

**Mémoire présenté devant l'ENSAE Paris
pour l'obtention du diplôme de la filière Actuariat
et l'admission à l'Institut des Actuares**

le 07/03/2025

Par : **Lou NICOLAS**

Titre : **Estimation des pertes associées au passage d'un
cyclone à La Réunion**

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

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

Membres présents du jury de la filière

Antoine HERANVAL

Entreprise : Generali IARD

Nom :

Signature :

*Membres présents du jury de l'Institut
des Actuares*

Directeur de mémoire en entreprise :

Nom : Malikane NIDAL MAHMOUD

Signature :

Malikane Nidal Mahmoud

✓ Certified by  youSign

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

Signature du responsable entreprise

Malikane Nidal Mahmoud

✓ Certified by  youSign

Secrétariat :

Signature du candidat

Bibliothèque :

Lou NICOLAS

✓ Certified by  youSign

Table des matières

Résumé	3
Abstract	4
Remerciements	5
Introduction	6
Partie 1 : Cadre de l'étude et contexte	8
Chapitre 1 : Présentation générale du phénomène cyclonique	8
1.1. Propriétés et mécanismes des cyclones	8
1.2. Typologie des sinistres engendrée par un cyclone	11
1.3. Historique et tendances d'évolution des cyclones à La Réunion	12
Chapitre 2 : Enjeux et adaptations de l'assurance aux risques cycloniques dans les Outre-Mer	15
2.1. Assurabilité dans les territoires ultramarins français	15
2.2. Régime CatNat dans les Outre-Mer françaises : cadre et spécificités régionales	16
2.3. Garantie « Tempêtes Ouragans Cyclones » : une solution assurantielle lorsque la garantie CatNat n'est pas mobilisée	19
Chapitre 3 : Contexte opérationnel et objectifs du mémoire	20
3.1. Impact des cyclones sur Prudence Créole : Enjeux et exposition aux risques	20
3.2. Mécanisme de réassurance entre Prudence Créole et Generali IARD : Structure et implications	21
3.3. Objectif de l'étude : Identification des événements réassurés et estimation de leurs coûts pour Prudence Créole et Generali IARD.	24
Partie 2 : Analyse des données et méthodologie de la modélisation du risque cyclonique à La Réunion	26
Chapitre 1 : Une modélisation jusqu'alors peu répandue	26
1.1. Rareté des modèles spécifiques à La Réunion	26
1.2. Des modèles marché se focalisant sur les effets du vent malgré l'importance des infiltrations à La Réunion	27
1.3. Méthodologie actuelle pour estimer les coûts des cyclones chez Prudence Créole	28
Chapitre 2 : Les données disponibles et leurs pré-traitements	30
2.1. Données de sinistralité historique	30
2.2. Données d'exposition	36
2.3. Données météorologique	42
Chapitre 3 : Présentation du modèle catastrophe simplifié	49
3.1. Les enjeux ayant conduit à l'élaboration d'un modèle CAT simplifié	49
3.2. Fonctionnement et application d'un modèle « CAT » simplifié	50

3.3. Limites et avantages d'un modèle CAT simplifié _____	52
Partie 3 : L'estimation des pertes associées à la survenance des cyclones à La Réunion _____	54
Chapitre 1 : Identification des cyclones nécessitant un suivi par le service de réassurance _____	54
1.1. Distribution théorique des excès et détection de seuil _____	55
1.2. Détermination de seuils météorologiques intenses _____	59
1.3. Critères de réassurance pour les cyclones : étude basée sur les sommes assurées impactées par des événements extrêmes _____	68
Chapitre 2 : Estimation de la charge d'un cyclone réassuré _____	72
2.1. Retraitements préalables à la modélisation _____	72
2.2. Modèle linéaire généralisé _____	81
2.3. Estimation du coût d'un cyclone après son passage proche ou sur La Réunion _____	88
Chapitre 3 : Limites et fiabilisation du modèle _____	96
3.1. Limites du modèle _____	96
3.2. Fiabilisation et perspectives d'amélioration _____	97
Conclusion _____	99
Bibliographie _____	102
Table des figures _____	106
Table des tableaux _____	108
Glossaire des acronymes _____	109
Note de synthèse _____	110
Executive summary _____	117
Annexes _____	124

Ce mémoire reprend certains concepts déjà rédigés par l'auteur dans une série d'articles publiés sur LinkedIn, tout en les approfondissant dans un cadre académique.

Par ailleurs, afin de respecter les obligations de confidentialité, certaines données ont été volontairement modifiées tout en préservant la cohérence et la pertinence des analyses. Les données modifiées sont présentées en gris et en italique, assurant une distinction tout en respectant la rigueur méthodologique et en protégeant les informations sensibles.

Résumé

À La Réunion, les cyclones représentent un risque majeur pour les assureurs, causant des dommages importants chaque année. Prudence Créole, filiale de Generali France et leader du marché local, voit ses résultats affectés lors des années marquées par des cyclones intenses. Pour limiter cet impact, Prudence Créole se réassure auprès de Generali France, qui supporte à son tour une partie des coûts à travers ses acceptations de réassurance.

Dans ce contexte, il est capital pour Generali France et Prudence Créole d'estimer rapidement les coûts de tels événements, d'autant plus que la modélisation des cyclones à La Réunion est très limitée sur le marché.

Ce mémoire propose une méthodologie en deux étapes pour modéliser les impacts financiers du risque cyclonique. Dans un premier temps, la théorie des valeurs extrêmes permet de déterminer si un cyclone dépassera un seuil de coût proche du seuil de cession de réassurance (75% du seuil). Ensuite, un modèle linéaire généralisé (GLM) estime la charge de sinistres associée à ces événements.

En adoptant une approche pragmatique et adaptée aux spécificités locales, cette démarche contribue à mieux anticiper les coûts des événements majeurs et à renforcer la résilience de Prudence Créole et de Generali France face aux risques cycloniques. Ce travail met ainsi en lumière le rôle clé d'une modélisation efficace pour faciliter les échanges avec les réassureurs et optimiser la gestion des risques.

Mots-clés : Cyclones, Catastrophes naturelles, La Réunion, Non-vie, Théorie des Valeurs Extrêmes (TVE), Modèle Linéaire Généralisé (GLM), Réassurance non-vie.

Abstract

On Reunion Island, cyclones represent a major risk for insurers, causing significant damage every year. Prudence Créole, a subsidiary of Generali France and the local market leader, sees its results affected in years marked by intense cyclones. To limit this impact, Prudence Créole reinsures with Generali France, which in turn bears part of the costs through its reinsurance acceptances.

In this context, it is vital for Generali France and Prudence Créole to rapidly estimate the costs of such events, especially as cyclone modeling in La Réunion is very limited on the market.

This paper proposes a two-stage methodology for modeling the financial impact of cyclonic risk. First, extreme value theory is used to determine whether a cyclone will exceed a cost threshold close to the reinsurance cession threshold (75% of the threshold). Next, a generalized linear model (GLM) estimates the claims burden associated with these events.

By adopting a pragmatic approach tailored to local conditions, this process helps us to better anticipate the costs of major events and strengthen the resilience of Prudence Créole and Generali France to cyclonic risks. This work highlights the key role of effective modeling in facilitating exchanges with reinsurers and optimizing risk management.

Keywords: Cyclones, Natural disasters, Reunion Island, Non-life, Extreme Value Theory (EVT), Generalized Linear Model (GLM), Non-life reinsurance.

Remerciements

Je suis profondément reconnaissante envers ma tutrice, Malikane NIDAL MAHMOUD, pour sa confiance et son accompagnement, qui ont largement dépassé le cadre de ce mémoire. Collaborer avec elle a été une véritable chance.

Un immense merci à mon tuteur de l'ENSAE, Antoine HERANVAL, pour sa réactivité, son accompagnement attentif et ses conseils précieux. Je tiens également à saluer l'ensemble du corps enseignant de l'ENSAE.

Je n'oublie pas l'université Paris Dauphine, qui a marqué mes premières années académiques et m'a inculqué la rigueur et la persévérance nécessaires à la poursuite de mes ambitions. Je remercie tout particulièrement les intervenants du Master 2 Assurance et Gestion des Risques, qui ont su transmettre leur expertise avec passion. Une mention spéciale à Kristiano BEJKO, qui m'a encouragée à croire en moi et à poursuivre mon parcours en actuariat à l'ENSAE.

Je tiens également à exprimer ma profonde gratitude à mes parents, qui m'ont soutenue tout au long de mes études, ainsi qu'à mes deux sœurs.

Je remercie également mes amis, et en particulier Zak, pour leur présence et leurs encouragements, qui ont été d'un grand soutien durant la rédaction de ce mémoire.

Mes remerciements vont aussi à mes collègues de Generali, Anas, Laura, Nathan, David et Yassine pour leur soutien, parfois discret mais toujours précieux, tout au long de ce projet.

À toutes ces personnes, je tiens à dire un grand merci. Chacun, à sa manière, a contribué à ce projet et à mon parcours, et je vous en suis profondément reconnaissante.

Introduction

Les cyclones, phénomènes météorologiques parmi les plus violents, constituent une menace majeure pour l'île de La Réunion, département français situé dans l'océan Indien. Avec des pluies torrentielles, des vents destructeurs et des houles cycloniques, ces événements naturels engendrent des dégâts humains, matériels et économiques considérables. Dans un contexte de réchauffement climatique, l'intensité et les impacts des cyclones ne cessent de croître, renforçant la vulnérabilité des territoires exposés. La Réunion illustre ces défis avec une intensité particulière en raison de son isolement géographique et de ses caractéristiques topographiques uniques.

Les sinistres liés aux cyclones peuvent atteindre des montants considérables, comme l'a montré le cyclone Belal en 2024, avec plus de 42 000 sinistres déclarés pour un coût total estimé à 100 millions d'euros pour l'ensemble du marché d'après France Assureur. Il est donc essentiel pour les assureurs locaux, dont Prudence Créole, filiale de Generali France et leader historique de l'assurance à La Réunion, d'estimer avec précision les coûts d'un événement cyclonique. Ces charges financières significatives obligent les assureurs à s'appuyer sur des mécanismes de réassurance pour limiter les impacts financiers et garantir la pérennité de l'activité.

Mais les réassureurs, souvent éloignés des réalités locales, tendent à percevoir les territoires d'Outre-mer comme homogènes et particulièrement risqués. Regroupant La Réunion avec des territoires aux profils bien différents et plus exposés, tels que les Antilles, ces analyses généralisées ne reflètent pas les spécificités des cyclones réunionnais ni l'efficacité des dispositifs de gestion des sinistres de Prudence Créole.

De plus, l'efficacité des modèles disponibles pour l'océan Indien est largement inférieure à celle des modèles développés pour les régions atlantiques et pacifiques. Dans les Caraïbes et le Pacifique, les modélisations bénéficient d'une infrastructure robuste soutenue par des centres de recherche en météorologie et modélisation climatique situés aux États-Unis, au Japon et en Australie. Ces pays investissent massivement dans la collecte de données et le développement d'outils de modélisation, permettant une calibration fine des modèles et une adaptation aux spécificités locales. En revanche, La Réunion, en raison de son isolement géographique et de l'absence de soutien de la part de grandes nations investissant dans la recherche sur les risques naturels de la région, dispose de modèles moins performants et moins nombreux. Les difficultés liées à la disponibilité des données historiques exacerbent cette situation ainsi que la topographie complexe de l'île. De plus, le coût élevé du développement de nouveaux modèles décourage les investissements, laissant les assureurs avec des outils moins efficaces conçus pour d'autres régions du globe, aux dynamiques climatiques et topographiques bien différentes.

Dans ce contexte, ce mémoire se concentre sur la problématique actuarielle suivante : comment développer un modèle simple et opérationnel permettant de déterminer rapidement si un cyclone nécessite une couverture de réassurance et d'estimer avec précision les coûts associés, en tenant compte des spécificités de La Réunion ? Cette problématique guide l'ensemble des travaux réalisés, de la sélection des données à la construction d'un modèle statistique robuste.

Ce mémoire se distingue par une démarche méthodologique adaptée aux particularités de La Réunion. Contrairement aux modèles classiques, généralement développés pour des régions comme les Antilles ou l'Atlantique Nord, ce travail intègre des éléments propres à l'île, notamment l'importance des précipitations intenses et des infiltrations dans la sinistralité.

L'intérêt du modèle construit repose également sur l'équilibre entre simplicité opérationnelle et pertinence locale. Plutôt que de s'appuyer sur des modèles complexes et peu exploitables en pratique, ce travail propose une approche pragmatique, conçue pour répondre rapidement aux besoins des différentes parties prenantes. En s'appuyant sur des données historiques de sinistres, des caractéristiques météorologiques et des spécificités d'exposition du portefeuille de Prudence Créole, ce modèle vise à fournir des résultats directement utilisables par Prudence Créole et Generali France, en tant que réassureur de sa filiale.

Afin de répondre à la problématique posée, le mémoire s'organise en trois parties. La première partie établit le contexte de l'étude en présentant les mécanismes des cyclones, leurs impacts spécifiques à La Réunion ainsi que les enjeux et adaptations des secteurs de l'assurance et de la réassurance face à ces risques.

La deuxième partie se concentre sur les données disponibles et la méthodologie de modélisation, en mettant en lumière les traitements effectués pour préparer les données et développer un modèle catastrophe simplifié adapté aux spécificités locales et aux contraintes opérationnelles de simplicité.

Enfin, la troisième partie explore les concepts théoriques nécessaires pour évaluer les impacts des cyclones et les applique afin de déterminer si un cyclone sera réassuré ou non, avant d'estimer précisément les coûts associés à ces événements.

Partie 1 : Cadre de l'étude et contexte

Chapitre 1 : Présentation générale du phénomène cyclonique

1.1. Propriétés et mécanismes des cyclones

Un cyclone, terme d'origine grecque signifiant "cercle", est une perturbation atmosphérique tourbillonnaire de diamètre réduit, généralement autour de 500 km. Ce phénomène engendre des vents violents et se déplace rapidement, couvrant une distance de 700 à 1000 km par jour.

Les cyclones se présentent sous forme de spirale avec des bandes nuageuses tournant autour d'un œil central calme mesurant entre 32 et 64 km de diamètre où les vents ne dépassent généralement pas 24 km/h. Autour de l'œil se trouve le mur de nuages, une zone de cumulonimbus atteignant 12 à 15 km d'altitude, où les vents peuvent atteindre jusqu'à 300 km/h, accompagnés de pluies torrentielles causant inondations et glissements de terrain. À environ 50 à 100 km de l'œil, des bandes nuageuses spiralées, pouvant s'étendre jusqu'à 500 km, apportent de fortes pluies et des vents puissants. Elles jouent un rôle majeur dans la structure du cyclone en maintenant sa puissance et sa rotation. De plus, les vents violents du cyclone peuvent générer une houle cyclonique, avec des vagues pouvant atteindre 20 mètres, augmentant le risque d'inondations côtières et de dégâts importants.

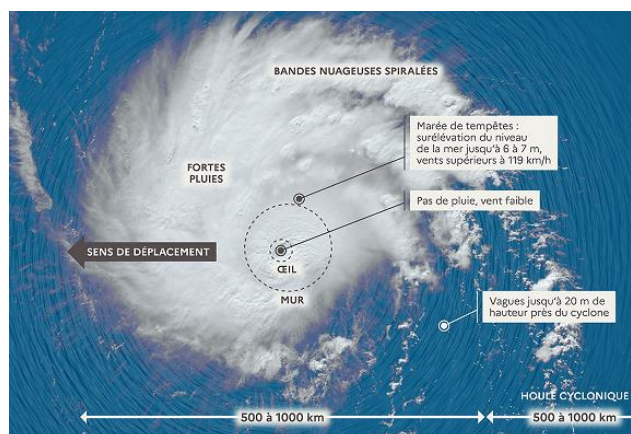


Figure 1 : Structure d'un cyclone. Source : Météo France.

Il est à noter que l'appellation du terme « cyclone » varie en fonction de la situation géographique : ouragan dans l'Atlantique et le nord-est du Pacifique, cyclone dans l'océan Indien et le sud du Pacifique, et typhon dans le nord-ouest du Pacifique. Etant donné que ce mémoire a pour but de modéliser les dégâts causés par les cyclones à La Réunion pour Prudence Créole, l'entité de Generali dans l'océan indien, seul le terme "cyclone" sera employé par la suite.

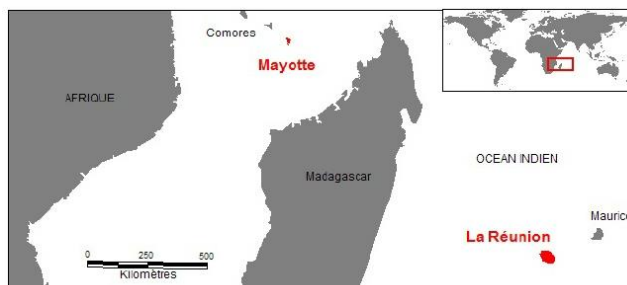


Figure 2 : Carte des Appellations Régionales des Cyclones. Carte réalisée à l'aide du logiciel Canva.

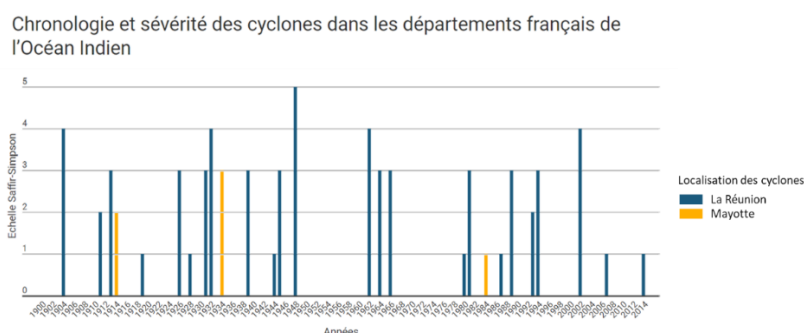
Encadré 1 – Territorialité du mémoire : La Réunion

Prudence Créole est implantée à La Réunion et à Mayotte mais seul le territoire de La Réunion sera modélisé. Voici les raisons principales.

Tout d'abord en étant plus près de l'équateur que La Réunion, Mayotte est moins affectée par les trajectoires cycloniques typiques qui se dirigent vers le sud-ouest de l'océan Indien. Mayotte est ainsi moins exposée que La Réunion au risque de vents violents.



De plus, Madagascar joue un rôle de barrière naturelle pour Mayotte. Lorsqu'un cyclone approche depuis l'est ou le sud-est, les reliefs montagneux et les terres malgaches réduisent l'intensité ou dévient la trajectoire des cyclones en perturbant leur structure et leur vitesse. Une exception notable reste le cyclone Chido de 2024, un événement rare car il n'a pas suivi la trajectoire habituelle. Avant lui, Mayotte n'avait été impactée que par trois cyclones depuis 1900, aucun n'atteignant la même intensité.



Le rapport de l'inspection générale des finances de 2020 sur le *phénomène de non-assurance dans les départements et collectivités d'outre-mer* classe ainsi le risque cyclonique à Mayotte comme modéré tandis qu'il serait fort à La Réunion.

Par ailleurs, les données historiques de sinistralité pour Mayotte sont inexistantes dans la base utilisée pour ce mémoire, qui commence en 2002. Le dernier cyclone significatif enregistré sur l'île remonte à 1984. Même si les deux territoires sont proches géographiquement, leurs conditions économiques diffèrent largement. Par exemple, le taux de souscription à une assurance habitation est estimé à 6 % à Mayotte, contre 68 % à La Réunion (rapport de l'Inspection générale des finances, 2020).

Enfin, d'après les données de collecte du portefeuille dommages aux biens de Prudence Créole vision juin 2024, Mayotte représente une part négligeable du portefeuille dommages aux biens de Prudence Créole tant en nombre d'assurés qu'en montant d'engagement.

