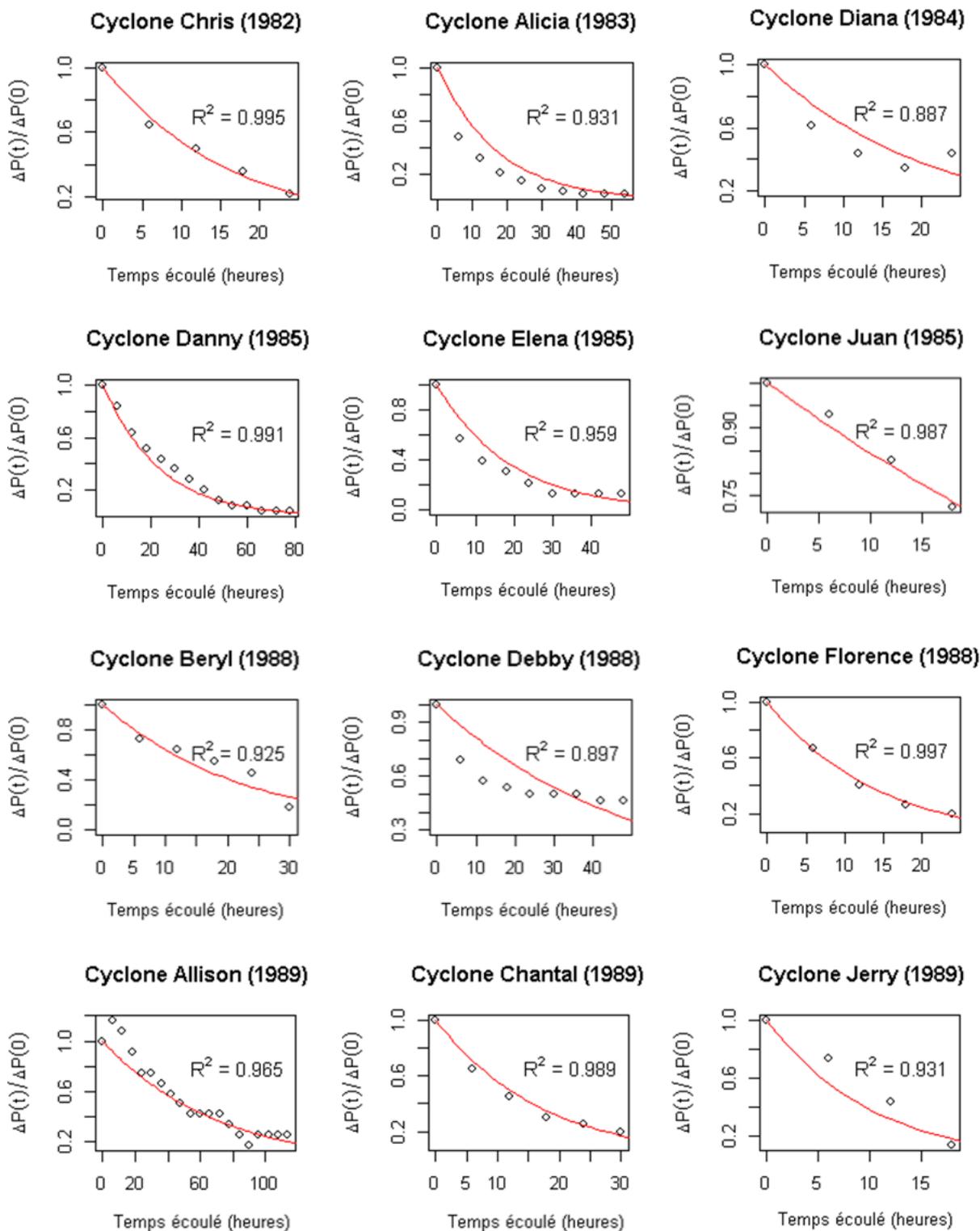
$$\widehat{AOB} = \arccos \left[\cos(Lat_A) \cos(Lat_B) \cos(Long_A - Long_B) + \sin(Lat_A) \sin(Lat_B) \right]$$

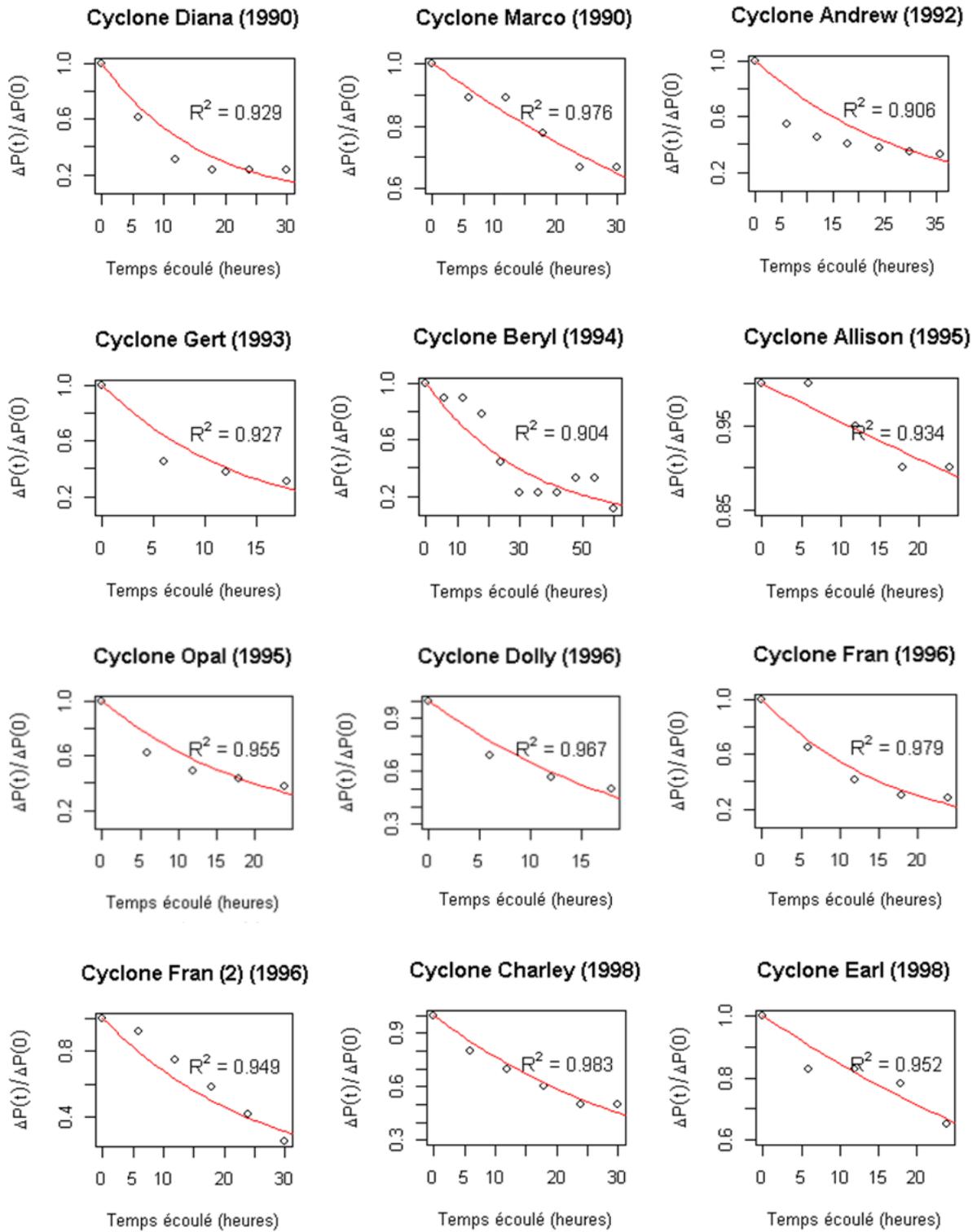
d'où :