Outre leur différence de nomenclature, les cyclones varient également en fonction des dates de saisons cycloniques, qui diffèrent selon les régions. Ces saisons correspondent aux périodes de l'année où les conditions météorologiques sont particulièrement propices à la formation de cyclones : des eaux chaudes, un faible cisaillement vertical du vent et une humidité atmosphérique élevée. Dans l'hémisphère nord, la saison cyclonique de l'Atlantique Nord s'étend du 1er juin au 30 novembre, avec un pic d'activité de la fin août à septembre. Dans l'hémisphère sud, elle s'étend du 15 novembre au 30 avril, avec un pic d'activité de janvier à mars. En moyenne, un cyclone s'étend de 5 à 7 jours.

Une fois formé, tant que le phénomène a accès à de l'eau chaude, il s'amplifie. Dans l'océan Atlantique et l'océan Pacifique, l'intensité des ouragans est graduée en cinq catégories en fonction de la vitesse des vents. Ce sont les fameuses catégories 1 à 5 de l'échelle de Saffir-Simpson. Cependant, il est à noter que, même si l'échelle Saffir-Simpson est régulièrement utilisée quelle que soit la zone par abus de langage, dans les faits, la classification des cyclones dans l'océan Indien est différente.

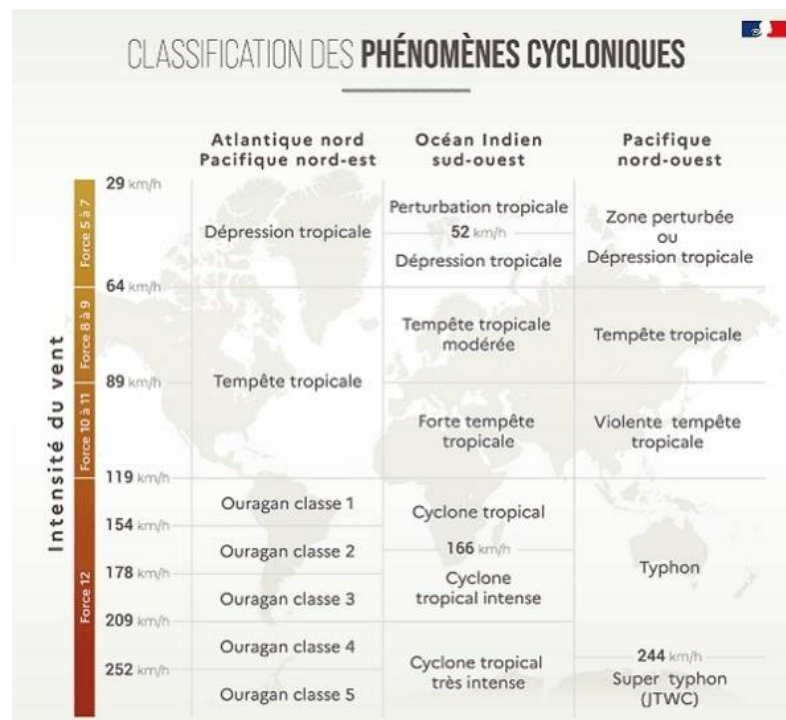


Figure 5 : Classification des phénomènes cycloniques en fonction de la zone géographique. Source : Météo France.

Ainsi, à La Réunion, lorsque les vents dépassent les 119 km/h, les phénomènes sont classés en trois catégories : cyclones tropicaux (jusqu'à 166 km/h), cyclones tropicaux intenses (jusqu'à 209 km/h), et cyclones tropicaux très intenses, avec des vents au-delà de 209 km/h. En dessous de 119 km/h, le terme cyclone est remplacé par perturbation, dépression ou tempête tropicale.

Après s'être intensifié, un cyclone peut également commencer à se dissiper lorsque certaines conditions météorologiques favorables cessent d'être réunies. Par exemple, lorsqu'un cyclone interagit avec des terres ou traverse des eaux plus froides, il perd l'accès à la chaleur et à l'humidité nécessaires à sa survie, ce qui désorganise sa structure. Ce désordre entraîne une diminution de l'intensité des vents et des précipitations. De plus, un cisaillement vertical accru du vent peut perturber l'alignement vertical de la tempête, accélérant ainsi son affaiblissement. Cette phase de déclin est critique pour évaluer les dommages finaux et les impacts économiques, car bien que l'intensité diminue, les effets résiduels tels que les inondations et les glissements de terrain peuvent continuer à causer des dégâts considérables.

1.2. Typologie des sinistres engendrée par un cyclone

Le cyclone est donc un événement météorologique complexe qui combine plusieurs actions destructrices.

Les **vents** des cyclones tropicaux causent des destructions considérables en raison de leur force et des changements brutaux de direction et d'intensité. Ces fluctuations entraînent des effets de vibrations et de percussions, accentuant ainsi la violence du cyclone et la gravité des dégâts. Les vents les plus forts soufflent près du centre du cyclone et diminuent rapidement en s'éloignant. Cependant, des vents violents peuvent également se produire dans les bandes nuageuses extérieures. En outre, lorsqu'un cyclone touche un continent, le relief terrestre et l'environnement des bâtiments modifient les vents. Le frottement avec le sol ralentit les vents moyens, mais en même temps, il génère des turbulences qui provoquent des variations soudaines de la vitesse du vent, ce qui intensifie les rafales. Par exemple, à La Réunion, où le relief est important, les rafales peuvent atteindre deux à trois fois la vitesse des vents moyens. Finalement, les vents violents peuvent transporter ou déplacer des objets et matériaux divers, les transformant en projectiles potentiellement très dangereux (panneaux, tuiles, branches d'arbres...). Par exemple, lors du cyclone de 1948, à Hell-Bourg située sur les hauteurs de La Réunion, les rafales étaient si violentes qu'elles ont arraché la statue en bronze de Carlo Sarrabezolles, « L'Âme de la France » de son piédestal, la projetant au sol où elle est restée pendant vingt ans.

Les cyclones tropicaux génèrent également des **pluies torrentielles** pouvant souvent entraîner des conséquences aussi dévastatrices que celles du vent, telles que des inondations, des infiltrations, des glissements de terrain, des coulées de boue et le débordement de torrents ou de ravines. La quantité de précipitations dépend de la vitesse de déplacement du cyclone (plus il se déplace lentement, plus les pluies peuvent durer) ainsi que de sa taille et de la structure nuageuse, souvent asymétrique. Lorsqu'un cyclone atteint une masse continentale, sa partie centrale se désintègre rapidement, réduisant l'intensité des vents et des pluies. Cependant, les bandes nuageuses extérieures persistent et provoquent des pluies intenses et durables à l'intérieur des terres. De plus, les vents peuvent rencontrer des obstacles tels que des montagnes, ce qui force l'air à monter. Ce soulèvement orographique¹ intensifie les précipitations, car l'air ascendant se refroidit, se condense et forme des nuages de pluie. À La Réunion, ce phénomène est amplifié par les hauts sommets, contribuant aux records mondiaux de précipitations mesurées sur plusieurs jours. Tous les records de précipitations pour des périodes de 12 heures à 15 jours sont d'ailleurs détenus par l'île et ils sont tous liés à des passages de tempêtes ou cyclones tropicaux.

Durée	Localités	Date	Hauteur (en mm)
1 mn	Unionville Md (152 m) USA	04/07/1956	31.2
15 mn	Plum Point Jamaïque	12/05/1916	198.1
1 heure (42 min.)	Holt (Montana) (263 m) USA	22/06/1947	304.8
12 heures	Foc Foc (2290 m) Réunion (Denise)	Janvier 1966	1144
24 heures	Foc Foc (2290 m) Réunion (Denise)	Janvier 1966	1825
48 heures	Aurère (940 m) Réunion (dép. non nommé)	Avril 1958	2467
3 jours	Commerson (2310 m) Réunion (Gamède)	Février 2007	3929
4 jours	Commerson (2310 m) Réunion (Gamède)	Février 2007	4869
5 jours	Commerson (2310 m) Réunion (Gamède)	Février 2007	4979
6 jours	Commerson (2310 m) Réunion (Gamède)	Février 2007	5075
7 jours	Commerson (2310 m) Réunion (Gamède)	Février 2007	5400
8 jours	Commerson (2310 m) Réunion (Gamède)	Février 2007	5510
9 jours	Commerson (2310 m) Réunion (Gamède)	Février 2007	5512
10 jours	Commerson (2310 m) Réunion (Hyacinthe)	Janvier 1980	5678
15 jours	Commerson (2310 m) Réunion (Hyacinthe)	Janvier 1980	6083

Tableau 1 : Records mondiaux de précipitations. Source : Meteo.fr.

¹ Tout phénomène lié à la présence et aux effets des montagnes

Ces pluies intenses sont susceptibles de provoquer des inondations², des éboulis, des coulées de boue ou encore des glissements de terrain³. Ces phénomènes sont d'autant plus fréquents si les sols des territoires sont déjà saturés en eau. A titre d'exemple, à la suite du passage de la tempête tropicale Diwa en mars 2006 à La Réunion, un glissement de terrain est survenu 15 jours après. Ce type de risques géologiques dus aux fortes pluies est très difficiles à prévoir et ajoute une complexité supplémentaire à la modélisation des destructions post-cycloniques.

Enfin, les cyclones ont aussi un impact significatif sur l'état de l'océan. Leur passage engendre une **houle cyclonique**, générée par les vents puissants associés au cyclone. Cette houle se déplace plus rapidement que le cyclone lui-même et peut engendrer une surélévation anormale du niveau de la mer. Ce phénomène est observable de plusieurs minutes à plusieurs heures et sa hauteur peut varier de quelques centimètres à plusieurs mètres, atteignant parfois entre quatre et huit mètres. Plus elle se rapproche des côtes, plus son intensité augmente, ce qui peut entraîner des inondations par submersion marine. Par ailleurs, la houle cyclonique peut accentuer le phénomène d'érosion côtière. En février 2007, lors du passage du cyclone Gamède, la houle cyclonique a causé des dégâts considérables dans la baie de Saint-Paul : le cimetière marin a été touché, la plage s'est rétrécie de manière significative et le front de mer a subi une érosion importante.

Les cyclones représentent donc une menace significative pour les zones côtières et insulaires, provoquant des dégâts considérables par le biais de vents violents, de précipitations intenses et de houles cycloniques. Ils laissent des traces profondes sur les biens et les populations de La Réunion, avec des événements marquants qui jalonnent l'histoire de l'île.

1.3. Historique et tendances d'évolution des cyclones à La Réunion

Depuis les premières observations météorologiques, La Réunion, située dans une zone particulièrement propice à la formation des cyclones, en a connu d'intensités variables. À travers les décennies, des cyclones emblématiques, comme Jenny en 1962, Hyacinthe en 1980, Firinga en 1989, Dina en 2002 ou encore Gamède en 2007, ont laissé des traces indélébiles, tant sur le paysage que dans la mémoire collective. Chacun de ces événements climatiques a occasionné des destructions massives des infrastructures, des pertes humaines et des bouleversements économiques, rappelant la vulnérabilité de l'île.

Le cyclone Jenny, en 1962, a frappé l'île avec des vents dépassant les 200 km/h. Les dégâts les plus importants ont été enregistrés dans le Nord-Est de l'île, où 92 % des maisons ont été touchées à Bras-Panon, causant 31 décès. Si un tel événement se produisait aujourd'hui, les dommages seraient estimés à environ 83 millions d'euros, avec 90 % des destructions causées par le vent et 10 % par les submersions⁴. En 1980, le cyclone Hyacinthe a apporté des pluies torrentielles pendant près de deux semaines, accumulant plus de 6 000 mm de pluie dans certaines régions. Ces précipitations exceptionnelles ont provoqué des inondations massives et des glissements de terrain, endommageant les infrastructures et isolant des communautés entières. Le cyclone Firinga, en 1989, a engendré des vents violents qui ont entraîné l'effondrement de ponts, la destruction de nombreux bâtiments et des pertes économiques importantes notamment dans le secteur agricole, laissant plus de 6 000 familles sinistrées⁵. Avec des rafales atteignant 300 km/h et une durée exceptionnelle de vents forts sur plus de

² Une inondation est caractérisée par la submersion d'une zone où habituellement l'eau n'est pas présente.

³ Les mouvements de terrain désignent les déplacements relatifs de volumes de sol ou de sous-sol, qui peuvent se produire à des vitesses variées. Cette catégorie englobe les éboulements, les coulées de boue, les glissements de terrain, ainsi que les chutes de pierres et de blocs.

⁴ Source cyclone Jenny : CCR

⁵ Source Firinga : France Info et INA

28 heures⁶, le cyclone Dina, en 2002, a causé des dommages majeurs aux infrastructures électriques et de communication, tout en provoquant des inondations significatives. Le cyclone Gamède de 2007 a causé des précipitations intenses avec plus de 4 979 mm de pluie en cinq jours sur le Piton de la Fournaise⁷, soit huit fois la quantité annuelle de pluie à Paris.

Ces exemples historiques soulignent que la dangerosité d'un cyclone ne réside pas uniquement dans la force de ses vents, mais aussi dans la quantité de précipitations qu'il apporte et les effets indirects qu'il peut engendrer. Ainsi, lors de la modélisation des cyclones pour estimer les dommages potentiels, il est essentiel de ne pas se limiter à l'analyse des vents violents. Une approche globale qui intègre à la fois les risques liés aux vents et aux précipitations permet d'obtenir une estimation plus précise des impacts possibles sur les infrastructures, l'économie et la population de La Réunion.

En revanche, bien que La Réunion ait été frappée par plusieurs cyclones dévastateurs, il est important de noter que tous les cyclones de la région n'ont pas la même ampleur ni les mêmes effets destructeurs. La zone du sud-ouest de l'océan Indien est particulièrement active en termes de cyclogenèse, avec une moyenne de 9 à 12 tempêtes tropicales par an, mais beaucoup d'entre elles finissent par toucher Madagascar, Maurice, ou simplement se dissiper en mer. Cette variabilité signifie que, bien que la zone soit exposée, l'île échappe souvent aux impacts directs des cyclones les plus violents. Ainsi, pour illustrer cette variabilité, la carte suivante présente les trajectoires de tous les cyclones ayant évolué près de La Réunion depuis 2002. Les cyclones ayant impacté directement Prudence Créole sont mis en évidence en rouge, tandis que les autres, qui ont suivi des trajectoires différentes ou se sont dissipés⁸, apparaissent en gris.

Trajectoires des cyclones dans le bassin indien sud depuis 2002

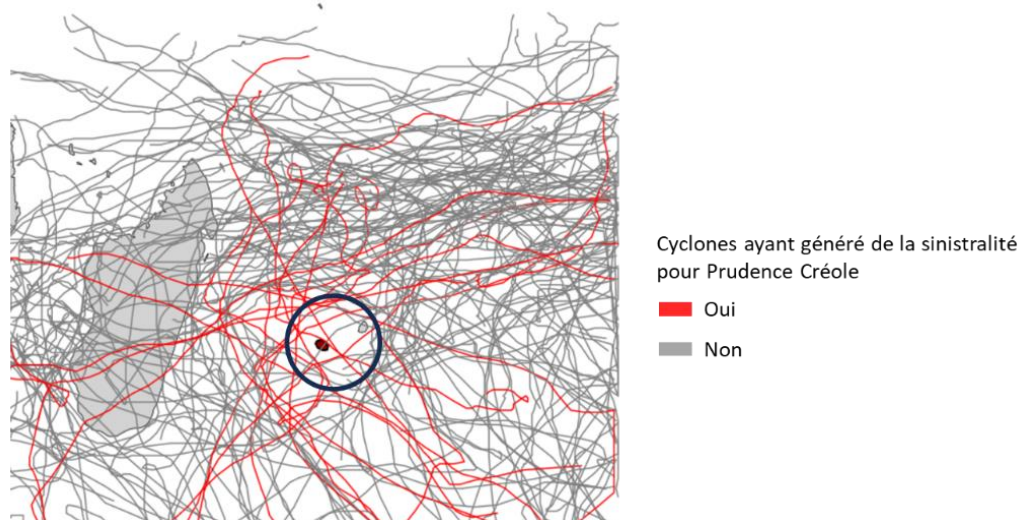


Figure 6 : Trajectoires des cyclones dans le bassin indien sud depuis 2002.

Mais les recherches scientifiques récentes⁹ suggèrent que si le nombre total de cyclones pourrait ne pas augmenter, leur intensité, elle, risque de s'intensifier. En effet, les simulations climatiques pour le XXI^e siècle du sixième rapport du GIEC montrent que les cyclones deviendront plus puissants sous l'effet du réchauffement climatique. Le changement climatique réchauffe la température des mers et des océans à la surface du globe et modifie la circulation atmosphérique globale. Bien qu'une température de surface de l'océan plus élevée ne facilite pas nécessairement la formation de cyclones,

⁶ Source : CCR

⁷ Source : Meteo.fr

⁸ Une trajectoire grise semble traverser La Réunion sans avoir impacté Prudence Créole. Il s'agit du cyclone Hondo de 2008, qui lorsqu'il a atteint La Réunion était à l'état de dépression tropicale avec des vents entre 50 et 65 km/h.

⁹ Emanuel, K. (2018). "100 years of progress in tropical cyclone research." *Meteorological Monographs*, 59, 15.1–15.68 et Tropical Cyclones and Global Climate Change : A Post-IPCC Assessment (1998)

un cyclone déjà formé puisera beaucoup plus d'énergie pour se renforcer dans une atmosphère plus humide au-dessus des océans réchauffés. En effet, plus l'atmosphère est chaude, plus elle peut contenir d'humidité. Cette humidité supplémentaire intensifie les pluies cycloniques, ce qui, à son tour, renforce encore le cyclone.

Ainsi, une augmentation de 5 % de l'intensité moyenne des cyclones et une hausse de 14 % de la proportion de cyclones très intenses sont prévues. De plus, les précipitations associées aux cyclones devraient augmenter d'environ 15 % pour un réchauffement global de 2 °C¹⁰. L'élévation du niveau des océans accroît également le risque d'inondations côtières provoquées par les cyclones, en raison de la montée des eaux et de l'intensification des tempêtes. Enfin, l'augmentation de la teneur en vapeur d'eau dans l'atmosphère due au réchauffement climatique contribue à des précipitations plus abondantes lors des cyclones, exacerbant ainsi les risques d'inondations à l'intérieur des terres.

Ces dernières années illustrent l'impact du changement climatique, La Réunion ayant été impactée par deux cyclones majeurs en un laps de temps réduit : Batsirai en 2022 et Belal en 2024. Batsirai a engendré des inondations localisées et des dégâts aux cultures, tandis que Belal a provoqué des rafales atteignant 217 km/h et des inondations sur les hauteurs de l'île, avec des coupures d'électricité massives. Selon les données de France Assureur, ce dernier cyclone aurait coûté près de 100 millions d'euros en dommages assurés.

Face à cette intensification des cyclones sous l'effet du changement climatique, les mécanismes d'assurance à La Réunion sont essentiels pour protéger les biens et les habitants contre ces risques accrus.

¹⁰Données : rapport *Intergovernmental Panel on Climate Change* (IPCC)

Chapitre 2 : Enjeux et adaptations de l'assurance aux risques cycloniques dans les Outre-Mer

2.1. Assurabilité dans les territoires ultramarins français

Contrairement à l'Hexagone, où le recours à l'assurance est largement répandu, les populations des territoires d'Outre-mer sont moins assurées, la moitié des habitants des territoires ultramarins bénéficient d'une protection similaire. Dans le rapport sur « *Le phénomène de non-assurance dans les départements et collectivités d'Outre-mer* », l'Inspection générale des finances détaille plusieurs facteurs expliquant cette plus faible couverture. Bien que les revenus y soient généralement inférieurs à ceux de l'Hexagone, ce n'est pas seulement une question de coût.