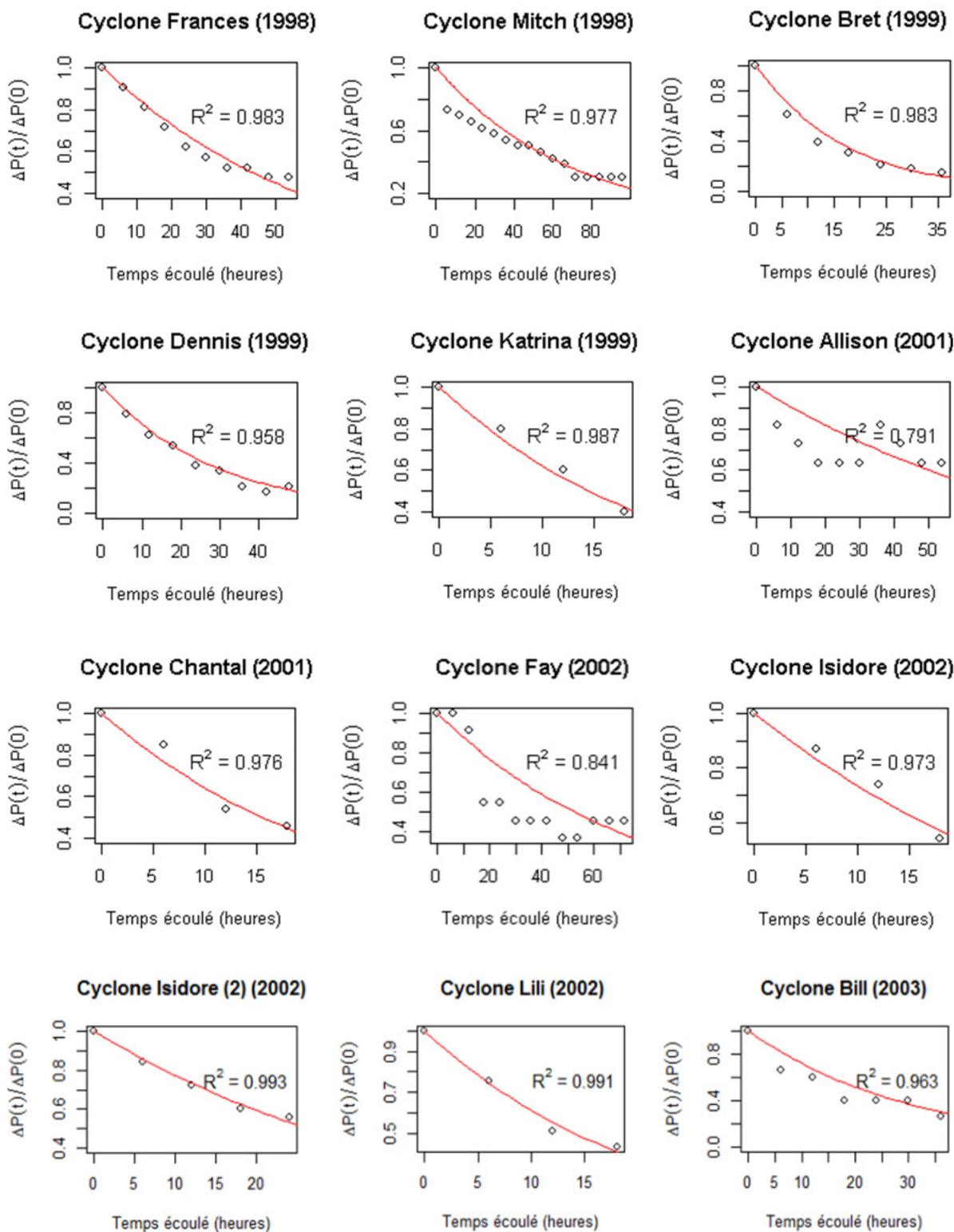
$$\widehat{AB} = R_T \times \arccos \left[\cos(Lat_A) \cos(Lat_B) \cos(Long_A - Long_B) + \sin(Lat_A) \sin(Lat_B) \right]$$