Historiquement, la discontinuité territoriale a retardé l'intégration des départements d'Outre-mer dans le système d'assurance national. De ce fait, les sociétés ultramarines étaient souvent résilientes et autosuffisantes face aux sinistres. L'isolement géographique des îles les contraignait à développer des mécanismes locaux de solidarité et d'entraide, comme les « case-à-vent¹¹ » pour résister aux cyclones, et des systèmes d'épargne communautaire comme les « sousou¹² » ou « tontine¹³ ».

De plus, une proportion importante de biens immobiliers est considérée comme non-assurable en raison de constructions irrégulières, d'incertitudes liées aux droits de propriété, ou de leur exposition à des risques élevés, comme les cyclones fréquents. Il existe également une méconnaissance de certains risques et de leurs conséquences parmi les populations locales, avec un faible niveau de sensibilisation aux pertes potentielles dues aux catastrophes naturelles.

Par ailleurs, les représentations négatives à l'égard de l'assurance, notamment les perceptions de complexité administrative et d'indemnisations tardives ou insuffisantes, contribuent aussi à la faible souscription. L'offre limitée d'assureurs dans ces régions, en raison des marchés étroits et de la perception de risques élevés, complique davantage l'accès à des couvertures d'assurance adaptées.

Enfin, les territoires ultramarins sont souvent représentés comme un ensemble homogène, malgré la mosaïque de réalités qui les compose. D'un point de vue climatique, la Guyane, Mayotte ou certaines îles du Pacifique sont peu affectées par les cyclones. À l'inverse, Saint-Martin, avec ses terres littorales basses et son urbanisation dense en zones à risque, se tient en première ligne face aux ouragans dévastateurs. Un cyclone de l'ampleur d'Irma n'aurait pas les mêmes conséquences sur une île comme La Réunion, dont la topographie et l'urbanisme diffèrent sensiblement. Il n'est donc pas judicieux de tirer des enseignements d'Irma pour les appliquer uniformément à tous les territoires ultramarins, tout comme il serait incohérent de le faire pour des régions en Hexagone régulièrement sinistrées.

D'un point de vue assurantiel, il existe également une forte disparité au sein des territoires ultramarins. Le taux de souscription s'élève à 68 % à La Réunion, 62 % en Martinique, 59 % en Guadeloupe, 49 % en Guyane et seulement 6 % à Mayotte. L'évolution de la pénétration de l'assurance au cours des 25 dernières années montre que La Réunion se trouve à un carrefour entre deux mondes : celui de la résilience communautaire traditionnelle et celui de l'assurance formelle.

Face à cette diversité, les défis d'assurabilité en Outre-Mer rendent le régime d'indemnisation des catastrophes naturelles, le régime « CatNat », particulièrement essentiel. D'après le rapport « Adapter

¹¹ Maison traditionnelle conçue pour résister aux cyclones et aux vents forts dans les régions insulaires.

¹² Système de collecte d'épargne dans lequel les membres d'un groupe contribuent régulièrement une somme d'argent à un pot commun. À chaque cycle, un membre différent reçoit la totalité du pot, offrant ainsi un moyen de financement sans passer par des institutions formelles.

¹³ Mécanisme similaire au « sousou », où chaque membre cotise régulièrement. Le pot est attribué périodiquement à l'un des membres, souvent par rotation ou tirage au sort.

le système assurantiel français face à l'évolution des risques climatiques » de décembre 2023 dit rapport « Langrenay », ce régime est un outil indispensable pour compenser les insuffisances du marché de l'assurance dans ces territoires fortement exposés aux risques naturels. Il permet une mutualisation nationale des risques, ce qui est crucial face à la fréquence et à l'intensité des catastrophes dans les régions ultramarines.

2.2. Régime CatNat dans les Outre-Mer françaises : cadre et spécificités régionales

Depuis sa création en 1982, le régime d'indemnisation des catastrophes naturelles dites « CatNat » s'est imposé comme une réponse aux extrêmes de la nature. Selon l'article L125-1 du Code des assurances, la garantie CatNat prend en charge les « *dommages matériels directs non assurables ayant eu pour cause déterminante l'intensité anormale d'un agent naturel, lorsque les mesures habituelles à prendre pour prévenir ces dommages n'ont pu empêcher leur survenance ou n'ont pu être prises* ». En raison de leur caractère exceptionnel et de leur potentiel destructeur élevé, les cyclones entrent donc dans la définition des périls non assurables et sont couverts par le régime CatNat.

Encadré 2 – L'inclusion des vents cycloniques dans le régime CatNat

L'intégration des Outre-mer au sein du régime CatNat rappelle que ces territoires évoluent souvent dans un temps décalé par rapport à l'Hexagone. Le législateur prévoyait en effet de répondre aux spécificités de ces territoires par une loi ultérieure. C'est après le passage de l'ouragan Hugo sur les Antilles en 1989 et du cyclone Firinga à La Réunion en 1989 que les territoires ultramarins ont été intégrés au régime, soit 8 années après l'Hexagone. Cependant, les dommages causés par les vents cycloniques restaient exclus du régime et continuaient d'être pris en charge par les contrats d'assurance classiques.

Ce n'est qu'avec la loi n° 2000-1207 du 13 décembre 2000 d'orientation pour l'outre-mer que les dommages liés aux vents cycloniques ont été intégrés au régime CatNat. Cette loi a modifié le Code des assurances pour inclure les effets du vent dus à un événement cyclonique « pour lequel les vents maximaux de surface enregistrés ou estimés sur la zone sinistrée atteignent ou dépassent 145 km/h en moyenne sur dix minutes ou 215 km/h en rafales ».

L'intégration des vents cycloniques dans le régime CatNat en 2000 a marqué une étape importante pour la protection des territoires d'outre-mer. Elle a reconnu l'ampleur des risques spécifiques auxquels ces régions sont exposées et a apporté une solution aux problèmes d'assurabilité rencontrés par les assureurs, tout en renforçant la sécurité économique des habitants.

Les ménages et les entreprises résidant sur le territoire français¹⁴ sont couverts contre les effets des catastrophes naturelles grâce à une extension obligatoire de leur contrat d'assurance. Pour financer le régime, chaque assuré verse une prime additionnelle de 12 % sur ses garanties dommages du contrat de base (par exemple, sur un contrat multirisques habitation) et de 6 % sur les garanties vol et incendie pour les automobiles. Ces taux sont les mêmes pour tous, quel que soit le degré d'exposition face aux CatNat. Ainsi, les habitants de la commune de Saint-Benoît (sud de l'île) et ceux de la commune de Sainte-Marie (nord de l'île) paient le même taux de surprime CatNat, alors même que leurs ratios sinistres à primes¹⁵ (S/P) sont respectivement inférieur à 10% et supérieur à 200%. Cette uniformité

¹⁴ Hors Pays d'Outre-mer : Nouvelle Calédonie, Polynésie Française, TAAF.

¹⁵ Le ratio sinistres/primes (S/P) mesure la proportion des sinistres payés par rapport aux primes collectées par un assureur. Il se calcule en divisant le montant total des sinistres indemnisés par le montant total des primes perçues. Un ratio inférieur à 100 % indique que les primes collectées suffisent à couvrir les sinistres, tandis qu'un ratio supérieur à 100 % reflète une insuffisance des primes par rapport aux sinistres, ce qui peut entraîner des pertes pour l'assureur.

des taux contribue à mutualiser les risques à l'échelle nationale. Ces taux passeront respectivement à 20 % et 9 % à partir du 1^{er} janvier 2025¹⁶.

Ratio sinistres à primes (S/P) tous périls confondus de 1995 à 2021 par commune

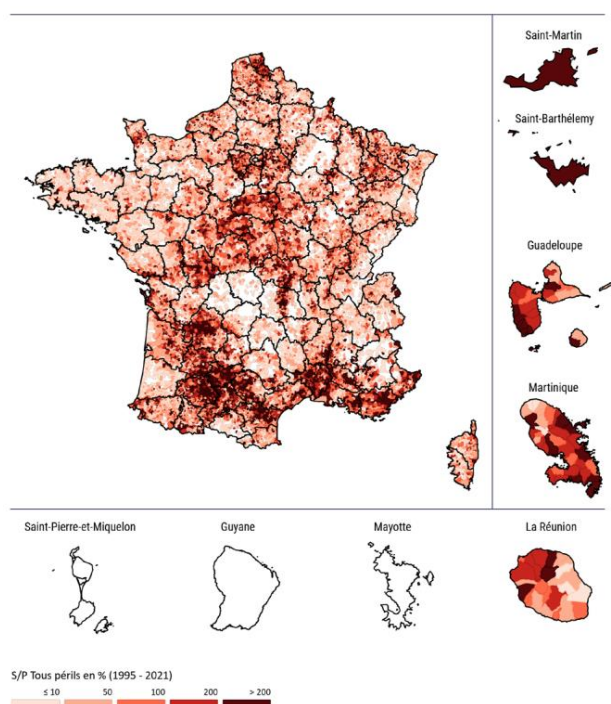


Figure 7 - Ratio sinistres à primes (S/P) tous périls confondus de 1995 à 2021. Source : CCR

En contrepartie de cette surprime, lorsqu'un événement exceptionnel survient, le maire dépose une demande de reconnaissance auprès du préfet, qui la transmet à une Commission interministérielle composée de représentants des ministères concernés, notamment du ministère de l'Intérieur et du Ministère de la Transition Écologique. Cette Commission a pour rôle de déterminer si l'événement présente une intensité anormale, en s'appuyant sur des analyses scientifiques réalisées entre autres par Météo France. Si c'est le cas, les indemnités de cet événement relèvent du régime d'indemnisation des catastrophes naturelles.

Pour les vents cycloniques, une intensité anormale est caractérisée par des seuils précis. Un cyclone est considéré comme exceptionnel si les vents enregistrés ou estimés sur la zone sinistrée atteignent ou dépassent **145 km/h en moyenne sur dix minutes ou 215 km/h en rafales**.

Pour les inondations consécutives, l'intensité anormale est évaluée en fonction de critères tels que le dépassement des niveaux de précipitations exceptionnels, le débordement de cours d'eau ou la submersion marine dépassant les seuils habituels. La classification d'une commune en CatNat est officialisée par un arrêté interministériel, dont la publication au Journal officiel est une condition préalable indispensable à l'indemnisation.

Une fois la catastrophe naturelle reconnue, la compagnie d'assurance missionne alors un(e) expert(e) pour qu'il/elle évalue les dommages et le montant d'indemnisation. L'assureur indemnise ensuite les sinistrés et la Caisse Centrale de Réassurance (CCR), la société de réassurance détenue à

¹⁶ Arrêté du 22 décembre 2023 modifiant le taux de la prime ou cotisation additionnelle relative à la garantie « catastrophe naturelle » aux contrats d'assurance mentionné à l'article L. 125-2 du code des assurances

100 % par l'État, prend en charge environ la moitié des coûts des catastrophes naturelles par le biais d'une couverture de réassurance publique. Ce montant peut augmenter significativement en fonction du caractère exceptionnel de l'événement. Si la CCR n'a pas les fonds nécessaires, l'Etat, par le biais d'une garantie illimitée accordée à la CCR, intervient en dernier ressort.

L'État garantit ainsi la solvabilité du système, intervenant lorsque la sinistralité excède 90 % des réserves de la CCR en une année. Une telle situation ne s'est produite qu'une seule fois depuis la création du régime CatNat, suite aux tempêtes Lothar et Martin en 1999, qui ont conduit à la reconnaissance de l'état de catastrophe naturelle pour 26 805 communes, touchées par des inondations, coulées de boue ou mouvements de terrain. Selon la Cour des comptes, le seuil d'intervention de l'État est actuellement fixé à 2,6 milliards d'euros, équivalant au coût d'une catastrophe telle qu'une crue centennale de la Seine ou un ouragan de type Irma. L'illustration ci-dessous détaille le processus complet de reconnaissance et d'indemnisation des catastrophes naturelles.

Couverture des risques naturels en France dans le cadre du régime Cat Nat institué par la loi n°82-600 du 13 juillet 1982

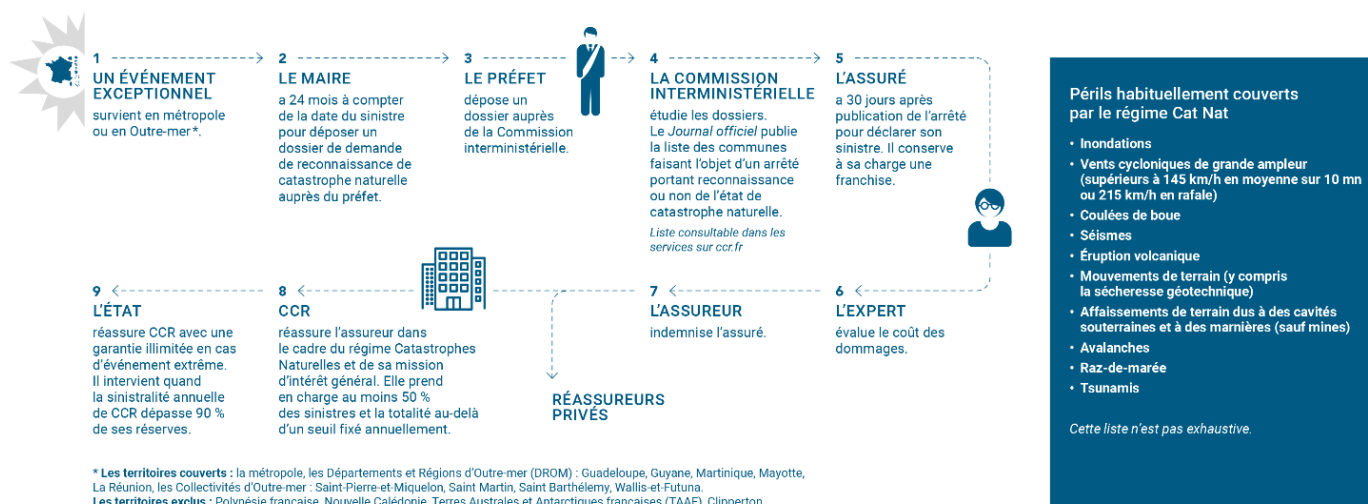


Figure 8 : Fonctionnement du régime CAT-NAT. Source : CCR

À La Réunion, le régime CatNat est particulièrement important en raison de la fréquence et de l'intensité des phénomènes naturels, notamment les cyclones tropicaux. Selon la CCR, sur la période 1995-2015, les dommages assurés liés aux catastrophes naturelles ont atteint en moyenne 9 millions d'euros par an, dont 6 millions d'euros par an pour les seuls vents cycloniques. Deux communes, Saint-Denis et Saint-Paul, concentrent à elles seules près de 60 % de la sinistralité historique sur la période 1995-2019. Cette concentration s'explique par la densité de population et l'exposition particulière de ces zones aux aléas naturels. Sur la période 1982-2019, l'île a bénéficié de 162 arrêtés CatNat, soit une moyenne d'environ 4 à 5 arrêtés par an. Les périls les plus fréquemment reconnus sont les inondations et les vents cycloniques. Ces données mettent en évidence l'importance du régime CatNat pour la protection économique de La Réunion face aux risques naturels majeurs. Elles soulignent également la nécessité d'une gestion du risque adaptée, intégrant la prévention et l'indemnisation pour renforcer la résilience de l'île.

Cependant, tous les sinistres ne sont pas toujours reconnus comme des catastrophes naturelles, notamment lorsque les vents ne dépassent pas le seuil d'intensité anormalement élevée. Pour couvrir ces types d'événements, la garantie Tempête, Ouragan et Cyclone (TOC) intervient.

2.3. Garantie « Tempêtes Ouragans Cyclones » : une solution assurantielle lorsque la garantie CatNat n'est pas mobilisée

La garantie « Tempêtes Ouragans Cyclones » dite « TOC » est incluse dans les contrats d'assurance multirisques de dommages aux biens en France, à condition que les biens soient également couverts contre l'incendie. Cette garantie vise à couvrir les **dégâts causés par des forces naturelles intenses, bien que non exceptionnelles**. Elle couvre principalement les dommages causés par le vent et les précipitations, y compris l'eau qui pénètre dans les bâtiments endommagés. Cette protection s'applique aux biens situés à l'intérieur de structures fermées, comme les bâtiments et les équipements, ainsi qu'aux pertes financières associées. Toutefois, les aménagements extérieurs et le mobilier extérieur ne sont pas couverts.

En 2024, le cyclone Belal a frappé La Réunion et s'est distingué par des précipitations exceptionnellement élevées : en quatre jours, 1 367 mm ont été enregistrés à Commerson et 1 019 mm à la Plaine des Palmistes. À titre de comparaison, lors de la tempête Alex en octobre 2020 dans les Alpes-Maritimes, les précipitations ont atteint des niveaux historiques, avec plus de 500 mm de précipitations en moins de 24 heures dans certaines zones. En plus des pluies torrentielles, les vents ont été particulièrement violents, notamment dans la moitié nord-est de l'île, où le mur de l'œil est passé. Les rafales ont dépassé les 150 km/h sur la plupart des points de mesure, atteignant même 216 km/h au Piton Maïdo, un point en hauteur particulièrement exposé. En revanche, ce dernier est le seul endroit à avoir enregistré des rafales aussi extrêmes dépassant le seuil des 215km/h, ce qui explique pourquoi la part de reconnaissance CatNat pour le cyclone Belal est restée limitée.

Trajectoire et vitesse de vent de Belal lors de son passage proche de la côte nord-est de La Réunion

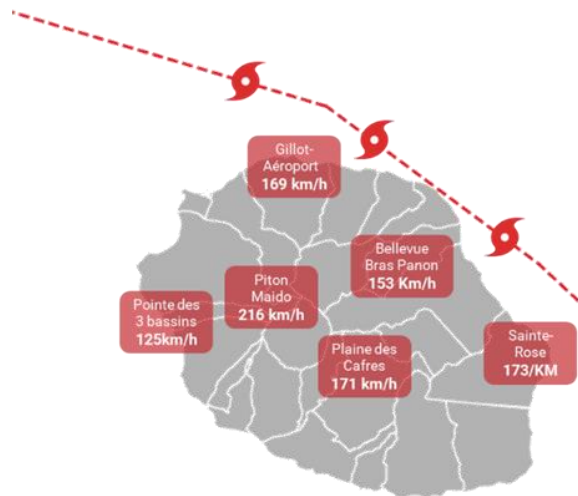


Figure 9 : Vitesse maximales du vent en km/h le 15 janvier 2024 enregistré à La Réunion. Source : Generali Climate Lab

Dans de tels cas, c'est la garantie TOC qui peut intervenir pour couvrir les dommages causés par ces vents violents. Cette situation peut représenter un défi financier important pour les assureurs, en particulier lorsque les dommages ne sont pas couverts par la reconnaissance en CatNat, car la CCR n'intervient pas. Le cyclone Belal, principalement couvert par la garantie TOC, est ainsi estimé à 100 millions d'euros. Un coût non négligeable pour les assureurs dont Prudence Créole, qui a une part de marché sur le dommage au bien de 17% à La Réunion.

Chapitre 3 : Contexte opérationnel et objectifs du mémoire

3.1. Impact des cyclones sur Prudence Créole : Enjeux et exposition aux risques

Prudence Créole, filiale de Generali France, est un acteur majeur de l'assurance à La Réunion et à Mayotte. Avec 17 % de part de marché en dommages aux biens¹⁷, elle se positionne comme le leader régional. En 2023, la compagnie a réalisé un chiffre d'affaires de 155 millions d'euros et un ratio combiné net (CoR) de 90,7 %¹⁸.

Encadré 3 - Le ratio combiné

Un ratio combiné est un indicateur utilisé dans le secteur de l'assurance pour évaluer la rentabilité d'une compagnie d'assurance. Il s'agit du rapport entre ce que l'assureur décaisse et ce qu'il encaisse. Ce ratio est calculé en additionnant les ratios de sinistralité (aussi appelé S/P) et de frais. Le ratio de sinistralité représente le rapport entre les sinistres payés (ou prévus) et les primes encaissées et le ratio de frais représente le rapport entre les frais de gestion (y compris les frais d'acquisition et d'administration) et les primes encaissées. Un ratio combiné inférieur à 100 % indique que l'assureur réalise un bénéfice, tandis qu'un ratio supérieur à 100 % signifie que l'assureur subit une perte.

L'analyse des CoR historiques de Prudence Créole permet de déceler un lien entre rentabilité et sinistralité cyclonique : lors des années avec une activité cyclonique élevée pour La Réunion, le CoR de Prudence Créole augmente. Le tableau ci-dessous illustre ce point.

Année	CoR	Cyclones
2018	95,5%	Fakir, Berguitta (et Dumazile)
2019	91,2%	
2020	87,5%	
2021	87,9%	
2022	94,2%	Batsirai, Emnati, Freddy
2023	90,7%	

Tableau 2 : Historiques de CoR de Prudence Créole depuis 2018. Source : Rapports RSSF de Prudence Créole.

En revanche, les années sans cyclone enregistrent une sinistralité climatique faible, limitée à quelques événements de moindre impact, tels que des pluies tropicales ou des mouvements de terrain. La sinistralité cyclonique est donc caractérisée par sa volatilité : les années sans cyclones majeurs ou sans succession de cyclones modérés se traduisent par une sinistralité climatique faible, tandis que les années avec des événements significatifs impactent fortement les résultats.

Parmi les branches d'activité de Prudence Créole, le dommage aux biens est la branche la plus exposée aux cyclones. En effet, Prudence Créole est peu implantée en agricole ou en transport maritime. Dans le cas de la branche automobile, les voitures peuvent être déplacées à l'abri avant l'arrivée d'un cyclone. Les propriétaires peuvent les garer dans des endroits plus sûrs, comme des garages ou des parkings souterrains, réduisant ainsi les risques de dommages directs. De plus, les pièces de rechange pour les voitures sont souvent plus facilement disponibles que celles nécessaires pour les réparations immobilières.

¹⁷ Donnée interne (étude menée dans le cadre d'une présentation à la CCR)

¹⁸ Rapport RSSF de Prudence Créole

D'où la nécessité pour Prudence Créole d'adapter sa stratégie de souscription et de gestion des risques dommages aux biens, notamment en recourant à la réassurance. La réassurance est une stratégie de transfert des risques, pouvant être considérée comme « l'assurance des compagnies d'assurance ». Il existe deux principaux types de réassurance : proportionnelle et non proportionnelle.

En réassurance proportionnelle, les primes et les sinistres sont partagés proportionnellement entre l'assureur et le réassureur. Le traité le plus courant dans cette catégorie est le traité en quote-part (QP). Par exemple, une quote-part à 50 % implique que le réassureur récupère 50 % des primes en contrepartie de la couverture de 50% des sinistres.

En réassurance non proportionnelle, le réassureur intervient uniquement au-delà d'un seuil prédéfini (la priorité) et jusqu'à une limite (la portée, qui peut être illimitée). Ces traités incluent :

- Les excédents de sinistres (XS), lorsque la priorité et la portée sont définies par un montant donné. Ces couvertures peuvent être par risque (un sinistre), par évènement (plusieurs sinistres ayant un même fait générateur : un cyclone par exemple) ou aggregate (plusieurs sinistres répondant à une même définition : l'ensemble des sinistres climatiques par exemple). Ces types de traités s'écrivent de la manière suivante : « *portée xs priorité* ». Par exemple un traité "120 M€ xs 5 M€" signifie que le réassureur intervient pour les montants supérieurs à 5 millions d'euros et jusqu'à 125 millions d'euros (5 + 120). Au-delà, la charge est à l'assureur.
- Les stop-loss (SL) sont similaires aux excédents de sinistres, mais avec des priorités et des portées définies en fonction des primes perçues par l'assureur.

3.2. Mécanisme de réassurance entre Prudence Créole et Generali IARD : Structure et implications

Pour se couvrir face aux cyclones, Prudence Créole bénéficie de la couverture climatique de deux réassureurs. Tout d'abord une couverture auprès de la Caisse Centrale de Réassurance (présentée dans la section 2.2). Deux traités successifs ont été conclus avec CCR, offrant une couverture spécifique contre les catastrophes naturelles liées aux cyclones : une quote-part 50 % suivie d'une couverture en stop-loss souscrite par Generali France consolidée (intégrant Generali France et ses filiales Outre-Mer), qui s'active si le ratio sinistres/primes (S/P) de la branche dommages aux biens dépasse *250 %*¹⁹ après application de la quote-part.

Ci-dessous, deux exemples illustrant la gestion d'un cyclone CatNat, selon l'activation ou non de la couverture stop-loss :

¹⁹ A partir de cette section, lorsqu'un chiffre est grisé et en italique, cela signifie qu'il a été modifié pour ne pas révéler d'informations confidentielles au sujet de Prudence Créole. Les mécanismes et les raisonnements derrière restent inchangés.

Survenance d'un cyclone lors d'une année sans Stop-Loss activé

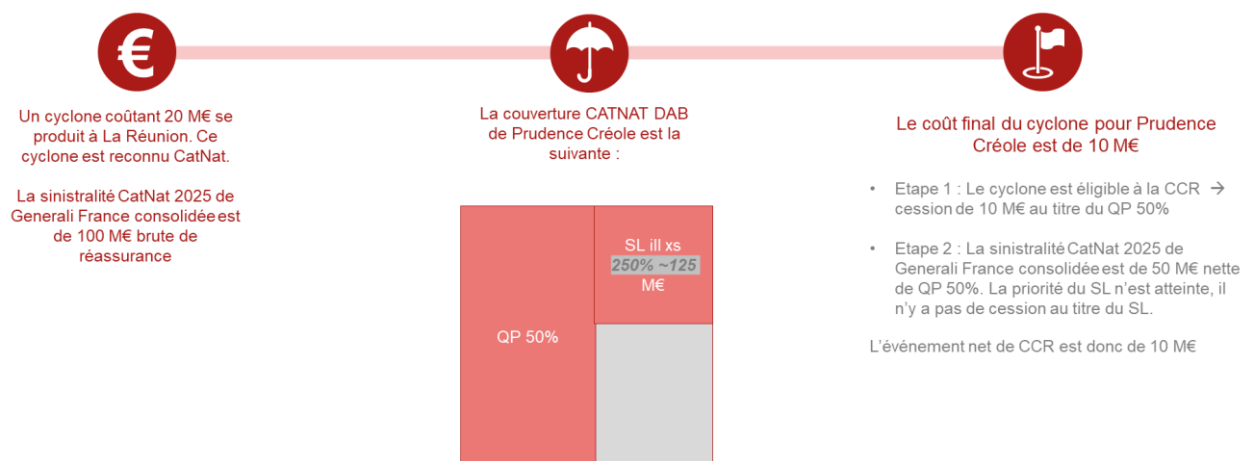


Figure 7 : Survenance d'un cyclone lors d'une année sans activation du Stop-Loss. Source : Formation réassurance de Generali France.

Survenance d'un cyclone lors d'une année avec activation du Stop-Loss

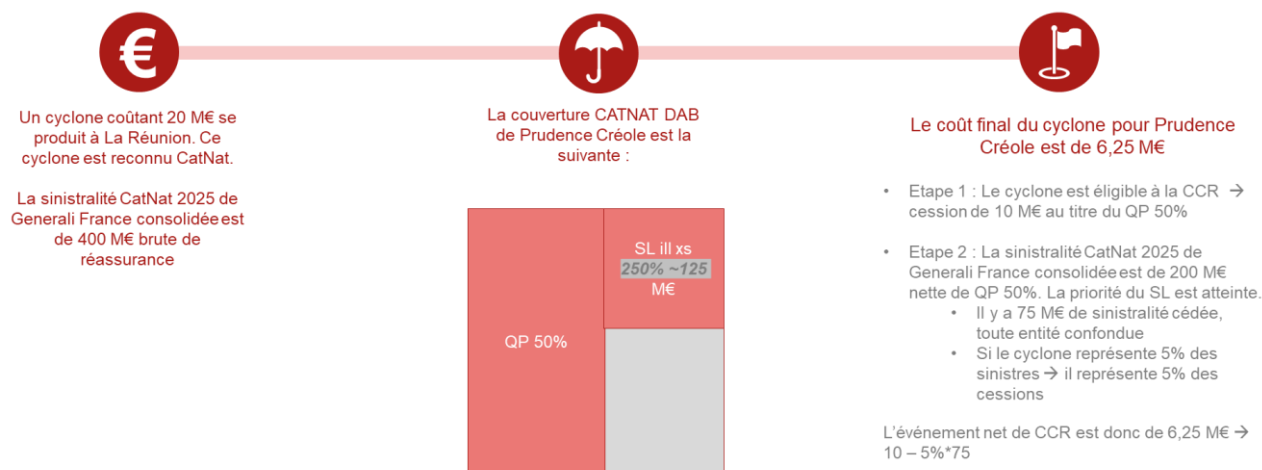


Figure 8 : Survenance d'un cyclone lors d'une année avec activation du Stop-Loss. Source : Formation réassurance de Generali France.

Cependant, à l'exception du cyclone Dina en 2002 dont l'ampleur était considérable, les cyclones ont une part négligeable de sinistralité CatNat. En effet, les vitesses de vent atteignent rarement les seuils requis pour le régime CatNat. En termes de « dégâts des eaux » climatiques, la majorité des dommages proviennent d'infiltration, péril non éligible au régime CatNat²⁰. Seules les inondations sont éligibles mais elles sont minoritaires. Par exemple, Batsirai et Belal, bien qu'importants pour Prudence Créole (Belal étant le deuxième plus gros cyclone après Dina), avaient une part de sinistralité CatNat inférieure à 5 %.

Generali France prend en charge les sinistres climatiques post cessions à la CCR. Cela équivaut à la somme, pour chaque cyclone, des sinistres CatNat après réassurance CCR et des autres sinistres climatiques exclus de la couverture de réassurance CatNat. Cette prise en charge est assurée grâce à un traité de réassurance de type 120 xs 5.

²⁰ Les infiltrations sont souvent considérées comme des dommages liés à l'entretien ou à des défauts de construction, ce qui les exclut du champ d'application du régime CatNat.

Structure de réassurance simplifiée pour les cyclones de Prudence Créole

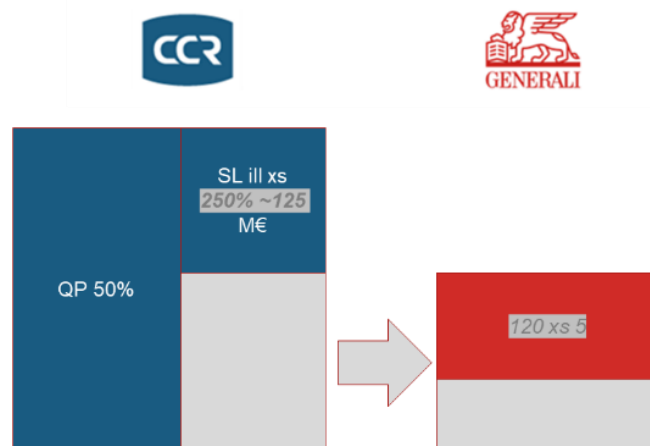


Figure 9 : Structure de réassurance simplifiée pour les cyclones de Prudence Créole. Source : documentation de la réassurance de Generali France.

Les cessions de Prudence Créole à Generali réduisent donc directement la sinistralité supportée par Prudence Créole et lui permettent de ne devoir supporter qu'au maximum 5 millions d'euros de sinistralité pour chaque cyclone (*modulo* les primes de REC²¹). En revanche, cela crée de la sinistralité climatique supplémentaire pour Generali France. A titre d'exemple, en 2024, les acceptations des cessions du cyclone Belal de Prudence Créole vers Generali France sont le deuxième événement climatique le plus coûteux pour Generali France (à la vision octobre 2024).

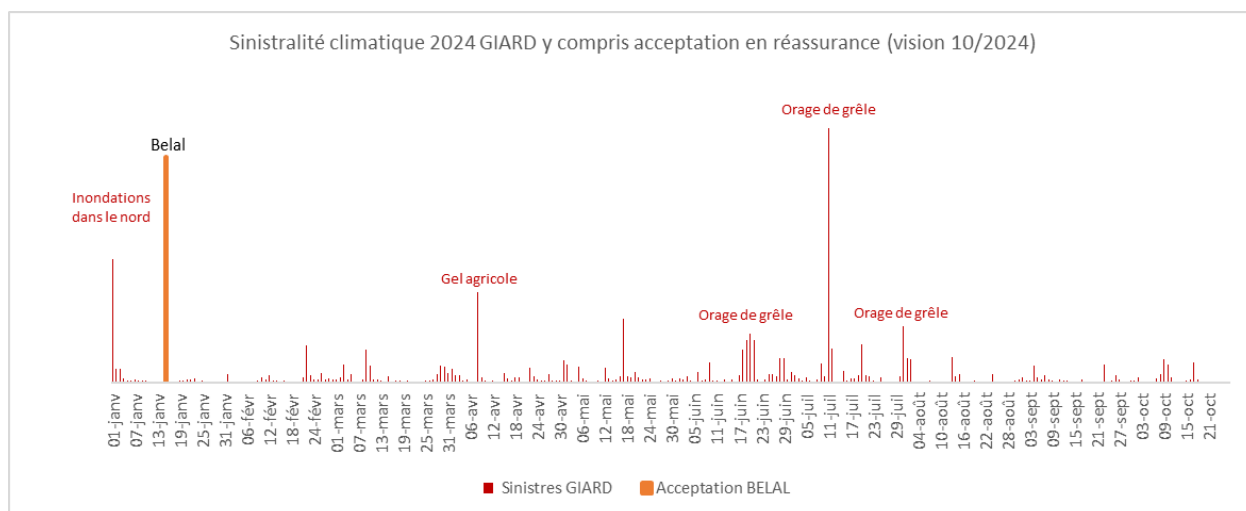


Figure 10 : Sinistralité climatique 2024 de Generali France y compris acceptation en réassurance. Source : données internes.

Cette situation souligne l'importance, pour Generali France de modéliser le coût des cyclones sur Prudence Créole, car ces événements exercent une influence significative sur sa propre sinistralité climatique.

²¹ Dès que la capacité d'une tranche (portion spécifique de la couverture assurée) est absorbée, l'assureur paye pour renouveler la portée unitaire d'une tranche : c'est la prime de reconstitution. Elle est déclenchée dès que la portée est impactée.

3.3. Objectif de l'étude : Identification des événements réassurés et estimation de leurs coûts pour Prudence Créole et Generali IARD.

Depuis 2002, La Réunion a été impactée par 17 cyclones majeurs. Parmi eux, 4 ont généré des charges dépassant le seuil de réassurance auprès de Generali France (une fois la sinistralité inflatée et ajustée en fonction du portefeuille de Prudence Créole en 2024 : notion d'AS-IF *détaillée dans une section ultérieure*).

Date	Nom	Phénomène	Cyclone réassuré en AS-IF ?
2002	Dina	Cyclone tropical très intense	Oui Cyclone le plus impactant pour Prudence Créole (et le marché).
2003	RAS		
2004			
2005			
2006	Diwa	Forte tempête tropicale	Non
2007	Gamède	Cyclone tropical intense	Oui
2008	RAS		
2009			
2010			
2011			
2012			
2013	Dumile	Cyclone tropical	Non
2014	Bejisa	Cyclone tropical intense	Non
2015	Haliba	Tempête tropicale	Non
2016	RAS		
2017	Carlos	Tempête tropicale	Non
2018	Berguitta	Tempête tropicale	Aucun cyclone réassuré en AS-IF Fakir proche du seuil Au global, 3 cyclones de moyenne ampleur ce qui est impactant
	Dumazile	Forte Tempête tropicale	
	Fakir	Forte Tempête tropicale	
2019	RAS		
2020	Diane	Forte tempête tropicale	Non
2021	Iman	Tempête tropicale	Non
2022	Batsirai	Cyclone tropical	Oui
	Emnati	Cyclone tropical	Non
	Freddy	Cyclone tropical intense	
2024	Belal	Cyclone tropical	Oui
	Candice	Tempête tropicale	Non

Tableau 3 : Tableau de l'activité cyclonique depuis 2002 et impact sur le portefeuille de Prudence Créole.

En moyenne, un cyclone nécessitant une réassurance AS-IF survient tous les quatre ans. Bien que ces événements soient peu fréquents, ils entraînent des coûts significatifs pour Prudence Créole et, par extension, pour Generali France.

De plus, la situation des départements d'outre-mer (DOM) alimente une certaine inquiétude chez les réassureurs, exacerbée depuis les dommages causés par le cyclone Irma²² dans les Antilles en 2017. Bien que La Réunion ne soit pas exposée de la même manière aux aléas climatiques, ces préoccupations ont un effet global sur les conditions de réassurance pour tous les DOM. Par ailleurs, les tensions sociales récentes, comme celles en Nouvelle-Calédonie et aux Antilles en 2024, renforcent cette

²² Les dommages assurés au titre de la garantie légale « catastrophes naturelles » hors automobile pour le cyclone Irma ont été estimés à 1,9 milliard d'euros (Source : CCR). Jusqu'à 2022, cet événement restait le plus impactant du régime CatNat.

méfiance. Cette situation constitue un enjeu pour Generali France, qui couvre Prudence Créole mais rétrocède une partie de la sinistralité à partir d'un certain seuil. Autrement dit, Generali est elle-même réassurée pour une partie des charges qu'elle accepte de Prudence Créole, comme le montre la structure de réassurance ci-dessous.

Structure de réassurance détaillée de Prudence Créole pour les cyclones

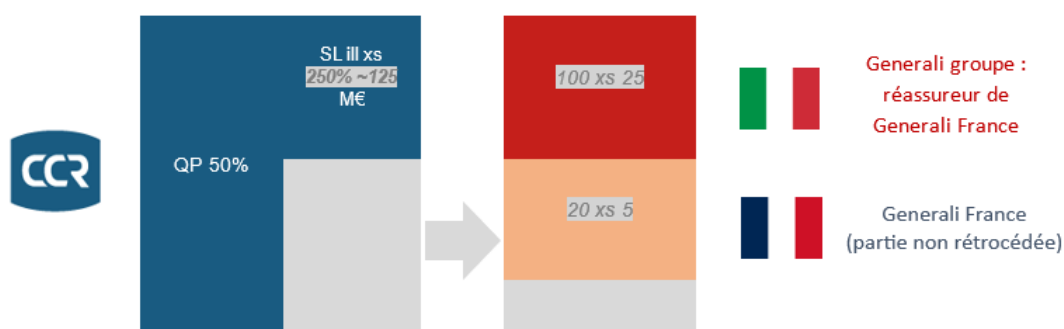


Figure 11 : Structure de réassurance détaillée de Prudence Créole.

Il convient de noter que Generali France est implantée dans tous les territoires d'outre-mer. Sa filiale GFA Caraïbes opère aux Antilles, tandis que ses activités s'étendent également à la Nouvelle-Calédonie et à la Polynésie française pour son propre compte. Cette présence dans des régions très variées complique la possibilité de mettre en avant la singularité de La Réunion auprès des réassureurs, qui perçoivent Generali France comme une seule entité regroupant l'ensemble des DOM, malgré leurs spécificités climatiques distinctes.

Cette dynamique génère une pression accrue de la part des rétrocessionnaires (réassureurs des réassureurs) pour obtenir des estimations rapides et précises des montants à céder lors d'un événement majeur en Outre-Mer. Cette exigence est d'autant plus primordiale dans le contexte de la réforme CatNat, où la charge de sinistralité, hors sécheresse, a augmenté de 6,5 % en raison de nouvelles mesures, comme la prise en charge des frais de relogement d'urgence (voir Annexe 1 : Réforme CatNat - Prise en charge des frais de relogement d'urgence: Réforme CatNat - Prise en charge des frais de relogement d'urgence).