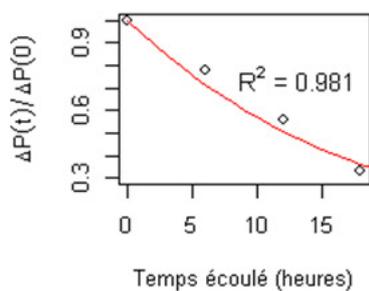
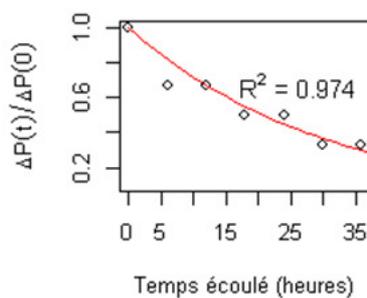
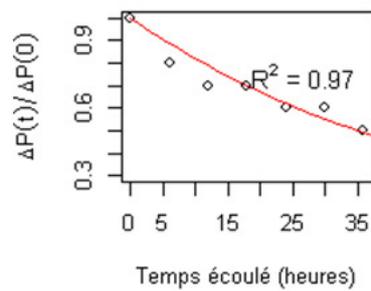
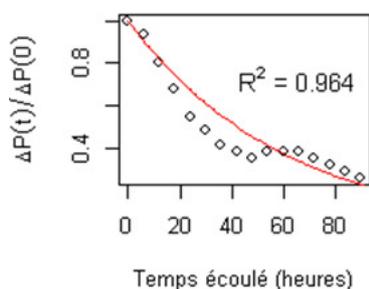
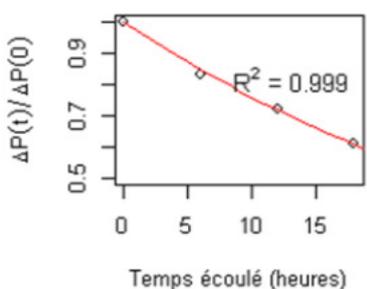
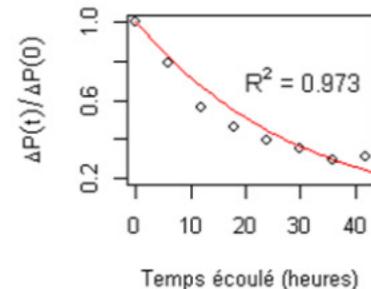
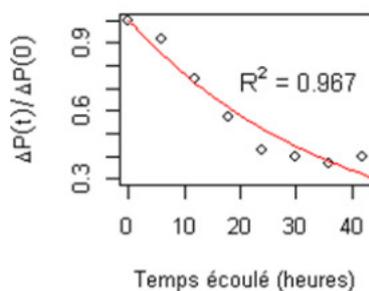
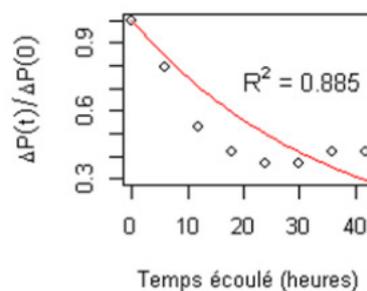
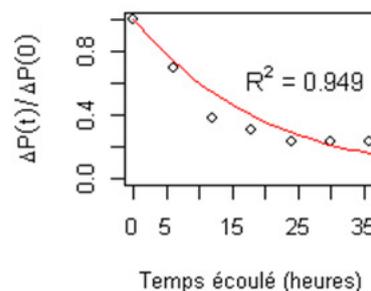
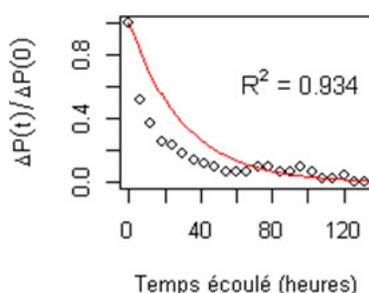
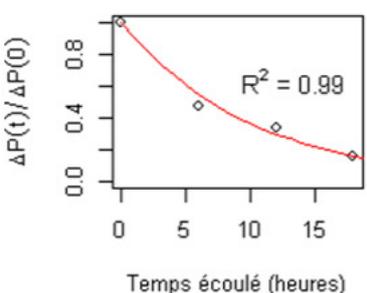
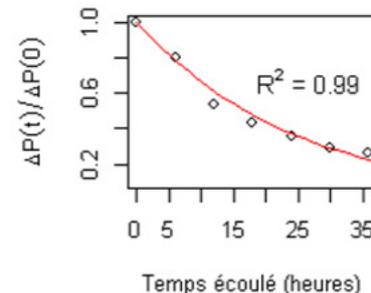
C Régressions retenues dans le cadre du modèle de désintégration

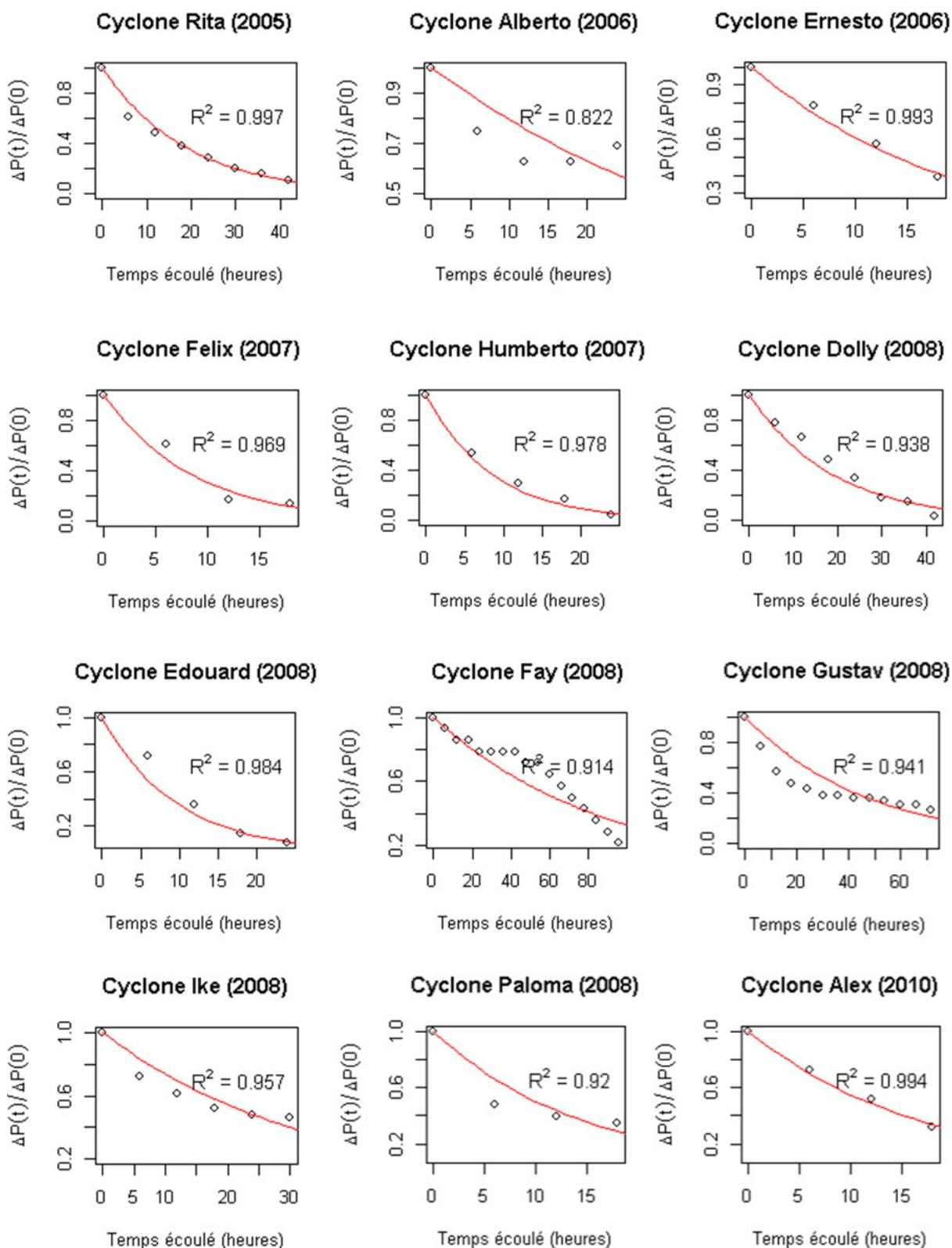
Les courbes ci-dessous représentent les 64 régressions exponentielles retenues dans le cadre du modèle de désintégration. Les points blancs correspondent aux données empiriques et les courbes rouges sont les ajustement au modèle. Dans chaque cas nous donnons aussi le coefficient de détermination ajusté associé.



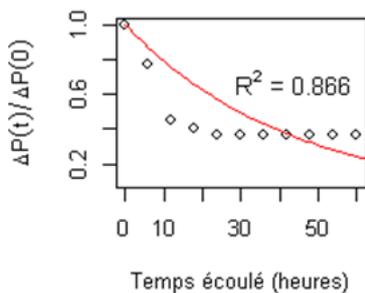




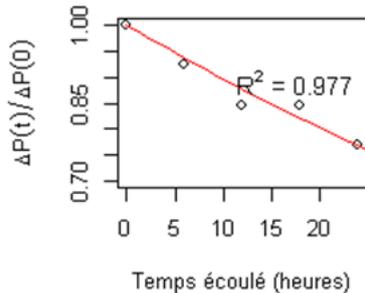
Cyclone Claudette (2003)**Cyclone Grace (2003)****Cyclone Larry (2003)****Cyclone Frances (2004)****Cyclone Gaston (2004)****Cyclone Ivan (2004)****Cyclone Jeanne (2004)****Cyclone Arlene (2005)****Cyclone Cindy (2005)****Cyclone Dennis (2005)****Cyclone Emily (2005)****Cyclone Katrina (2005)**



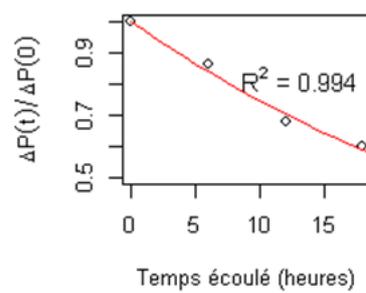
Cyclone Hermine (2010)



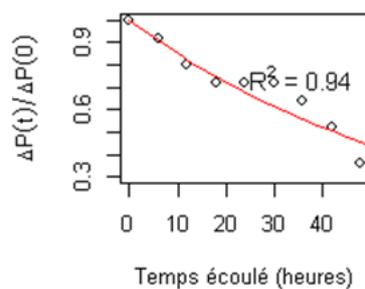
Cyclone Matthew (2010)



Cyclone Irene (2011)



Cyclone Lee (2011)



D Equations associées aux courbes de vulnérabilité

Nous donnons ci-dessous les équations retenues pour chacune des quatre courbes de vulnérabilité utilisées dans la troisième partie du mémoire et cela pour chaque classe de bâtiments (1, 2 et 3).

D.1 Classe 1

Structure des bâtiments résidentiels :

$$taux = \begin{cases} 0 & \text{si } v \leq 161 \\ 19,4884 - 0,636695 \cdot v - 8,37325 \cdot 10^{-3} \cdot v^2 - 5,67397 \cdot 10^{-5} \cdot v^3 \\ -2,085194 \cdot 10^{-7} \cdot v^4 - 3,930172 \cdot 10^{-10} \cdot v^5 + 2,97327 \cdot 10^{-13} \cdot v^6 & \text{si } 161 < v \leq 339 \\ 1 & \text{si } v \geq 339 \end{cases}$$

Contenu des bâtiments résidentiels :

$$taux = \begin{cases} 0 & \text{si } v \leq 192 \\ -2,733159 + 4,020344 \cdot 10^{-2} \cdot v - 1,974725 \cdot 10^{-4} \cdot v^2 \\ +3,240469 \cdot 10^{-7} \cdot v^3 & \text{si } 100 < v \leq 274,6252 \\ 1,267159 - 5,059481 \cdot 10^{-3} \cdot v & \text{si } 274,6252 < v < 344 \\ 1 & \text{si } v \geq 448 \end{cases}$$

Structure des bâtiments industriels/commerciaux/gouvernementaux :

$$taux = \begin{cases} 0 & \text{si } v \leq 112 \\ -0,9111455 + 2,281205 \cdot 10^{-2} \cdot v - 1,993365 \cdot 10^{-4} \cdot v^2 \\ +6,89741 \cdot 10^{-7} \cdot v^3 - 6,307864 \cdot 10^{-10} \cdot v^4 - 3,37137 \cdot 10^{-13} \cdot v^5 & \text{si } 112 < v \leq 352 \\ 1 & \text{si } v \geq 352 \end{cases}$$

Contenu des bâtiments industriels/commerciaux/gouvernementaux :

$$taux = \begin{cases} 0 & \text{si } v \leq 142 \\ 0,506905 - 1,220410 \cdot 10^{-2} \cdot v + 1,105903 \cdot 10^{-4} \cdot v^2 \\ -4,582755 \cdot 10^{-7} \cdot v^3 + 7,527075 \cdot 10^{-10} \cdot v^4 & \text{si } 142 < v \leq 331 \\ 1 & \text{si } v \geq 331 \end{cases}$$

D.2 Classe 2

Structure des bâtiments résidentiels :

$$taux = \begin{cases} 0 & \text{si } v \leq 144 \\ 11,2167 - 0,357435 \cdot v + 4,67397 \cdot 10^{-3} \cdot v^2 - 3,19013 \cdot 10^{-5} \cdot v^3 \\ +1,18836 \cdot 10^{-7} \cdot v^4 - 2,26953 \cdot 10^{-10} \cdot v^5 + 1,73449 \cdot 10^{-13} \cdot v^6 & \text{si } 144 < v \leq 343 \\ 1 & \text{si } v \geq 343 \end{cases}$$

Contenu des bâtiments résidentiels :

$$taux = \begin{cases} 0 & \text{si } v \leq 176 \\ -0,4608301 + 8,415864 \cdot 10^{-3} \cdot v - 5,089902 \cdot 10^{-5} \cdot v^2 \\ \quad + 1,020747 \cdot 10^{-7} \cdot v^3 & \text{si } 176 < v \leq 262 \\ -1,0973179 + 4,516427 \cdot 10^{-3} \cdot v & \text{si } 262 < v < 464 \\ 1 & \text{si } v \geq 464 \end{cases}$$

Structure des bâtiments industriels/commerciaux/gouvernementaux :

$$taux = \begin{cases} 0 & \text{si } v \leq 100 \\ -1,046995 + 2,704828 \cdot 10^{-2} \cdot v - 2,512016 \cdot 10^{-4} \cdot v^2 \\ + 9,99059 \cdot 10^{-7} \cdot v^3 - 1,507006 \cdot 10^{-10} \cdot v^4 + 6,01281 \cdot 10^{-13} \cdot v^5 & \text{si } 100 < v \leq 344 \\ 1 & \text{si } v \geq 344 \end{cases}$$

Contenu des bâtiments industriels/commerciaux/gouvernementaux :

$$taux = \begin{cases} 0 & \text{si } v \leq 152 \\ 0,9193589 - 2,143188 \cdot 10^{-2} \cdot v + 1,839008 \cdot 10^{-4} \cdot v^2 \\ \quad - 7,025593 \cdot 10^{-7} \cdot v^3 + 1,04344 \cdot 10^{-9} \cdot v^4 & \text{si } 142 < v \leq 329,7 \\ 1 & \text{si } v \geq 329,7 \end{cases}$$