Actuellement, Generali dispose de modèles robustes pour les périls majeurs en France hexagonale tels que les inondations, les tempêtes, la grêle ou la sécheresse. Toutefois, il n'existe pas encore de modélisation dédiée aux cyclones à La Réunion.

Ce mémoire s'inscrit dans une démarche visant à combler cette lacune en répondant à deux questions fondamentales :

- 1- Comment identifier rapidement, si un cyclone impactera Prudence Créole de sorte à déclencher un traité de réassurance ?
- 2- Comment estimer les coûts engendrés par ces événements pour Prudence Créole et Generali IARD ?

Ces réponses permettront non seulement d'approfondir la compréhension du risque cyclonique et d'anticiper les impacts financiers, mais aussi de démontrer une gestion rigoureuse et maîtrisée du portefeuille d'assurés. En fournissant des estimations rapides et précises, Generali France pourra renforcer la confiance des réassureurs. Cette transparence et cette réactivité favoriseront des négociations plus sereines et des conditions de réassurance optimales, tout en consolidant la position de Generali France en tant qu'acteur fiable dans les territoires d'outre-mer.

Partie 2 : Analyse des données et méthodologie de la modélisation du risque cyclonique à La Réunion

Chapitre 1 : Une modélisation jusqu'alors peu répandue

1.1. Rareté des modèles spécifiques à La Réunion

Le développement de modèles cycloniques spécifiques à La Réunion est primordial pour Prudence Créole, car ces outils restent rares en raison de nombreuses contraintes. Selon l'étude "*Estimating Tropical Cyclone Vulnerability: A Review of Different Open-Source Approaches*" (Ward et al., 2020), l'efficacité des modèles disponibles pour l'océan Indien est largement inférieure à celle des modèles développés pour les régions atlantiques et pacifiques.

L'étude souligne que, dans les Caraïbes et le Pacifique, les modélisations bénéficient d'une infrastructure robuste soutenue par des centres de recherche en météorologie et modélisation climatique situés aux États-Unis, au Japon et en Australie. Ces pays investissent massivement dans la collecte de données et le développement d'outils de modélisation, permettant une calibration fine des modèles et une adaptation aux spécificités locales. En revanche, La Réunion, en raison de son isolement géographique et de l'absence de soutien de la part de grandes nations investissant dans la recherche sur les risques naturels de la région, dispose de modèles moins performants et moins nombreux.

Les difficultés liées à la disponibilité des données historiques exacerbent cette situation. Avant les années 1990, les données sur la vitesse du vent et les caractéristiques des cyclones dans l'océan Indien sud manquaient de cohérence, en raison de l'absence de suivi par des systèmes satellitaires et radars sophistiqués, contrairement aux Caraïbes, qui bénéficient de décennies de données fiables grâce à un réseau de surveillance dense et diversifié. Cette lacune de données limite la capacité des modèles à bien refléter l'intensité, la trajectoire et la distribution spatiale des dégâts potentiels des cyclones.

Par ailleurs, l'hétérogénéité géographique de La Réunion, avec ses reliefs abrupts, ses microclimats variés et son réseau hydrologique sensible, complique l'adaptation des modèles développés pour d'autres bassins. Les précipitations cycloniques sont amplifiées par la topographie de l'île, ce qui peut engendrer des crues soudaines et des glissements de terrain. Les interactions entre les microclimats locaux et les perturbations cycloniques diffèrent significativement des caractéristiques météorologiques des bassins caribéens ou pacifiques, nécessitant des ajustements spécifiques que les modèles génériques ne peuvent pas intégrer sans modifications majeures.

Enfin, le coût élevé et la complexité du développement de nouveaux modèles adaptés à La Réunion ne justifient pas, pour de nombreuses entreprises de modélisation, un investissement significatif compte tenu de la taille réduite du marché local. Cela contraste avec les investissements massifs réalisés pour les modélisations dans les Caraïbes et l'Atlantique nord, où le volume de risques assurés est plus important et où les modèles sont calibrés sur des décennies de sinistralité. En conséquence, les assureurs locaux doivent s'appuyer sur des outils de modélisation moins performants, ce qui les expose à des estimations biaisées des sinistres cycloniques et rend difficile la gestion proactive des risques.

1.2. Des modèles marché se focalisant sur les effets du vent malgré l'importance des infiltrations à La Réunion

L'étude de Ward et al. (2020) mentionne également que, même lorsque des modèles de marché sont disponibles, leur transposition à La Réunion est souvent incomplète car ils se concentrent essentiellement sur les effets du vent et négligent d'autres phénomènes tout aussi importants, comme les infiltrations d'eau ou les glissements de terrain.

Si ces modèles se focalisent sur le vent, c'est parce qu'ils ont été développés pour les régions où les vents violents représentent le péril principal, comme les Caraïbes, le Golfe du Mexique et le sud-est des États-Unis. Cette approche est justifiée dans les territoires plats et homogènes, où les vents violents causent la plupart des dommages matériels. Cependant, elle s'avère inadaptée pour des régions comme La Réunion, où les caractéristiques topographiques et hydrologiques locales entraînent une plus grande vulnérabilité aux infiltrations d'eau et aux crues soudaines.

De plus, les résultats de l'étude de 2013 SYCODES de l'Agence Qualité Construction (AQC)²³ montrent que les désordres liés aux infiltrations d'eau sont particulièrement fréquents à La Réunion. Par exemple, les infiltrations aux points singuliers, tels que les jonctions entre différents matériaux ou parties d'un bâtiment, ainsi que les défauts d'étanchéité (notamment au niveau des toitures terrasses et des façades), représentent une part significative des sinistres observés sur le territoire. La complexité de ces sinistres, résultant de la combinaison de pluies intenses, de ruissellements et de la configuration géographique de l'île, rend leur modélisation particulièrement difficile avec les modèles centrés sur les vents.

Illustrations de failles de construction pouvant provoquer des infiltrations d'eau



Figure 12 : Deux illustrations de failles de construction pouvant provoquer des infiltrations d'eau. Source : Ootravau, bricozone.

Malgré ces observations, très peu de modèles de marché incluent les infiltrations et inondations comme des périls principaux dans le cadre des risques cycloniques.

Il apparaît donc essentiel de développer un modèle de risque multi-périls simplifié, capable de représenter correctement les divers impacts cycloniques spécifiques à La Réunion, et de ne pas se limiter aux seuls effets du vent, comme c'est souvent le cas avec les modèles de marché actuels.

²³ https://www.reunion.developpement-durable.gouv.fr/IMG/pdf/Comite_technique_CRC_20juin2013_donnees_desordres_Reunion_cle26fdfd.pdf

1.3. Méthodologie actuelle pour estimer les coûts des cyclones chez Prudence Créole

En l'absence de modèles adaptés à La Réunion, Prudence Créole a développé, à partir de son expérience terrain, une méthode simple d'estimation des coûts cycloniques auprès de ses réassureurs.

Pour communiquer un coût à la CCR, la démarche est généralement rapide, car les critères d'éligibilité CatNat sont rarement atteints. Les vitesses de vent mesurées, communiquées par Météo France dans la semaine suivant l'événement, sont généralement en deçà des seuils requis pour activer le régime CatNat (cf. *Partie 1 : Cadre de l'étude et contexte*). Les inondations, lorsqu'elles surviennent, restent limitées en sinistralité et ne constituent pas un facteur dominant dans les sinistres causés par un cyclone. Seuls les événements exceptionnels, combinant des vents et des inondations reconnues CatNat, justifient une communication plus complexe auprès de la CCR. Depuis 2002, le cyclone Dina est le seul exemple où ces deux critères ont été réunis. Dans ce cas particulier, Prudence Créole applique une méthode d'évaluation des coûts similaire à celle utilisée pour les estimations auprès de Generali France.

Pour communiquer un coût à Generali France, Prudence Créole adopte une méthode qui peut être facilement explicable en instance. Selon Sophia Aly²⁴, responsable du provisionnement des cyclones chez Prudence Créole, cette méthode repose sur plusieurs étapes.

1. Identification d'un cyclone intense

Les employés de Prudence Créole, subissant à titre personnel les cyclones, savent assez rapidement si un cyclone sera d'ampleur ou non. Cette évaluation initiale, reposant sur un jugement d'expert, est appuyée par des observations terrain. Ce « pressenti » a été fiable à chaque cyclone que Prudence Créole a dû suivre mais pourrait être complété par une approche scientifique afin de justifier rapidement le fait de ne pas déployer un suivi en gestion de crise auprès des réassureurs, qui de l'extérieur, ont parfois une vision déformée de la réelle ampleur des dommages.

2. Collecte d'une partie des données via un numéro vert

Lorsqu'un cyclone survient, un numéro vert est rapidement activé pour faciliter les déclarations de sinistres par les assurés. Par exemple, lors du cyclone Belal en janvier 2024, un centre d'appel dédié a été ouvert dès le lendemain de l'événement pour centraliser les déclarations par téléphone et via un portail en ligne.



Figure 13 : Communication de Prudence Créole suite au passage du cyclone BELAL.

²⁴ Cette section est le compte rendu d'une interview qui a été effectuée auprès de Sophia Aly afin de mieux comprendre les méthodes actuelles de provisionnement des cyclones chez Prudence Créole et les attentes de la compagnie sur des éventuels compléments.

3. Utilisation d'un triangle sur les nombres d'appel

Une fois le numéro vert ouvert, Prudence Créole suit quotidiennement le nombre d'appels reçus.

Ensuite, un cyclone de référence est sélectionné, présentant le profil le plus similaire possible au cyclone en cours, en termes de ressenti et d'impact local observé. Par exemple, pour le cyclone Belal de 2024, le cyclone Batsirai de 2022 a été retenu comme référence.

Prudence Créole compare alors les appels reçus pour le cyclone en cours à ceux observés lors du cyclone de référence. Avec des triangles de développement²⁵, Prudence Créole applique une dynamique d'évolution similaire à celle du cyclone de référence pour le nombre d'appels par jour et détermine ainsi le nombre total d'appels attendus pour le cyclone en cours.

4. Application de coûts moyens

Une fois le nombre d'appels estimé, des coûts moyens spécifiques sont appliqués pour les particuliers et pour les professionnels. Cela permet d'obtenir une estimation initiale du coût global du cyclone pour les sinistres déclarés en numéro vert.

5. Complément par les sinistres déclarés en agence

Cependant les sinistres ne sont pas exclusivement déclarés via le numéro vert. Les déclarations en agence sont également possibles. La méthode pour les estimer est d'appliquer des triangles de coûts basés là encore sur les données du cyclone de référence.

Bien que cette méthodologie donne des résultats relativement proches de la réalité Prudence Créole a identifié des axes d'amélioration. Tout d'abord, l'estimation repose fortement sur le cyclone pris comme référence, ce qui peut entraîner des biais dans des contextes exceptionnels ou atypiques. Le fait d'utiliser Batsirai, un cyclone très marqué par les pluies, pour « modéliser » Belal, un cyclone très marqué par le vent, en est une illustration. Ensuite, la compagnie souhaiterait limiter les écarts entre l'estimation initiale et le coût réel. Dans le cas de Belal, le delta était de 15% par rapport aux montants réels. Du point de vue de Generali France, en tant que réassureur, le delta était acceptable car ce qui compte en début d'événement, c'est d'avoir un bon ordre de grandeur, puis de stabiliser les estimations par la suite, ce qui a été le cas pour le cyclone Belal. La priorité des deux compagnies est plutôt d'intégrer les données météorologiques et de portefeuille dans les calculs pour renforcer la robustesse des projections.

Dans ce contexte, ce mémoire propose de développer une méthode simple et complémentaire de modélisation, combinant les données historiques des sinistres et les caractéristiques de portefeuille de Prudence Créole ainsi que les données météorologiques de Météo France. Cette approche vise à répondre aux axes d'amélioration de la méthode actuelle. La première étape du travail consiste à identifier et organiser les données disponibles pour renforcer les estimations de Prudence Créole.

²⁵ Un triangle pour les nombres de sinistres est un outil actuariel qui organise les données des sinistres par année de survenance et par année de développement. Il permet de suivre le nombre de sinistres déclarés au fil du temps. Cela aide les actuaires à estimer les réserves nécessaires pour couvrir les sinistres futurs en appliquant les tendances historiques

Chapitre 2 : Les données disponibles et leurs pré-traitements

2.1. Données de sinistralité historique

L'objectif principal de cette étape est de consolider et fiabiliser les données relatives à la sinistralité cyclonique afin de les intégrer dans le modèle.

Les données de sinistralité de Prudence Créole ont été fournies par Prudence Créole et couvrent une période allant de 2002 à 2024. En général, lors de l'utilisation d'une base de données de sinistres, l'année en cours est exclue car les sinistres pourraient encore évoluer. Ici, l'inclusion de 2024 a été rendue possible grâce à la stabilisation rapide de la sinistralité associée aux événements cycloniques (entre 3 et 4 mois, les cyclones de 2024 ayant eu lieu en janvier et le début de l'analyse en août).

Onze bases de données ont été fournies, avec des formats différents, et se composant d'un fichier consolidé regroupant les sinistres cycloniques survenus entre 2005 et 2018, des fichiers spécifiques pour les événements plus récents ainsi qu'un fichier sur le cyclone Dina en 2002. Chaque fichier contient des informations variées relatives aux contrats, aux caractéristiques des sinistres et aux charges financières.

Pour cette étude, seules les variables essentielles à la modélisation et suffisamment renseignées ont été retenues. Cela inclut les caractéristiques relatives aux contrats, telles que les numéros d'identification, le code postal et le produit d'assurance, ainsi que les caractéristiques propres aux sinistres, à savoir les dates de survenance, les charges sinistres, et l'identification d'un sinistre en tant que CatNat.

En revanche, certaines variables ont été exclues de la modélisation en raison de leur incohérence, de leur faible pertinence ou du fait qu'elles soient très peu renseignées. Par exemple, la garantie appliquée n'a pas été conservée, car les informations disponibles variaient au fil des années, avec des modifications fréquentes de nomenclature qui rendaient leur harmonisation trop complexe. De plus, certains cyclones, notamment les plus anciens, n'avaient aucune information sur les garanties. De même, la variable identifiant le péril principal, c'est à dire la cause du sinistre, s'est révélée inexploitable, car son contenu, souvent saisi manuellement par les gestionnaires, était trop subjectif ou incomplet. À titre d'illustration, pour le cyclone Gamède, 68% des périls étaient identifiés "Voir libellé libre", sans qu'il soit possible de distinguer s'il s'agissait de dégâts causés par le vent ou par des infiltrations. Cette variable a en revanche permis de pouvoir exclure les sinistres qui n'étaient pas climatiques (identifiés "Responsabilité civile" par exemple).

Le tableau ci-dessous liste les variables reçues et si elles ont servi à la modélisation.

Variable	Libellé	Exemple	Variable conservée pour la modélisation ?
ENVCODEV	Nom de la compagnie	PRUDENCE	OUI
DTARET	Date de dernière MAJ	2020-11-08	OUI
NOPERS	Identifiant client	Confidentiel	OUI
NOMPRS	Nom du client	Confidentiel	OUI
PRODUI	Code produit	3A01	OUI
LIBPRD	Libellé produit	Incendie Risque Simple	OUI
LoB	Ligne de Business	GMRH	OUI
DEPTSI	Département du sinistre	RUN	OUI
ANNESI	Année du sinistre	2006	OUI
NDOSSI	Numéro du sinistre	Confidentiel	OUI
Doublon	Variable créée par Prudence Créole pour supprimer les doublons	Ne pas enlever	OUI (Grâce à cette variable un filtre a été effectué pour ne garder aucun doublon)
DTSURV	Date de survenance	2006-03-05	OUI

Type EE	Type d'événement	Cyclone	OUI
Personnalisation EE	Nom de l'événement	Gamède	OUI
LBCIRC	Cause du sinistre	Bien assuré endommagé par tiers	OUI pour enlever le non climatique
CODGAR	Code garantie	80000	NON, majoritairement non renseigné
LBGARA	Libellé produit	Incendie Explosion Bâtiment	NON, majoritairement non renseigné
LBETAT	Etat du dossier	Clôturé	OUI pour enlever les sinistres sans suite
CN / HCN	Sinistre CatNat ?	HCN	OUI
Charge sinistre	Coût du sinistre	24 553,13 €	OUI
WOD1LIG1	Adresse partie 1		NON, majoritairement non renseigné
WOD1LIG2	Adresse partie 2		NON, majoritairement non renseigné
WOD1LIG3	Adresse partie 3		NON, majoritairement non renseigné
WOD1LIG4	Commune	97434 ST GILLES LES BAINS	OUI
Quartier	Quartier	Saint Gilles Les Bains	NON, majoritairement non renseigné
WOD1CPOS	Code postal	97434	OUI
WOD1LPY	Code pays iso	REUNION	OUI

Tableau 4 : Tableau de variables reçues pour la base de sinistralité historique. Source : Prudence Créole.

Comme onze fichiers ont été reçus, les données présentaient des disparités importantes entre les différents fichiers, en raison de la diversité des événements cycloniques et de l'évolution des méthodes de gestion des bases au fil des années. Ces fichiers ont donc nécessité des traitements spécifiques pour uniformiser les variables, gérer les incohérences et rassembler l'ensemble des informations dans une base unique adaptée à la modélisation.

Les fichiers ont été importés dans R pour un premier travail de formatage, notamment sur les dates.

Chaque base a ensuite été passée en revue pour harmoniser les variables essentielles et combler certaines valeurs manquantes. Par exemple pour le cyclone Dina, l'identification d'un sinistre en CatNat n'étaient pas renseignées. En analysant les montants bruts et nets de réassurance CCR dans les templates de réassurance²⁶, il a été possible de déduire que la quasi-totalité des sinistres Dina étaient reconnus CatNat et ainsi de remplir la variable car la charge brute de Dina était égale au double de sa charge nette, caractéristique du quote-part 50 % avec la CCR. Un proxy a alors été d'identifier l'ensemble des sinistres de Dina en CatNat.

En ce qui concerne les produits et LoB (Line of Business, exemple : multirisque habitation) non renseignés sur certaines bases, des valeurs temporaires "Non attribué" ont été assignées sur R en attendant un traitement sur Excel qui sera présenté ultérieurement.

Enfin, les sinistres ont été vérifiés pour s'assurer qu'ils étaient correctement rattachés à leur année de survenance, avec correction des erreurs ponctuelles dans les dates. Il est facile de corriger ce type d'erreur car elle portait à chaque fois sur l'année de survenance et les sinistres sont rattachés à un cyclone nommé et que les jours et mois coïncident avec le cyclone en question.

Ce travail d'uniformisation a abouti à une base consolidée regroupant l'ensemble des cyclones. Une vérification a ensuite été effectuée à l'aide des templates de réassurance, afin de s'assurer que les montants financiers de chaque cyclone étaient conformes aux données transmises aux réassureurs. Cette étape a validé la qualité et la cohérence de la base pour la modélisation et les analyses futures. Il est à noter que la base des sinistres était déjà de très bonne qualité au moment de la réception car à l'exception de quelques fautes de frappes, elle ne comportait pas d'erreurs bloquante, aucune donnée

²⁶ Un template de réassurance est un format standardisé utilisé pour structurer les informations nécessaires à la négociation et à la gestion des contrats de réassurance. Il inclut des données comme les sinistres historiques, les primes, les montants réassurés et les limites des couvertures. Ce format facilite les échanges entre assureurs et réassureurs en assurant une présentation claire et cohérente des données. C'est donc la communication auprès des réassureurs qui sert de référence.

de charge de sinistre vide ou incohérente et les doublons étaient pré-identifiés par Prudence Créole, ce qui a facilité leur gestion.