D.3 Classe 3

Structure des bâtiments résidentiels :

$$taux = \begin{cases} 0 & \text{si } v \leq 96 \\ -4,808 \cdot 10^{-2} + 5,008 \cdot 10^{-4} \cdot v & \text{si } 96 < v \leq 148,116 \\ 31,7 - 0,9276 \cdot v + 1,114 \cdot 10^{-2} \cdot v^2 - 7,018 \cdot 10^{-5} \cdot v^3 \\ + 2,438 \cdot 10^{-7} \cdot v^4 - 4,412 \cdot 10^{-10} \cdot v^5 + 3,243 \cdot 10^{-13} \cdot v^6 & \text{si } 148,116 < v < 335 \\ 1 & \text{si } v \geq 335 \end{cases}$$

Contenu des bâtiments résidentiels :

$$taux = \begin{cases} 0 & \text{si } v \leq 106 \\ -0,2178 + 4,551 \cdot 10^{-3} \cdot v - 3,088 \cdot 10^{-5} \cdot v^2 + 6,952 \cdot 10^{-8} \cdot v^3 & \text{si } 106 < v < 391 \\ 1 & \text{si } v \geq 391 \end{cases}$$

Structure des bâtiments industriels/commerciaux/gouvernementaux :

$$taux = \begin{cases} 0 & \text{si } v \leq 100 \\ -3,357 \cdot 10^{-2} + 3,385 \cdot 10^{-4} \cdot v & \text{si } 100 < v \leq 152,361 \\ 29,53 - 0,8366 \cdot v + 9,708 \cdot 10^{-3} \cdot v^2 - 5,899 \cdot 10^{-5} \cdot v^3 \\ + 1,976 \cdot 10^{-7} \cdot v^4 - 3,450 \cdot 10^{-10} \cdot v^5 + 2,451 \cdot 10^{-13} \cdot v^6 & \text{si } 152,3613 < v < 344 \\ 1 & \text{si } v \geq 344 \end{cases}$$

Contenu des bâtiments industriels/commerciaux/gouvernementaux :

$$taux = \begin{cases} 0 & \text{si } v \leq 151 \\ -4,720 + 0,1678 \cdot v - 2,389 \cdot 10^{-3} \cdot v^2 + 1,743 \cdot 10^{-5} \cdot v^3 \\ - 6,890 \cdot 10^{-8} \cdot v^4 + 1,401 \cdot 10^{-10} \cdot v^5 - 1,141 \cdot 10^{-13} \cdot v^6 & \text{si } 151 < v < 344 \\ 1 & \text{si } v \geq 344 \end{cases}$$

E Ajustement d'une distribution sur les coûts bruts dans le cadre du modèle simplifié

Nous reprenons ici la modélisation des coûts bruts générés par une tempête. Nous avons montré dans la section 8.2 que les excès au delà du seuil $S_1 = 250$ M€ suivaient une loi de Pareto généralisée avec ξ négatif (c'est-à-dire une loi de Pareto de type II). Il reste alors à modéliser les coûts inférieurs à S_1 . Pour cela nous traçons à nouveau le *mean excess plot* de l'échantillon, puis nous déterminons le seuil S_2 à partir duquel ce dernier possède une tendance linéaire et nous ajustons une loi de Pareto généralisée sur les excès au delà de S_2 . Afin de tester l'adéquation des excès avec la loi choisie nous effectuons deux tests : l'un graphique (Q-Q plot) et l'autre statistique (Kolmogorov-Smirnov). Cette démarche a également été mise en œuvre pour déterminer S_3, S_4, S_5, S_6, S_7 et S_8 ainsi que les ajustements de Pareto associés au excès au delà de ces seuils. Notons que nous nous sommes arrêté au huitième seuil car celui-ci vaut 0.

Dans les éléments ci-dessous nous donnons successivement et pour $j = 2, \dots, 8$:

- le *mean excess plot* ayant permis de déterminer le seuil S_j
- le graphe quantile-quantile des excès contre la loi de Pareto ajustée sur les excès au delà du seuil S_j
- le test de Kolmogorov-Smirnov d'adéquation des excès

Détermination du seuil S_2 et ajustement des excès au delà de S_2 :

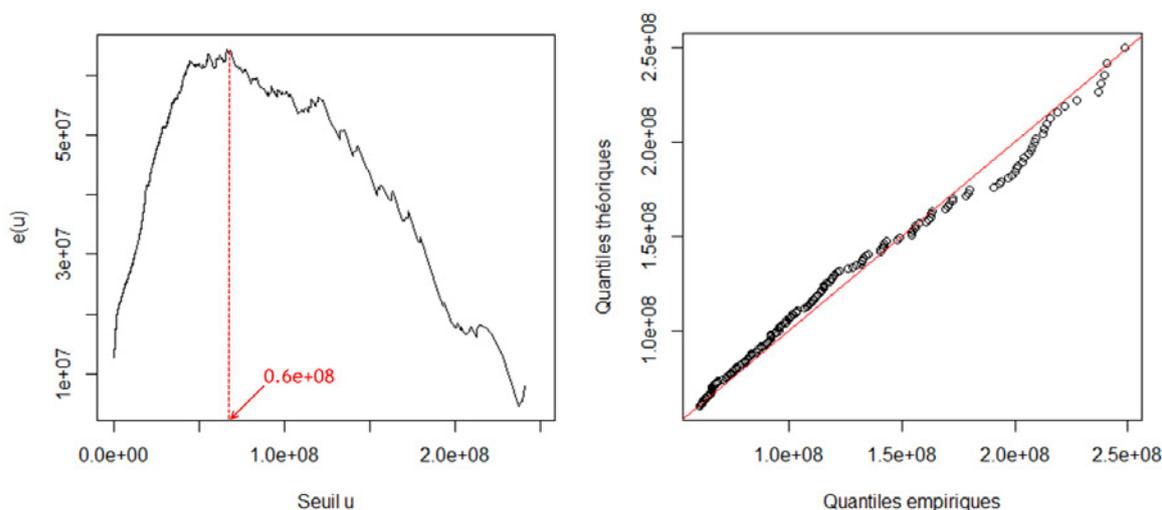


FIGURE 73 – *Mean excess plot* et Q-Q plot des excès au delà de 60 M€

Statistique	Seuil	p-valeur
0,0696	0,106	0,406

Détermination du seuil S_3 et ajustement des excès au delà de S_3 :

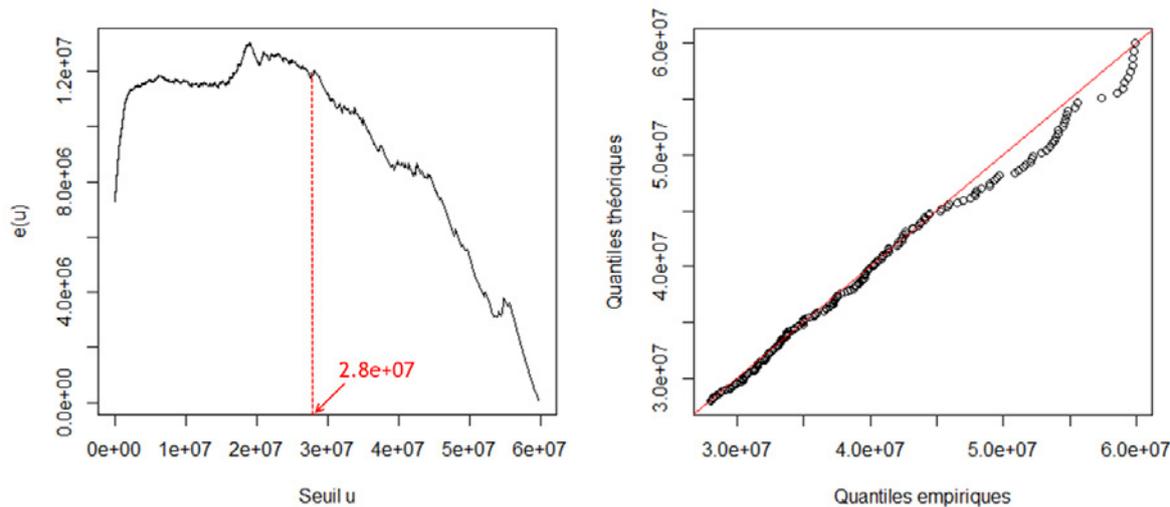


FIGURE 74 – Mean excess plot et Q-Q plot des excès au delà de 28 M€

Statistique	Seuil	p-valeur
0,0593	0,099	0,530

Détermination du seuil S_4 et ajustement des excès au delà de S_4 :

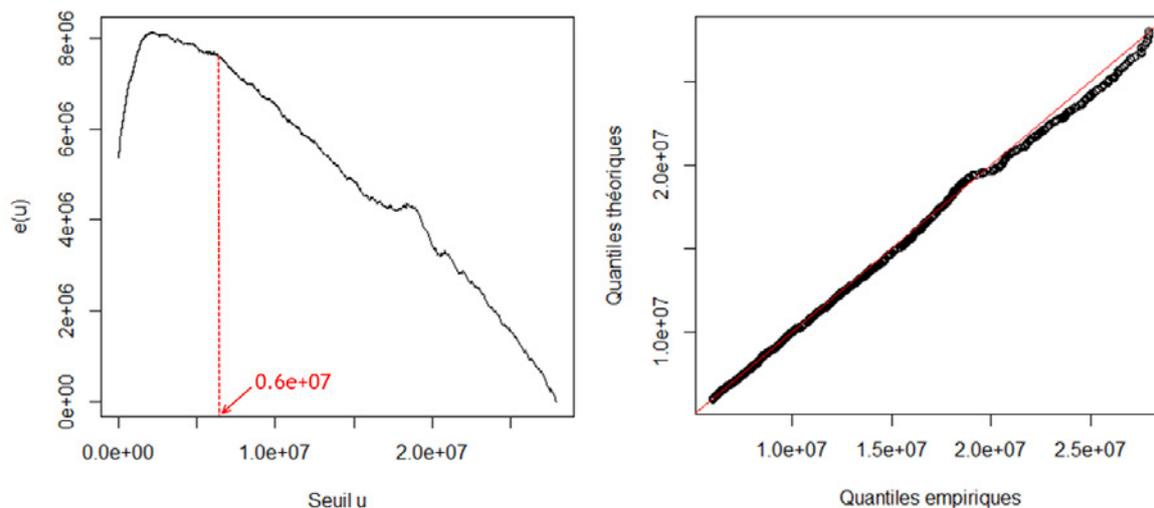


FIGURE 75 – Mean excess plot et Q-Q plot des excès au delà de 6 M€

Statistique	Seuil	p-valeur
0,0218	0,0429	0,728

Détermination du seuil S_5 et ajustement des excès au delà de S_5 :

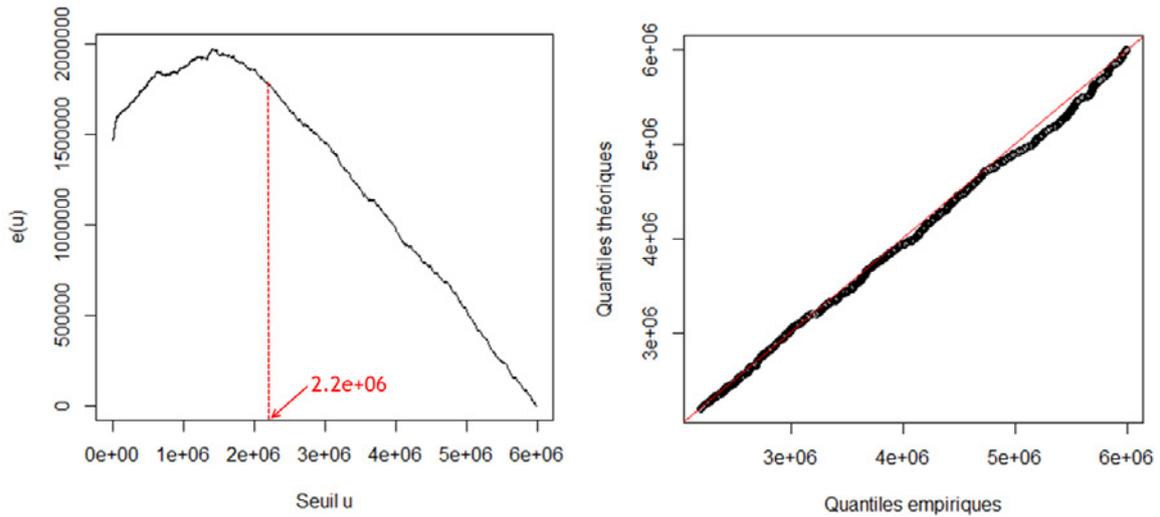


FIGURE 76 – *Mean excess plot* et *Q-Q plot* des excès au delà de 2,2 M€

Statistique	Seuil	p-valeur
0,0419	0,0591	0,315

Détermination du seuil S_6 et ajustement des excès au delà de S_6 :