Après l'uniformisation et l'intégration des données dans une base consolidée, un filtrage a été effectué pour écarter les derniers sinistres non pertinents. Les sinistres liés à l'auto ont été exclus, car cette catégorie relève d'un traité de réassurance spécifique et reste négligeable dans le cadre de la réassurance cyclonique. De plus, les dossiers dont l'état du sinistre est "sans suite" ont été éliminés. Ces derniers, après analyse, ne sont pas éligibles à une indemnité climatique et, par conséquent, ne peuvent être réassurés. De même pour les sinistres présents dans la base alors qu'ils n'étaient pas dus à une cause climatique. Enfin, les doublons identifiés par Prudence Créole ont été supprimés.

Après ce tri, un fichier de travail a été créé sur Excel pour faciliter les analyses et les vérifications. Ce fichier comprend plusieurs feuilles détaillées ci-dessous.

Feuille de tables de correspondances

La première feuille contient les tables de correspondance créées pour agréger plus facilement les informations.

La première table de correspondance concerne les informations géographiques, qui présentaient initialement 421 combinaisons différentes de communes et de codes postaux, en raison de saisies manuelles disparates par les gestionnaires de sinistres au fil du temps. En croisant les codes postaux et les noms de communes, chaque sinistre a été rattaché à l'une des 24 communes officielles de La Réunion. Le tableau ci-dessous illustre le processus de normalisation des données géographiques, en prenant l'exemple de la commune de Sainte-Suzanne :

Base de données de Prudence Créole		Infos harmonisées	
Code postal	Commune	Code Postal	Commune
97441	97441 BAGATELLE ²⁷	97441	Sainte Suzanne
97441	97441 SAINTE SUZANNE	97441	Sainte Suzanne
97441	97441 SAINTE-SUZANNE	97441	Sainte Suzanne
97441	97441 SAINYTE SUZANNE	97441	Sainte Suzanne
97441	97441 SANTE SUZANNE	97441	Sainte Suzanne
97441	97441 STE SUZANNE	97441	Sainte Suzanne
97441	97441 STE-SUZANNE	97441	Sainte Suzanne
97441	Sainte Suzanne	97441	Sainte Suzanne
97441	STE SUZANNE	97441	Sainte Suzanne

Tableau 5 : Tableau illustratif des saisies manuelles des communes. Source : base de données de Prudence Créole.

Les communes incohérentes ou non attribuables (par exemple, "00000 ILE DE LA REUNION") ont été exclues, représentant environ 1 % des sinistres.

La deuxième table regroupe les informations sur les produits et les branches d'assurance. Les données initiales, comportaient le nom du produit, qui n'était parfois pas assez précis pour être utilisé tel quel (exemple : Multirisque) et la LoB avec un niveau de détails cohérent avec les reportings de Generali mais souvent vide. Ces deux variables ont permis d'en créer une troisième, la branche, regroupant les informations sur le produit en trois grandes catégories : Multirisque Habitation (MRH), Multirisque Commerce (MRC), et Risque Industriel (RI). La table de correspondance est présentée ci-dessous.

²⁷ Quartier de la commune de Sainte Suzanne

Produit	LoB	BRANCHE
100% Pro Artisans- commerçants	na	MRC
100% Pro Artisans- commerçants	GENERALI multi-risques commerçants	MRC
100% Pro Services	na	MRC
100% Pro Services	GENERALI multi-risques commerçants	MRC
Incendie Risque Simple	na	MRC
Incendie Risque Simple	GENERALI multi-risques commerçants	MRC
M.Risque Professionnelle MCA	na	MRC
M.Risque Professionnelle MCA	GENERALI multi-risques commerçants	MRC
Multirisque	na	MRC
Multirisque	GENERALI multi-risques commerçants	MRC
Inc. Rsq Simple Propriétaire	na	MRC
INC. RSQ COMMERC LOCATAIRE	na	MRC
INC. RSQ COMMERC PROPRIETAIRE	na	MRC
MULTIRISQUE PROFESSIONNELLE	na	MRC
MULTI-RISQUES COM.LOCATAIRE	na	MRC
MULTI-RISQUES COM.PROP.NON-OCC	na	MRC
MULTI-RISQUES COM.PROPRIET.OCC	na	MRC
MULTI-RISQUES GARAGISTES	na	MRC
M.R. Propriétaire non Occupant	na	MRH
M.R. Propriétaire non Occupant	GENERALI MRH	MRH
M.R.H PROP.OCCUPANT/LOCATAIRE	na	MRH
M.R.H PROP.OCCUPANT/LOCATAIRE	GENERALI MRH	MRH
Multirisque Habitation	na	MRH
Multirisque Habitation	GENERALI MRH	MRH
Multirisque Habitation SHLMR	GENERALI MRH	MRH
MULTI. TROPIQUE LOCATAIRE	na	MRH
MULTI. TROPIQUE LOG. FONCTION	na	MRH
MULTI. TROPIQUE PROP.OCC.	na	MRH
MULTI-RISQUES HAB.LOC/PROP.OCC	na	MRH
MULTI-RISQUES HAB.LOG.FONCTION	na	MRH
MULTI-RISQUES MIXTE LOCATAIRE	na	MRH
MULTI-RISQUES PROP. NON. OCC.	na	MRH
INC. RSQ INDUST LOCATAIRE	na	RI
INC. RSQ INDUST PROPRIETAIRE	na	RI
MULTI-RISQUE INDUSTRIELLE	na	RI
Multirisque Risq.Entreprise	na	RI
Multirisque Risq.Entreprise	GENERALI MRI	RI

Figure 14 : Table de correspondance permettant d'obtenir la branche.

Feuille des données d'inflation et d'évolution du portefeuille

La feuille dédiée à l'inflation et à l'évolution du portefeuille permet de revaloriser les sinistres historiques selon une approche AS-IF. Cela consiste à adapter les sinistres passés pour qu'ils reflètent les conditions économiques et assurantielles actuelles. En d'autres termes, cette méthode simule un scénario où ces sinistres auraient eu lieu dans le contexte d'aujourd'hui, en tenant compte de deux facteurs principaux :

1. **L'évolution du portefeuille.** La structure actuelle du portefeuille de Prudence Créole diffère considérablement de celle observée il y a 20 ans, en raison de la croissance démographique (+ de 20%²⁸) de La Réunion et de la stratégie de développement menée par la compagnie. Pour intégrer cette évolution, un coefficient de revalorisation a été calculé pour chaque branche de chaque année étudiée. Ce dernier correspond au rapport entre le nombre de risques souscrits en 2024 et le nombre de risques correspondant à l'année considérée. L'utilisation du nombre de risques comme base de calcul s'explique par les données historiques disponibles dans les templates de réassurance, qui comprennent les primes et le nombre de risques. Les primes étant déjà ajustées pour l'inflation d'une année sur l'autre, le nombre de risques représente une mesure plus fiable pour isoler l'évolution du portefeuille au fil du temps. La recherche des données anciennes pour la revalorisation des sinistres a nécessité un travail considérable. Il n'existait pas de fichier historisé des nombres de sites, seulement des fichiers N/N-1. Pour chaque année de cyclone étudiée, il fallait donc consulter le fichier de l'année du cyclone et récupérer l'information. Entre 2002 et 2024, les méthodologies de rangement, les formats des fichiers et parfois même les informations disponibles changeaient, nécessitant des vérifications de cohérence. La méthode la plus simple consistait à comparer le nombre de sites de l'année N dans le reporting de l'année N avec celui de l'année N+1. Cependant, cette information

²⁸ Source : Insee

n'était pas toujours disponible. Par exemple, l'année 2002 a été particulièrement difficile à retrouver. En raison du passage à l'euro, les reportings se concentraient davantage sur la conversion des montants en francs en euros que sur le nombre de sites. L'information sur le nombre de sites en 2002 n'est disponible que dans le reporting de 2003. Pour la vérifier, les données de 2003 à 2005 ont été rapatriées, bien qu'il n'y ait pas eu de cyclone pendant ces années. Cela a permis de vérifier que l'évolution du nombre de sites entre 2002 et 2006, année du cyclone Gamède, était cohérente.

2. L'inflation. Pour prendre en compte les évolutions économiques sur la période étudiée, un coefficient d'inflation a été calculé à partir des indices spécifiques à chaque branche d'assurance.

Pour les branches Multirisque Habitation et Multirisque Commerce, l'indice FFB (Fédération Française du Bâtiment) a été utilisé, car il reflète les variations des coûts dans le secteur de la construction. Cet indice, dont l'évolution est présentée ci-dessous, prend en compte les fluctuations des prix des matériaux, des salaires, ainsi que des équipements nécessaires aux travaux.

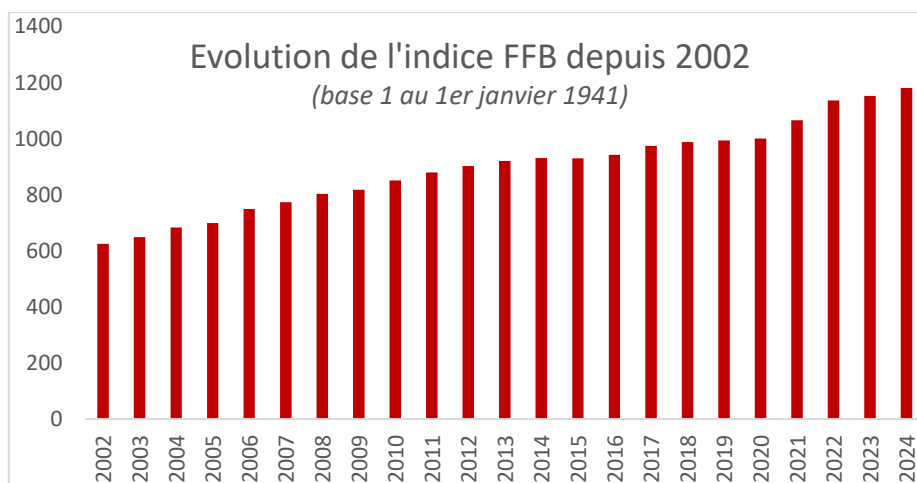


Figure 15 : Évolution de l'indice FFB (2002-2024). Réalisé sur Excel à partir des données ffbatiment.

Pour la branche Risque Industriel, l'ajustement a été effectué à l'aide de l'indice Risque Industriel, qui mesure l'évolution des coûts directement liés aux activités industrielles.

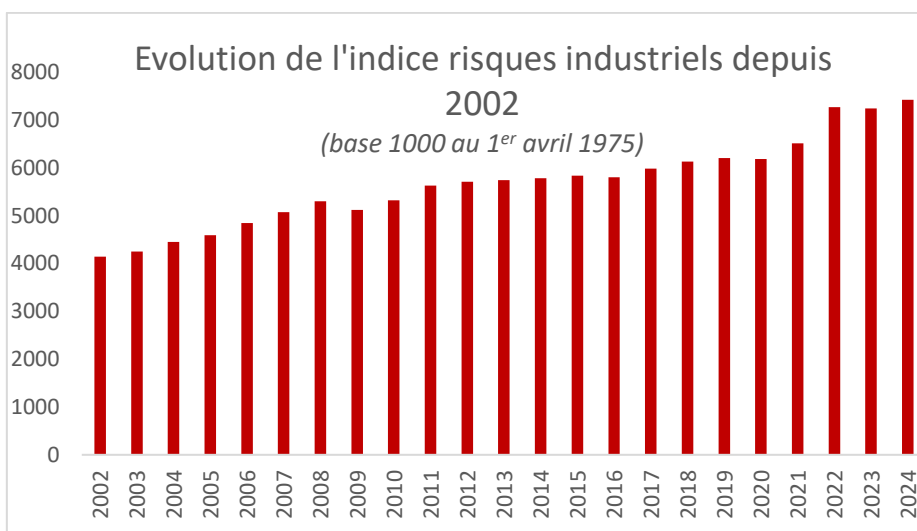


Figure 16 : Évolution de l'indice Risque Industriel (2002-2024). Réalisé sur Excel à partir des données assurances-gesco.

Le coefficient d'inflation a été calculé en divisant la valeur de 2024 de l'indice sélectionné par celle de l'année concernée. Les indices FFB et RI n'étant pas encore disponibles pour 2024 au moment des travaux de mise en AS-IF, une estimation a été réalisée en prenant l'indice de 2023 majoré de 2,5 %²⁹, conformément aux prévisions d'inflation publiées par la Banque de France en juin 2024.

Finalement, en multipliant le coefficient d'inflation avec le coefficient d'évolution du portefeuille, un coefficient global a été calculé pour ajuster les charges sinistres aux évolutions de portefeuille et d'inflation.

Le tableau ci-dessous synthétise les coefficients de revalorisation appliqués aux sinistres MRH pour chaque année, en combinant les ajustements liés à l'évolution du portefeuille et à l'inflation.

Pour des raisons de confidentialité, certaines colonnes ont été modifiées en affichage. La colonne des coefficients de portefeuille a été remplacée par des valeurs reflétant directement l'indice FFB, afin d'éviter de communiquer sur l'évolution du portefeuille de Prudence Créole. Les valeurs utilisées dans le modèle pour les revalorisations sont cependant bien celles qui comportent les données de Prudence Créole.

Annee	Branche	Revalo_ptf marginale	REVALO_ptf	Inflation	REVALO_inflation	REVALO_ptf x inflation
2002	MRH	Nb police	190%	FFB 624,9	190%	361%
2003	MRH	Nb police	183%	FFB 648,7	183%	335%
2004	MRH	Nb police	174%	FFB 683,2	174%	302%
2005	MRH	Nb police	170%	FFB 699,4	170%	288%
2006	MRH	Nb police	158%	FFB 749,1	158%	251%
2007	MRH	Nb police	153%	FFB 773,5	153%	236%
2008	MRH	Nb police	148%	FFB 802,9	148%	219%
2009	MRH	Nb police	145%	FFB 817,9	145%	211%
2010	MRH	Nb police	139%	FFB 851,2	139%	195%
2011	MRH	Nb police	135%	FFB 879,8	135%	182%
2012	MRH	Nb police	131%	FFB 903,1	131%	173%
2013	MRH	Nb police	129%	FFB 920,8	129%	166%
2014	MRH	Nb police	128%	FFB 930,8	128%	163%
2015	MRH	Nb police	128%	FFB 929,5	128%	163%
2016	MRH	Nb police	126%	FFB 942	126%	159%
2017	MRH	Nb police	122%	FFB 974,8	122%	148%
2018	MRH	Nb police	120%	FFB 988,2	120%	144%
2019	MRH	Nb police	119%	FFB 994,3	119%	143%
2020	MRH	Nb police	119%	FFB 1000,5	119%	141%
2021	MRH	Nb police	111%	FFB 1066,4	111%	124%
2022	MRH	Nb police	104%	FFB 1137	104%	109%
2023	MRH	Nb police	103%	FFB 1152,6	103%	106%
2024	MRH	Nb police	100%	FFB +2,5%	1181,4	100%

Tableau 6 : Coefficients de revalorisation AS-IF par année pour la branche MRH. Réalisé avec Excel.

Feuille pour la base de données agrégée finale

La dernière étape a consisté à enrichir la base de données agrégée en créant des variables clés nécessaires à la modélisation, à partir des tables de correspondance et des coefficients de revalorisation.

1. Branche d'assurance : Chaque sinistre a été rattaché à l'une des trois branches définies (MRH, MRC, RI) en utilisant les correspondances établies entre les produits et les LoB. Les branches ainsi que l'année de survenance permettent alors la mise en AS-IF.
2. Charges sinistres revalorisées : Une variable *Charge_sinistre_inflatée* a été calculée en appliquant les coefficients AS-IF aux charges sinistres.
3. Réforme CatNat : Pour les sinistres éligibles au régime CatNat, une majoration de 6,5 % a été ajoutée aux charges inflatées pour tenir compte de l'impact de la réforme, créant la variable *Charge_sinistre_inflatée_Reforme*. Cette réforme a déjà été évoquée dans la première partie et pour rappel ne concerne que très peu de cyclone (principalement Dina).

²⁹ Source : <https://www.banque-france.fr/fr/actualites/projections-macroeconomiques-juin-2024>

Le graphique ci-dessous illustre l'évolution de la charge de sinistre après avoir pris en compte l'évolution du portefeuille, l'inflation et la réforme CatNat.

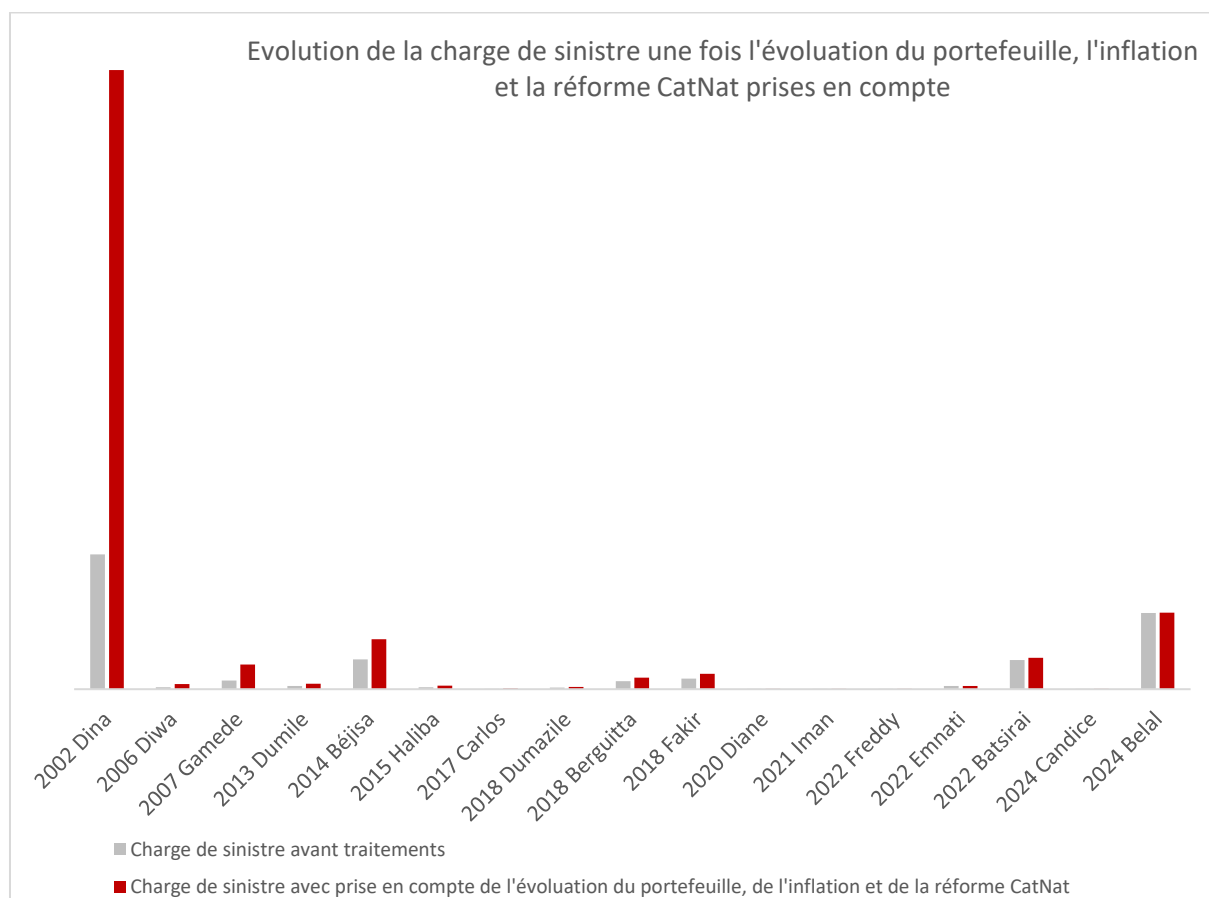


Figure 17 : Evolution de la charge de sinistre une fois l'évolution du portefeuille, l'inflation et la réforme CatNat prises en compte. Réalisé avec Excel.