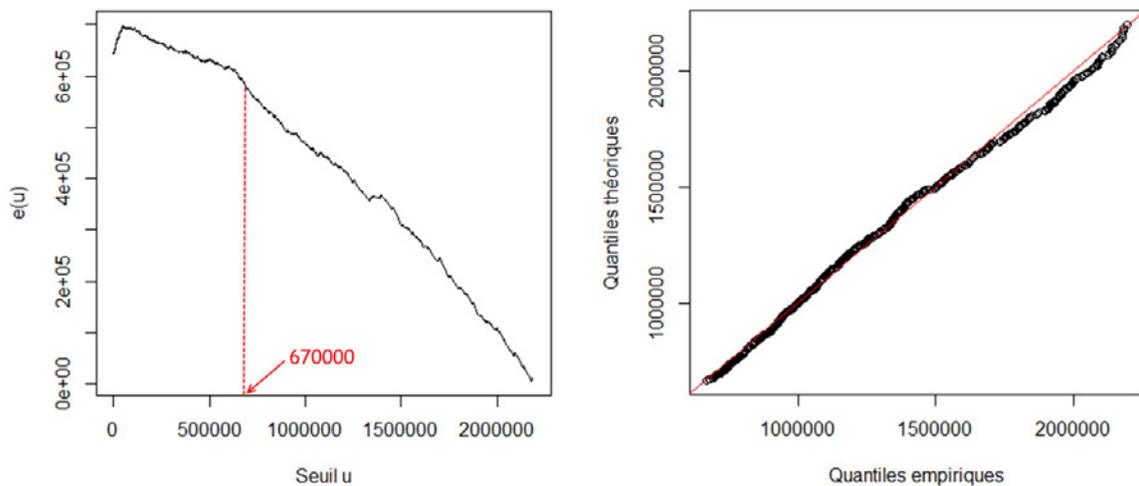


FIGURE 77 – *Mean excess plot* et *Q-Q plot* des excès au delà de 670 k€

Statistique	Seuil	p-valeur
0,0294	0,0547	0,664

Détermination du seuil S_7 et ajustement des excès au delà de S_7 :

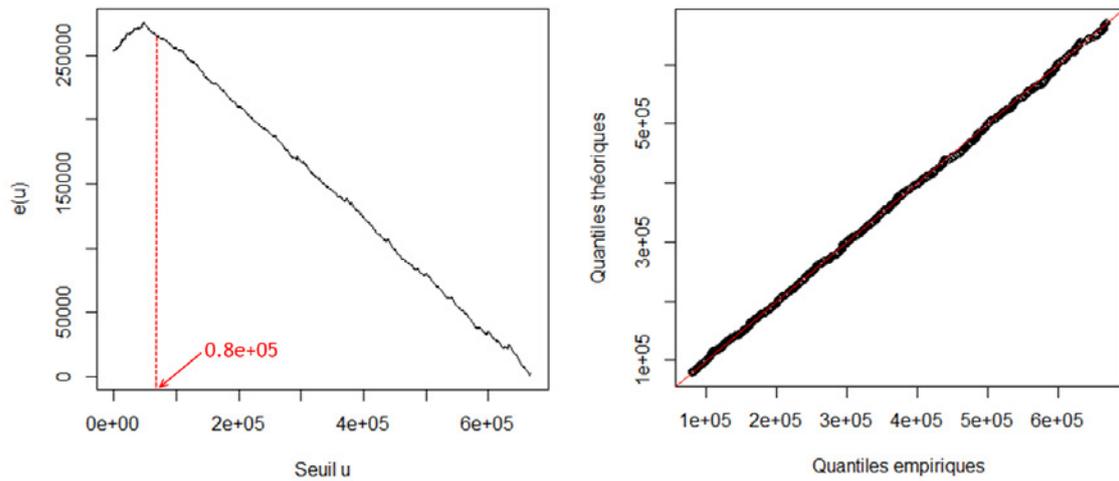


FIGURE 78 – *Mean excess plot* et *Q-Q plot* des excès au delà de 80 k€

Statistique	Seuil	p-valeur
0,0212	0,0512	0,912

Détermination du seuil S_8 et ajustement des excès au delà de S_8 :