Cette approche a le mérite d'offrir une vision actuelle des sinistres, y compris changements réglementaires. Elle comporte cependant certaines limites : elle ne prend pas en compte les changements structurels tels que l'évolution de la vulnérabilité des bâtiments ou des comportements, ni l'impact du changement climatique sur les coûts des sinistres.

La base sinistres finale inclut, pour chaque sinistre, l'année de survenance, le nom de l'événement, la commune, la branche et les coûts revalorisés avec réforme CatNat. Les informations sur l'adresse étant insuffisamment renseignées dans les données de sinistralité, contrairement aux communes qui le sont systématiquement, l'ensemble des données sera donc traité à l'échelle communale pour la suite de l'analyse.

2.2. Données d'exposition

L'objectif de cette section est de structurer les données d'exposition afin d'obtenir, pour chaque cyclone et chaque commune, les informations nécessaires à la modélisation des risques cycloniques.

Les données d'exposition désignent l'ensemble des informations relatives aux biens assurés, nécessaires pour évaluer l'exposition financière d'un assureur en cas de sinistre. Issues du portefeuille dommages aux biens de Prudence Créole, elles rassemblent des informations telles que la branche

d'assurance, le type de bâtiment, le mode d'occupation (propriétaire, locataire, etc.), les caractéristiques structurelles du bâtiment (béton, etc.), les sommes assurées, ainsi que certains éléments contractuels, comme les franchises et les taux de coassurances³⁰. Le tableau en Annexe 2 les énumère.

Pré-traitement pour Generali France

Le groupe Generali France, composé de plusieurs filiales³¹, dispose de nomenclatures spécifiques pour chacune d'entre elles (format de dates différents, etc.). Cette diversité rend nécessaire une harmonisation des données avant leur intégration dans les systèmes d'analyse de Generali France.

Avant tout traitement, les variables reçues sont analysées pour détecter d'éventuelles valeurs manquantes. Si le taux de valeurs manquantes se dégrade de plus de 5 %, par rapport au dernier fichier de portefeuille transmis, Prudence Créole est sollicitée pour fournir une version actualisée et plus complète de la base.

Control ID	Control Name	Control Description	Data Instance Name	Pourcentage d'anomalies	Pourcentage d'anomalies N-1	Commentaire	Période	Type	Company
DQC - FR015 - adresse - 041	Zeros	Vérifie s'il y a des valeurs nulles	adresse	0%	0%	Stable	062024	DAB	PC
DQC - FR015 - code_org - 041	Zeros	Vérifie s'il y a des valeurs nulles	code_org	0%	0%	Stable	062024	DAB	PC
DQC - FR015 - coass - 041	Zeros	Vérifie s'il y a des valeurs nulles	coass	0,2%	1%	Better	062024	DAB	PC

Tableau 7 : Exemple d'analyse des valeurs manquantes pour les variables adresse et coassurance et code organisation. Source : données internes.

Une fois ce contrôle réalisé, les valeurs manquantes sont remplacées en s'appuyant sur des règles définies à dire d'expert, adaptées aux spécificités des données d'assurance. Voici trois exemples d'imputation :

- Taux de coassurance : Lorsque cette information est absente, il est supposé qu'aucune coassurance n'a été mise en place, et la valeur est imputée à 100 %.
- Somme assurée pour les propriétaires : Si la somme assurée du bâtiment n'est pas renseignée, elle est estimée en multipliant une valeur moyenne au m² par la surface du bien, fournissant ainsi une approximation cohérente.
- Adresse manquante : Lorsque l'adresse complète n'est pas disponible, mais que les coordonnées géographiques (latitude et longitude) sont connues, ces dernières sont utilisées pour déduire l'adresse du bien.

Dans l'ensemble, la qualité des données disponibles pour le portefeuille dommages aux biens de Prudence Créole est très élevée et il y a peu de valeurs manquantes. Le tableau suivant analyse le taux de complétude des variables de portefeuille étudiées pour modéliser la vulnérabilité face aux cyclones par les modèles marché. Si le taux est supérieur à 90%, la base de données est considérée comme étant de qualité (standard RMS³²).

³⁰ La coassurance est une technique d'assurance où plusieurs compagnies se partagent la couverture d'un même risque. Chaque assureur prend en charge une partie du risque, proportionnellement à la quote-part qu'il a acceptée. Cette méthode est souvent utilisée pour des risques élevés ou des biens de grande valeur.

³¹ L'Équité en Hexagone, GFA Caraïbes aux Antilles et en Guyane et Prudence Créole à La Réunion et Mayotte.

³² RMS (Risk Management Solutions) est un acteur de référence dans le développement de modèles probabilistes de risques pour les catastrophes climatiques.

Variables	Taux de complétude supérieur à 90%
Sommes assurées (SA)	OUI
Primes	OUI
Mode d'occupation (propriétaire, locataire, etc.)	OUI
Type de construction ³³	NON
Nombre d'étage	OUI
Type d'usage (résidentiel, industriel, etc.)	OUI
Géocodage (latitude + longitude)	OUI

Tableau 8 : Tableau synthétique sur la qualité des données. Source : données internes.

Parmi ces variables, seule celle relative au type de construction présente une qualité inférieure à 90 %. Cette lacune sera traitée dans les paragraphes suivants.

Ensuite, le nombre de sites, le nombre de contrats et toutes les variables quantitatives (sommes assurées, primes, etc.) sont comparés à l'année précédente. Si une variation de plus de 5 % est détectée, une explication qualitative est demandée à Prudence Créole. De même, si des variations contradictoires sont observées (par exemple, une augmentation du nombre de sites et des primes, mais une baisse des sommes assurées), une explication est requise. Ces analyses sont effectuées par branche (MRH, MRC, RI) pour mieux comprendre le comportement du portefeuille de Prudence Créole. Par exemple, si Prudence Créole décide de se développer en MRH et de résilier certains contrats RI, le portefeuille global pourrait montrer une augmentation du nombre de sites malgré une baisse des sommes assurées, les risques industriels ayant des sommes assurées très élevées par rapport aux habitations. D'où l'importance d'une analyse par branche.

De plus, les enregistrements présentant des valeurs par défaut incohérentes, telles que 999999999 pour les sommes assurées, ont été remplacées par zéro pour éviter toute surestimation de l'exposition aux risques dans les analyses.

Pour éviter tout problème d'encodage les caractères spéciaux présents dans les variables textuelles, notamment les noms de communes, ont été normalisés. Par exemple, les accents ont été supprimés ou remplacés par des caractères standards, ce qui a conduit à transformer *Saint-André* en *Saint-Andre*.

Un travail similaire a été effectué sur les dates. Elles ont été converties au format standard (YYYY-MM-DD), permettant une gestion des périodes de couverture et facilitant les analyses temporelles ultérieures.

Enfin, les engagements bruts ont été recalculés à partir des données de sommes assurées. Ils représentent l'ensemble des sommes assurées pour un même risque (bâtiment, contenu...) prévues dans le contrat. Par exemple, l'engagement brut pour la garantie tempête ouragan cyclone (TOC) pour un particulier propriétaire est obtenu en additionnant les sommes assurées (SA) comme ceci :

$$\text{Engagement dommage direct TOC brut pour un particulier propriétaire} = SA_{BAT} + SA_{CONT}$$

Dans le cas d'un locataire, l'engagement se limite aux sommes assurées "Contenu", les dommages sur les bâtiments étant propre aux propriétaires. Enfin, si l'assuré est un professionnel, il faut également ajouter les sommes assurées pertes d'exploitation, frais et pertes, etc.

³³ Le type de construction compris est : structure tout béton, structure Béton avec toiture différente, Structure acier et toiture autre que béton, Inconnu

Une fois les engagements bruts harmonisés, les engagements nets ont été calculés en appliquant les limites contractuelles d'indemnisation (LCI) et les pourcentages de coassurance (Coass). Une LCI est une clause dans certains contrats d'assurance qui limite le montant maximal que l'assureur est tenu de verser en cas de sinistre. Cette clause permet à l'assureur de définir une limite de couverture, réduisant ainsi sa responsabilité financière. Ces ajustements permettent de refléter précisément les montants réellement supportés par Prudence Créole, après application des clauses spécifiques à chaque contrat.

Par exemple, pour un contrat présentant une somme assurée brute de 500 000 €, une limite d'indemnisation de 400 000 € et une coassurance de 80 %, le calcul des engagements nets s'effectue en deux étapes :

1. Application de la limite d'indemnisation

Engagement net après limite

$$= \min(\text{somme assurée brute}, \text{limite d'indemnisation})$$

$$= \min(500\,000, 400\,000) = 400\,000\text{€}$$

2. Application de la coassurance

Engagement net après coassurance

$$= \text{Engagement net après limite} * \frac{\text{Coassurance}}{100} = 400\,000 * 0,8 = 320\,000\text{€}$$

Les primes ont également été ajustées pour tenir compte des clauses de coassurance, et connaître la part réellement perçue par Prudence Créole. La méthode de calcul est simple : la prime nette est obtenue en multipliant la prime totale du contrat par le pourcentage de coassurance.

Par exemple, pour un contrat ayant une prime totale de 2 000 € avec une coassurance fixée à 80 %, la prime nette est calculée comme suit :

$$\text{Prime totale net} = \text{prime totale} * \frac{\text{Coassurance}}{100} = 2\,000 * 80\% = 1\,600$$

Traitements spécifiques pour le modèle cyclonique

À partir des données harmonisées et retraitées par Generali France, deux axes de travail majeurs ont été mis en œuvre pour adapter les données d'exposition aux besoins spécifiques du modèle développé dans ce mémoire : l'amélioration de la qualité des données relatives au type de bâtiment, seule variable avec un taux de complétude inférieur à 90%, et le traitement des informations géographiques pour une agrégation par commune, conforme à la maille d'étude des sinistres.

D'une part, le niveau de complétude de la classification des constructions, inférieur à 90%, s'explique par une mauvaise transposition des données initiales de Prudence Créole. Bien que ces données comprenaient 57 modalités détaillées de types de construction, elles n'ont pas été correctement regroupées dans les quatre grandes catégories utilisées pour analyser de manière harmonisée l'ensemble des entités de Generali France.

En effet, la granularité initiale, bien qu'informatrice, rendait les données difficiles à exploiter pour la modélisation. Pour pallier cette complexité, une classification a été établie regroupant les constructions en quatre grandes catégories inspirées des méthodologies développées par RMS. Cette simplification permet de mieux prendre en compte les caractéristiques clés de vulnérabilité des bâtiments, notamment leur résistance aux vents violents. Le tableau ci-dessous présente la classification des types de construction.

Nom	Niveau de risque	Commentaire
Structure tout béton	Faible	Cette catégorie regroupe les bâtiments dont la structure principale et le toit sont entièrement en béton. Ils offrent une grande résistance aux vents violents typiques des cyclones, mais aussi aux infiltrations et inondations. Grâce à leur construction solide, ces bâtiments ont un risque d'infiltration limité.
Structure béton avec toiture différente	Intermédiaire	Ces bâtiments ont une structure en béton mais avec une toiture en matériaux plus légers, comme des tôles métalliques ou des matières plastiques. Bien que la structure en béton offre une bonne résistance aux vents et inondations, les matériaux plus légers utilisés pour le toit peuvent être plus vulnérables aux infiltrations et aux détériorations lors des fortes rafales ou en cas de pluie intense. Les toitures en métal, notamment, sont plus sensibles aux fuites d'eau.
Structure acier et toiture autre que béton	Important	Les bâtiments avec une ossature en acier et une toiture en matériaux légers (comme les tôles ou les matières plastiques) sont particulièrement vulnérables aux cyclones. Ces toits peuvent être facilement endommagés par les vents violents, ce qui expose la structure en dessous à des infiltrations d'eau, aggravant ainsi les risques d'inondation à l'intérieur.
Inconnu	Important	Cette catégorie regroupe les bâtiments pour lesquels les informations sur la structure ou la toiture ne sont pas complètes ou sont incohérentes. Ils peuvent présenter des vulnérabilités inconnues face aux vents, inondations et infiltrations, ils seront donc considérés comme dangereux.

Tableau 9 : Tableau de catégories de type de construction.

À titre d'illustration, l'exemple suivant montre comment un type de construction initial, renseigné sous le code "BEBA" dans la base de Prudence Créole, a été reclassé dans l'une des quatre catégories définies alors qu'il n'était pas attribué auparavant. Grâce à une table de correspondance détaillée fournie par Prudence Créole, il est possible de décomposer ce code pour en déduire que :

- BE, correspond à une structure en béton,
- BA, correspond à une toiture en bardeaux d'asphalte.

À partir de ces deux éléments, une nouvelle classification est attribuée : "Structure béton et toiture différente", en conformité avec les catégories standardisées établies pour simplifier et homogénéiser l'analyse des types de construction.

Code renseigné par Prudence Créole	Type de structure (obtenu grâce à une table de correspondance)	Type de toit (obtenu grâce à une table de correspondance)	Nouvelle classification
BEBA	BE Béton	BA Bardeaux d'asphaltes	Structure béton et toiture différente

Tableau 10 : Exemple d'application de la nouvelle classification des types de construction.

À l'issue de ce travail, le taux de complétude de la variable détaillant la structure du bâtiment est passée au-dessus du seuil des 90%. D'autre part, les sommes assurées dont le type de construction est inconnu représentent une part négligeable du portefeuille, elles ne seront donc pas exploitées par la suite.

D'autre part, les données d'exposition incluaient des coordonnées géographiques précises (latitude et longitude), permettant une localisation fine des biens assurés. Cependant, les données de sinistralité disponibles sont, quant à elles, agrégées à l'échelle communale. En effet, il n'est possible qu'à partir de 2018, de relier directement les sinistres cycloniques aux données d'exposition grâce à un identifiant unique (numéro client). Cette approche aurait impliqué de limiter l'analyse aux événements survenus après cette date. Cela aurait significativement réduit la quantité de données disponibles pour construire le modèle, excluant des cyclones majeurs comme Dina en 2002 ou Gamede en 2007 et limitant le nombre de cyclones analysés à 10.

En conséquence, un choix méthodologique a été fait : travailler à l'échelle communale. Les données d'exposition, initialement détaillées à l'échelle des bâtiments, ont été agrégées par commune afin d'être compatibles avec les données de sinistralité. Ce processus a permis de regrouper les informations sur les sommes assurées et les caractéristiques associées par commune.

Par exemple, dans le tableau ci-dessous :

- Les sommes assurées (SA) ne sont plus calculées par risques, mais sont désormais agrégées au niveau communal.
- Des variables spécifiques ont été créées pour représenter les sommes assurées par type de construction et par commune.

Identifiant	Commune	Type de construction	Somme assurée
1	Saint Paul	Tout béton	10
2	Sainte Suzanne	Structure acier et toiture béton	30
3	Sainte Suzanne	Structure acier et toiture béton	12
4	Saint Paul	Structure Béton avec toiture différente	15

↓

Commune	SA totale	SA en béton	SA avec structure acier et toiture béton	SA avec Structure Béton avec toiture différente
Saint Paul	25	10	0	15
Sainte Suzanne	42	0	42	0

Figure 18 : Exemple d'agrégation des données par commune

Cette approche s'applique également à d'autres caractéristiques, telles que le mode d'occupation (propriétaire, locataire) ou l'usage des bâtiments (résidentiel, commercial, industriel).

Le nombre de risques a été écarté car il ne reflète pas l'exposition financière réelle du portefeuille.

Il n'y a pas non plus d'analyse sur le nombre d'étage car la quasi-totalité du portefeuille étudié est en rez-de-chaussée, ce qui n'apportait pas d'information pour le modèle.

Ainsi, à ce stade, la base d'exposition contient les informations ci-dessous pour chaque cyclone et chaque commune, inspirée des variables d'analyse de la vulnérabilité standard sur le marché (cf. Tableau 8) :

- Sommes assurées (SA) totale
- Sommes assurées béton
- SA structure béton avec toiture différente
- SA acier
- SA Propriétaire
- SA Locataire

- SA Appartement
- SA Immeuble
- SA Maison
- SA usage Commercial
- SA usage Industriel

En complément, les données météorologiques, essentielles pour caractériser l'intensité des cyclones, enrichissent cette base. Elles permettent d'établir des liens entre les conditions climatiques observées et les dommages assurés, constituant ainsi une base solide pour les analyses et modélisations à venir.

2.3. Données météorologique

L'objectif de cette section est de décrire la provenance des données météorologiques, les variables sélectionnées, ainsi que les traitements spécifiques réalisés pour les rendre compatibles avec les besoins du modèle. La méthodologie qui va suivre a été présentée et validée par Laurent Boissier, responsable du pôle recherche climatique chez Generali France.

Les données climatiques utilisées dans ce mémoire sont issues des bases en open source mises à disposition par Météo France³⁴. Ces données comprennent des enregistrements quotidiens de données climatiques pour les stations météorologiques de La Réunion, depuis 1950. Les stations sélectionnées pour cette étude sont celles encore en activité, garantissant la cohérence et la continuité des informations au cours du temps. Les variables pour cette analyse incluent :

- NOM_USUEL : Nom de la station d'enregistrement,
- LAT et LON : Coordonnées géographiques de la station,
- ALTI : Altitude de la station (en mètres),
- AAAAMMJHH : Date et heure de la mesure,
- RR : Quantité de précipitations sur 24 heures (en mm),
- FXI : Vitesse maximale instantanée des rafales de vent sur 24 heures (en m/s),
- FXY : Vitesse moyenne sur 10 minutes des vents soutenus (en m/s).

Bien que les rafales de vents expliquent mieux les dégâts que les vents soutenus sur 10 minutes, les deux caractéristiques de vent sont gardées car elles sont toutes les deux critères de reconnaissance CatNat (215km/h en rafale et 145km/h en vent soutenu sur 10 minutes).

Les données météorologiques ont été filtrées à partir de 2002 pour ne retenir que les périodes correspondant aux passages des cyclones étudiés. En effet, l'objectif n'est pas de prédire la survenance d'un cyclone, mais d'évaluer son impact une fois qu'il est survenu à proximité de La Réunion. Il est donc plus pertinent de s'attarder sur les conditions météorologiques extrêmes. De plus, comme étagé dans la partie 1, en dehors des cyclones, la sinistralité climatique est négligeable. Il n'y aurait donc pas de données de sinistres à comparer aux données climatiques en dehors des cyclones.

Un filtre supplémentaire a été appliqué pour exclure certaines stations météorologiques situées à des altitudes supérieures à 1000 mètres. Ces stations présentent souvent des vitesses de vent très élevées, qui ne sont pas représentatives des zones habitées où se concentre l'essentiel des biens assurés. Leur inclusion dans le calcul des données météorologiques par commune, auraient artificiellement amplifié l'intensité des aléas dans des zones où peu ou pas d'assurés résident.

Par exemple, lors du cyclone Belal, la station Piton Maïdo, situé à 2150 mètres d'altitude, a enregistré une rafale de vent à 217 km/h, dépassant ainsi les critères de reconnaissance CatNat pour les vents. Cependant, aucune reconnaissance CatNat n'a été attribuée pour le vent pour cet

³⁴ <https://meteo.data.gouv.fr/form>

événement, car peu de personnes vivent dans cette zone. Cet exemple illustre l'importance d'exclure ces stations pour garantir une modélisation cohérente et adaptée à la réalité des zones assurées.

Pour illustrer cela, la carte ci-dessous présente la région de Saint-Denis, où les stations de la *Plaine des Chicots* et du *Brûlé*, affichées en gris sur la carte ci-dessous, ont été écartées. Tandis que les stations proches des zones côtières, où se concentre la majorité des assurés de Prudence Créole (en bleu sur la carte), ont été conservées.