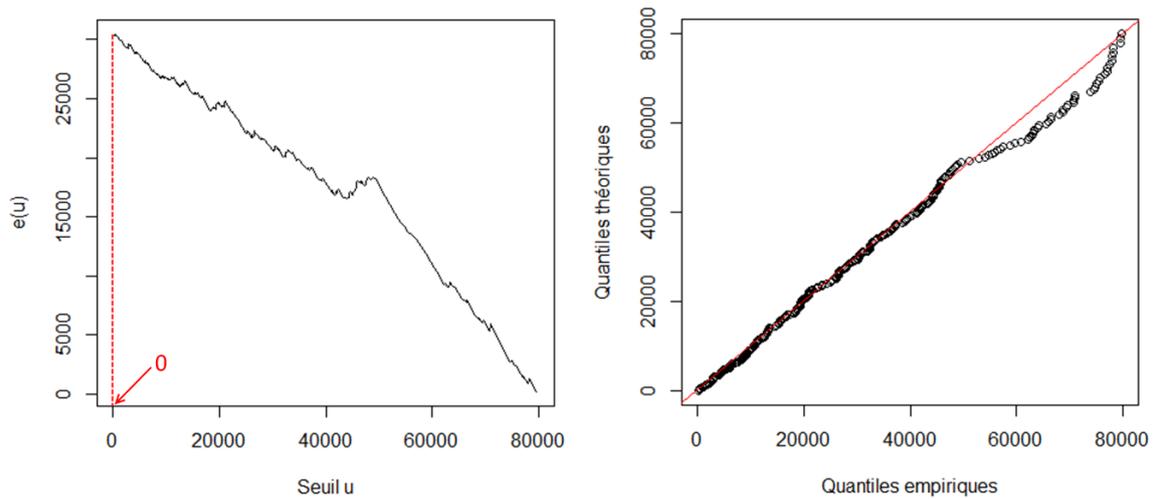


FIGURE 79 – *Mean excess plot* et *Q-Q plot* des excès au delà de 0 €

Statistique	Seuil	p-valeur
0,0469	0,0810	0,568

Bibliographie

- [1] APREF [2011] : *Réforme du régime des Catastrophes Naturelles*. Réassurance, pp 12.
- [2] CALVET L., GRISLAIN-LETRÉMY C. [2011] : *L'assurance habitation dans les départements d'Outre-mer : une faible souscription*. INSEE, Economie et Statistique numéro 447.
- [3] DARLING R.W.R. [1991] : *Estimating probabilities of hurricane wind speeds using a large-scale empirical model*. Journal of Climate, volume 4, pp 1035–1046.
- [4] DVORAK V.F. [1975] : *Tropical cyclone intensity analysis and forecasting from satellite imagery*. Monthly weather review, volume 103, pp 420-430.
- [5] EIOPA [2010] : *QIS5 Technical Specifications*. European Commission, Internal Market and Services DG.
- [6] EMMANUEL K. A. [1988] : *The maximum intensity of hurricanes*. Journal of Atmospheric Sciences, volume 45, pp 1143–1155.
- [7] EMMANUEL K. A. [1988] : *Toward a general theory of hurricanes*. American Scientist, volume 76, pp 371-379.
- [8] EARLE A. [1997] : *Hazards of the British Virgin Islands*. Final report of the Hazard and Risk Assessment Project (HRAP).
- [9] HENKA P., RUCK B. [2008] : *A damage model for the assessment of storm damage to buildings*. Engineering Structures, volume 30, pp 3603-3609.
- [10] HOLLAND G. [2010] : *A revised hurricane pressure–wind model*. Monthly weather review, volume 136, pp 3432-3445.
- [11] JELESNIANSKI C.P., CHEN J., SHAFFER W.A. [1992] : *SLOSH : Sea, Lake, and Overland Surges from Hurricanes*. NOAA Technical Report NWS 48.
- [12] KAPLAN J., DEMARIA M. [1995] : *A simple empirical model for predicting the decay of tropical cyclone winds after landfall*. Journal of Applied Meteorology., volume 34, pp 2499-2512.
- [13] NEUMANN C.J. [1987] : *The National Hurricane Center risk analysis program (HURISK)*. Journal of structural engineering, pp 1222-1237.
- [14] SCHOTT T., LANDSEA C., HAFELE G., LORENS J., TAYLOR A., THURM H., WARD B., WILLIS M., ZALESKI W. [2012] : *The Saffir-Simpson Hurricane Wind Scale*. The Saffir-Simpson Team, National Oceanic and Atmospheric Administration (NOAA).
- [15] VICKERY P.J., SKERLJ P.F., TWISDALE L.A. [2000] : *Simulation of hurricane risk in the US using empirical track model*. NOAA Technical Memorandum NWS NHC 38.
- [16] WANG Y., ROSOWSKY D.V. [2011] : *Joint distribution model for prediction of hurricane wind speed and size*. Structural Safety, volume 35, pp 40-51

Table des figures

1	Implantation de GFA Caraïbes	5
2	Architectures des risques de GFA Caraïbes	7
3	Le bassin Atlantique Nord	8
4	Les différents stades d'une perturbation tropicale	10
5	Image satellitaire du cyclone Katrina (2005) dans le golfe du Mexique	11
6	Mois de formation des tempêtes tropicales (1982-2011)	11
7	Forces engendrant la rotation d'un cyclone	12
8	Structure d'un cyclone tropical à maturité	13
9	Trajectoires habituelles des cyclones dans le bassin Atlantique Nord	14
10	Evolution de la pression et de la vitesse des vents dans un cyclone	15
11	Echelle de SAFFIR-SIMPSON	16
12	Illustration des dommages par catégorie	17
13	Allure de la base HURDAT	19
14	Modélisation du risque tempête	20
15	Points d'impact à 6 heures d'intervalle (1982-2011)	23
16	Algorithme de simulation d'une saison complète de trajectoires	24
17	Nombre d'occurrence à horizon un an et histogramme associé (1982 – 2011)	25
18	Histogrammes théorique et empirique du nombre d'occurrences	27
19	Quadrillage du bassin Atlantique selon une grille $1^\circ \times 1^\circ$	28
20	Positions initiales des tempêtes tropicales (1982-2011)	29
21	Echantillon de 50 trajectoires empiriques (à gauche) et simulées (à droite)	31
22	Boîtes à moustache des nombres de points formant les trajectoires	32
23	Densités des impacts empiriques (à gauche) et simulés (à droite)	33
24	Tempêtes simulées traversant les îles assurées	34
25	Evolution de la pression au centre du cyclone Jeanne (2004)	36
26	Algorithme de modélisation de la pression centrale le long d'une trajectoire	37
27	Filtre océan-terre utilisé	37
28	Intensité relative, pression centrale et température mensuelle moyenne à la surface de l'océan pour Andrew (1992)	42
29	Température moyenne à la surface de l'océan Atlantique Nord (août 2011)	43
30	Résolution choisie pour la détermination des coefficients de la régression	45
31	Nombre de données disponibles pour effectuer les régressions	45
32	Détermination de la pression centrale le long des trajectoires dans l'océan	46
33	Résultat des tests pour les régressions en résolution $5^\circ \times 5^\circ$	46
34	Résultat des tests pour les régressions en résolution $2,5^\circ \times 2,5^\circ$	47
35	Pressions centrales moyennes empiriques (à gauche) et modélisées (à droite)	48
36	Histogramme des pressions centrales	49
37	Q-Q plot entre pressions empiriques et pressions modélisées	50

38	Décroissance du déficit de pression au centre de Jeanne (2004)	51
39	Nombre d'observations à disposition	52
40	Histogramme et boîte à moustache des constantes de désintégration	52
41	Q-Q plot de la constante de désintégration	54
42	Pressions centrales moyennes empiriques (à gauche) et modélisées (à droite) . . .	54
43	Vitesse maximale des vents en fonction de la pression centrale (HURDAT)	55
44	Q-Q plot des erreurs	56
45	Représentation des modèles liant v_{max} et P	57
46	Représentation du logarithme de R_{max} en fonction de diverses variables	58
47	Représentation de R_{max}^i en fonction du R_{max}^{i-1}	59
48	Histogramme des R_{max} initiaux	60
49	Exemple d'un gradient de vitesse (Katrina, 2005)	61
50	Localisation des îles en résolution $1^\circ \times 1^\circ$	62
51	Exemple d'une trajectoire traversant une île	62
52	Histogrammes des vitesses de vent simulées	63
53	Sommes assurées par GFA Caraïbes	66
54	Classes proposées en fonction de la qualité des bâtiments	67
55	Fenêtres munies de jalousies métalliques	67
56	Courbes de vulnérabilité retenues	68
57	Courbes de vulnérabilité ajustées	69
58	Résultats bruts issus des simulations	71
59	Illustration du traité quote-part	74
60	Illustration du traité en excédent de sinistre	75
61	Fonctionnement du programme de réassurance de GFA Caraïbes	77
62	Evolution du SCR net en fonction du nombre de simulations	78
63	Calcul du SCR net par la formule standard	80
64	Histogramme fréquentiel du nombre de tempêtes impactant GFA Caraïbes	83
65	Adéquation graphique	84
66	Coûts non nuls et histogramme fréquentiel associé	85
67	<i>Mean excess plot</i> des coûts bruts	86
68	Q-Q plot des excès au delà de 250 M€	86
69	Algorithme de simulation du coût brut annuel	88
70	Q-Q plot des deux échantillons	88
71	Evolution des quantiles à 99,5%	89
72	Distance entre deux points d'une sphère	97
73	<i>Mean excess plot</i> et Q-Q plot des excès au delà de 60 M€	107
74	<i>Mean excess plot</i> et Q-Q plot des excès au delà de 28 M€	108
75	<i>Mean excess plot</i> et Q-Q plot des excès au delà de 6 M€	108
76	<i>Mean excess plot</i> et Q-Q plot des excès au delà de 2,2 M€	109
77	<i>Mean excess plot</i> et Q-Q plot des excès au delà de 670 k€	109
78	<i>Mean excess plot</i> et Q-Q plot des excès au delà de 80 k€	110
79	<i>Mean excess plot</i> et Q-Q plot des excès au delà de 0 €	110