Répartition des assurés et des stations météorologiques retenues et exclues dans la région de Saint-Denis

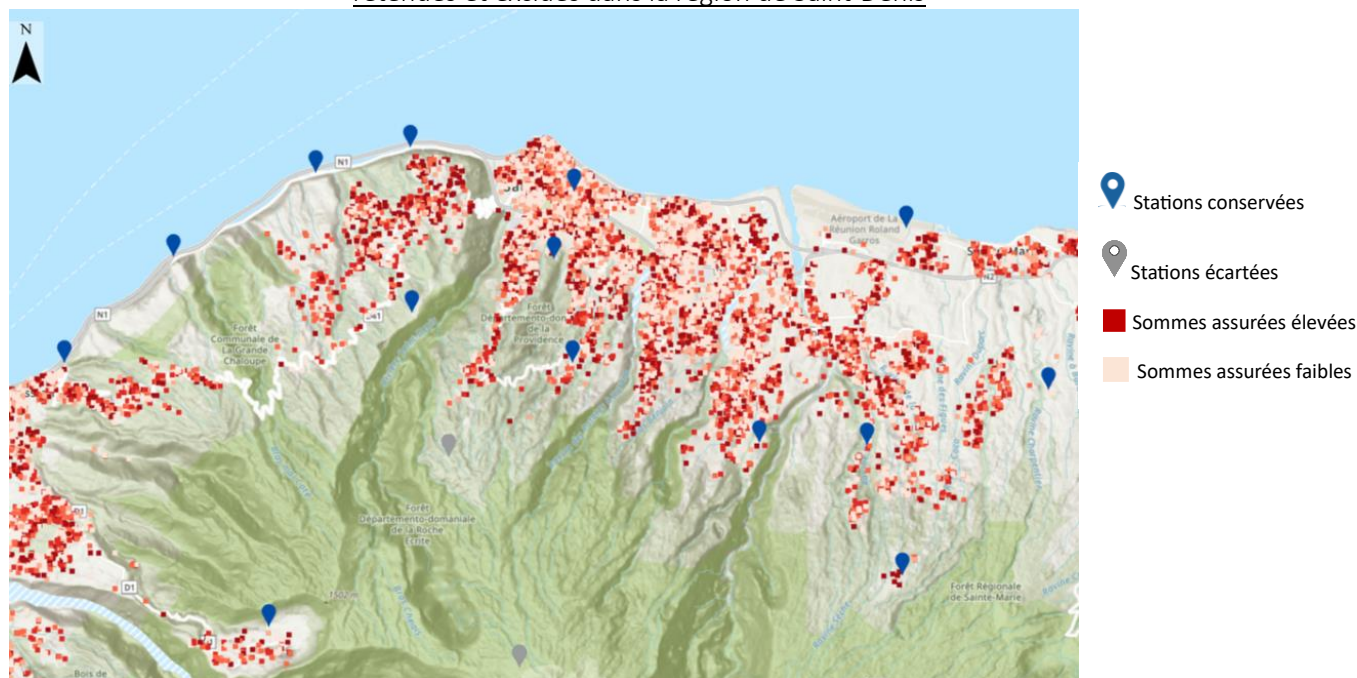


Figure 19 : Carte des stations météorologiques retenues et exclues pour la région de Saint-Denis. Carte réalisée avec ArcGIS.

Traitements des données de vents

Les vitesses de vent, initialement mesurées en m/s, ont été converties en km/h pour faciliter leur interprétation.

Les rafales de vent ont été retenues comme indicateur pour le vent pour plusieurs raisons. D'une part, les données de rafales fournies par Météo France sont plus complètes que celles des vents soutenus sur 10 minutes, ce qui garantit une meilleure analyse des événements. D'autre part, les rafales sont souvent publiées très rapidement par Météo France après le passage des cyclones majeurs, ce qui permet de pouvoir les utiliser plus facilement et rapidement dans le modèle.

Enfin, leur intensité explique de manière plus directe les préjudices causés aux bâtiments, et donc les impacts sur le portefeuille dommages aux biens de Prudence Créole. Les vitesses de vent moyennes sur dix minutes ne sont donc pas utilisées pour la modélisation des coûts bruts de réassurance mais elles restent une variable d'analyse lors du passage du calcul du coût brut au coût net de réassurance CCR.

Toutefois, en raison du nombre limité de stations météorologiques sur l'île, comme le montre la carte ci-dessous, une interpolation spatiale a été nécessaire pour estimer les vitesses de vent en tout point de La Réunion.

Carte des stations météorologiques ayant enregistré
des vitesses de vents lors des cyclones à La Réunion

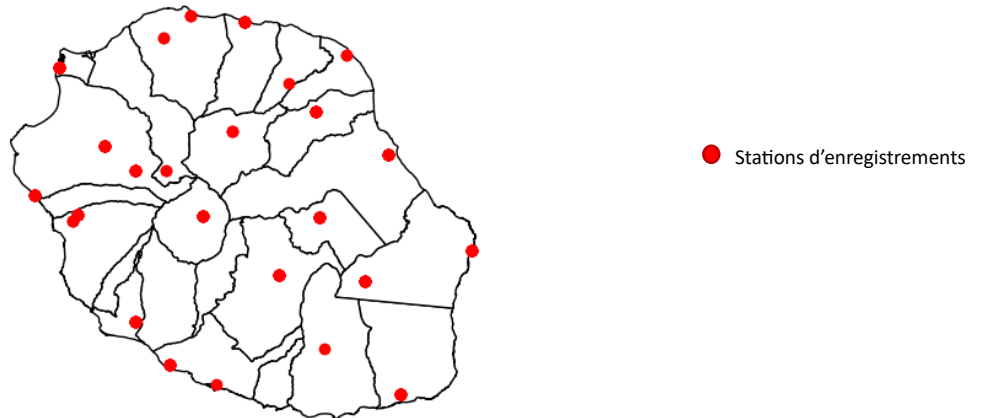


Figure 20 : Carte des stations météorologiques de La Réunion ayant enregistré des vitesses de vents lors de cyclones. Réalisé sous python grâce au package ArcPy.

La méthode utilisée est l'interpolation spatiale par pondération inverse de la distance (IDW), avec une pondération au cube. Cette technique permet de modéliser les vitesses de vent en fonction de la proximité des stations.

Encadré 4 - L'interpolation spatiale par pondération inverse de la distance (IDW)

L'interpolation spatiale est une technique utilisée pour estimer les valeurs de variables géographiques à des emplacements non mesurés, en se basant sur les valeurs connues à des points de mesure voisins. Cette méthode repose sur le principe que des points proches géographiquement ont des valeurs similaires

La méthode de pondération inverse, inspirée de la notion de moyenne pondérée, permet de modéliser des variables continues en attribuant un poids inversement proportionnel à la distance entre le point à interpoler et les points de mesure disponibles.

L'équation pour calculer la valeur interpolée $u(x)$ au point x s'exprime comme suit :

$$u(x) = \frac{\sum_{k=1}^N \frac{u_k}{d(x, x_k)^3}}{\sum_{k=1}^N \frac{1}{d(x, x_k)^3}}$$

Où :

- $w_k(x) = \frac{1}{d(x, x_k)^3}$ est la fonction de pondération,
 - x est le point à interpoler,
 - x_k représente les points d'interpolation connus,
 - $d(x, x_k)$ est la distance entre le point x et x_k ,
- u_k représente la valeur de la vitesse du vent au point connu x_k ,
- N est le nombre total de points pris en compte dans l'interpolation.

Cette méthode repose sur le principe selon lequel plus un point connu est éloigné du point d'interpolation, plus son poids diminue, réduisant ainsi son influence sur la valeur interpolée. Elle est particulièrement utile pour lisser les informations météorologiques, telles que les vitesses de vent, à partir des données collectées par les stations météorologiques.

L'IDW présente certaines limites dans les zones montagneuses, car elle ne prend pas directement en compte des facteurs comme l'altitude ou la topographie. Cela peut engendrer des biais dans des régions de relief complexe, comme à La Réunion.

Néanmoins, ce problème est atténué par le fait que les stations en altitude ont été majoritairement écartées de l'analyse. De plus, l'utilisation d'une distance au cube améliore la précision en reflétant mieux les variations locales, essentielles dans une région comme La Réunion où les phénomènes météorologiques, tels que les vents et les précipitations, évoluent rapidement sur de courtes distances. En diminuant significativement l'influence des points éloignés et en renforçant celle des points proches, cette méthode permet de mieux représenter les spécificités locales tout en limitant un lissage excessif des résultats. Elle garantit ainsi une interprétation plus fidèle des phénomènes cycloniques à l'échelle territoriale.

Le graphique suivant présente une carte des vitesses de vent interpolées pour le cyclone Béjisa sur l'ensemble de l'île, réalisée à l'aide de la méthode IDW.

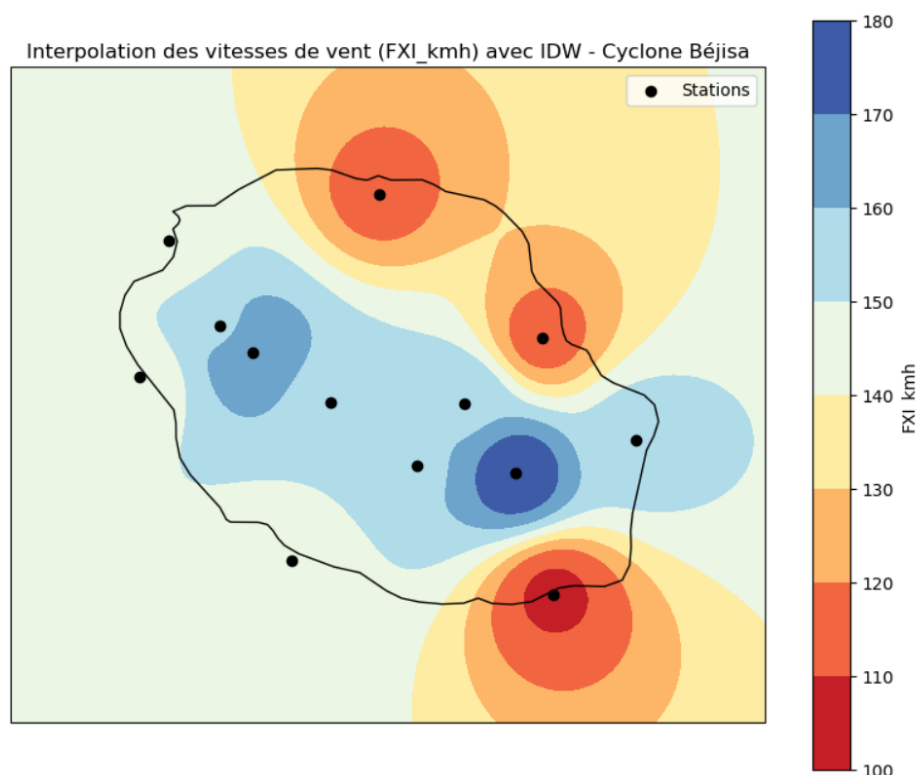


Figure 21 : Interpolation des rafales de vents lors du cyclone Béjisa grâce à une interpolation IDW. Carte réalisée sur python à l'aide de la librairie ArcPy.

Une fois les vitesses de rafales interpolées pour chaque point de La Réunion, elles ont été regroupées à l'échelle communale afin de correspondre aux données de sinistralité et d'exposition. Cette agrégation a été réalisée à l'aide d'une jointure géospatiale³⁵ sous python avec le package ArcPy. La moyenne des vitesses de vent par commune a été privilégiée, car elle offre une représentation plus fidèle de l'exposition réelle des assurés, contrairement aux valeurs maximales souvent observées dans des zones peu habitées.

³⁵ Une jointure géospatiale est une méthode utilisée pour associer des données à des entités géographiques en fonction de leur position dans l'espace. Dans ce cas, elle permet de relier les vitesses de vent calculées pour chaque point de La Réunion aux communes correspondantes, en fonction de la localisation de ces points à l'intérieur des limites géographiques des communes.

Traitements des données de cumuls de précipitations

Contrairement aux données de vent, les précipitations disposent quant à elles d'un réseau de stations suffisamment dense pour ne pas nécessiter d'interpolation.

Carte des stations météorologiques ayant enregistré des cumuls de pluie lors des cyclones à La Réunion



Figure 22 : Carte des stations météorologiques de La Réunion ayant enregistré des cumuls de pluie lors de cyclones. Réalisé sous python grâce au package ArcPy.

Pour chaque événement cyclonique, les cumuls de précipitations ont été calculés sur toute la durée de l'événement, en sommant les enregistrements quotidiens. Cette approche est cohérente avec les pratiques de reconnaissance CatNat et permet de capturer l'impact total des précipitations lors d'un cyclone.

Les précipitations ont ensuite été agrégées à la maille communale en retenant le cumul maximal enregistré par commune. Ce choix méthodologique repose sur plusieurs considérations. Tout d'abord, les stations situées en altitude, susceptibles de surestimer les précipitations pour certaines zones, ont été exclues, ce qui rend le cumul maximal par commune plus représentatif des conditions extrêmes réellement subies par les assurés. Ensuite, bien que l'identification précise du péril à l'origine des sinistres ne soit pas systématiquement possible pour les événements historiques, des échanges avec Lionel LAFFIN, expert construction chez Saretec³⁶, ont mis en évidence que les infiltrations liées aux précipitations représentent une part significative des dommages survenus ces dernières années à La Réunion.

Ainsi, adopter une approche « sévère » pour les précipitations, en retenant les conditions les plus pénalisantes par commune, permet de mieux refléter la réalité des risques cycloniques spécifiques au portefeuille dommages aux biens de Prudence Créole. Cette démarche ajuste les analyses aux particularités du contexte local et renforce la pertinence des résultats obtenus.

Validation et intégration finale

Pour assurer la fiabilité des données météorologiques retravaillées, une validation a été effectuée en comparant les résultats des retraitements aux publications officielles de Météo France et de la CCR, en particulier pour les événements cycloniques majeurs dont la documentation est facilement accessible. Pour les cyclones moins médiatisés ou ayant généré peu de documentation

³⁶ Saretec est une société spécialisée dans l'expertise après sinistre. Elle accompagne les assureurs et les entreprises en offrant des services d'analyse, de conseil, et d'intervention sur le terrain pour évaluer les dommages et proposer des solutions adaptées.

officielle, des vérifications ont été effectuées à partir de checks logiques, comme la cohérence entre les vitesses de vent, les dommages constatés et le niveau de médiatisation de l'événement.

En illustration, les données de vent interpolées et les cumuls de précipitations sont comparées aux observations publiées par Météo France pour le cyclone Batsirai. Cette comparaison met en évidence que les données climatiques retravaillées sont cohérentes avec celles des cartes officielles. Ce résultat valide la qualité de notre méthodologie, en confirmant que les données retravaillées reflètent fidèlement les conditions réelles observées sur le terrain.

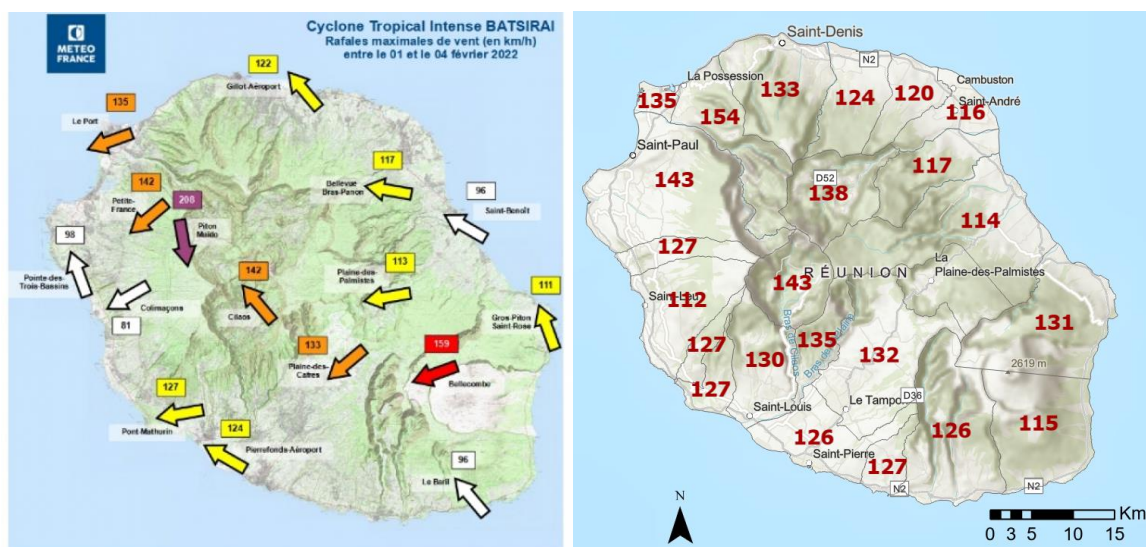


Figure 23 : Comparaison des vitesses de rafales enregistrées lors du cyclone Batsirai : carte de Météo France (à gauche) et des données interpolées (à droite). Carte de droite réalisée avec ArcGIS.

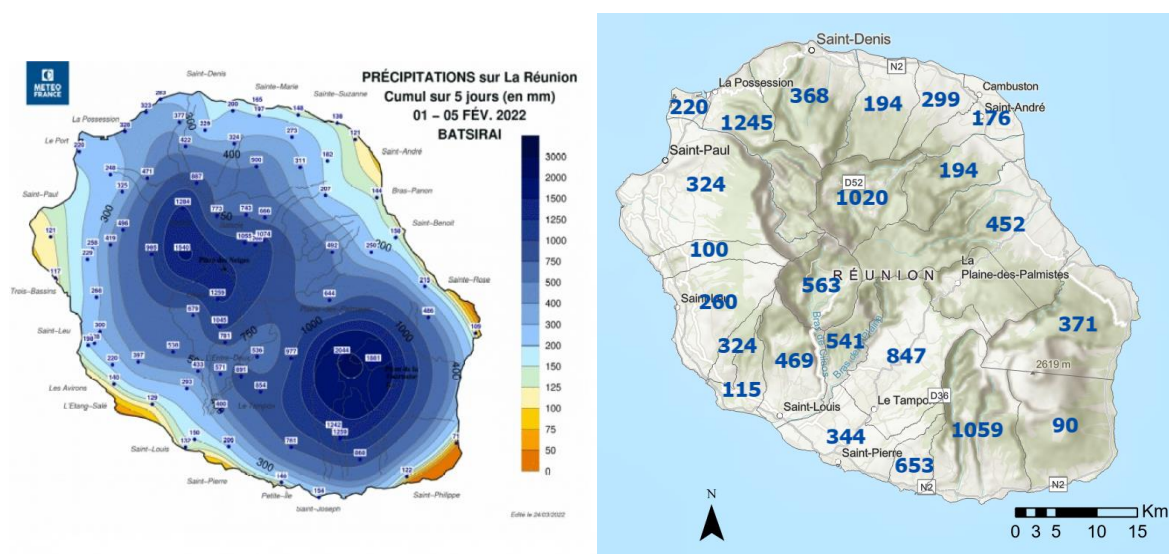


Figure 24 : Comparaison des cumuls de précipitations enregistrées lors du cyclone Batsirai : carte de Météo France (à gauche) et données de notre base (à droite). Carte de droite réalisée avec ArcGIS.

Une fois toutes les données traitées, une base consolidée incluant les données de sinistres, de portefeuille et météorologiques a été créée pour la modélisation du risque cyclonique à La Réunion. Cette base est à l'échelle communale, car c'est la plus fine disponible pour les sinistres.

Il reste important de noter que les données de portefeuille sont géolocalisées à l'adresse et que, grâce au travail d'interpolation des données météorologiques réalisé dans le cadre du mémoire, les données météorologiques peuvent également être géolocalisées à l'adresse si nécessaire. Si le modèle

devait évoluer vers une échelle plus précise, les travaux entrepris dans le cadre du mémoire faciliteraient cette transition, notamment grâce à l'amélioration significative de la qualité des données sur les informations relatives à la structure des bâtiments.

Chapitre 3 : Présentation du modèle catastrophe simplifié

3.1. Les enjeux ayant conduit à l'élaboration d'un modèle CAT simplifié

Comme des données météorologiques, de sinistres et de portefeuille sont disponibles, la première piste de modélisation étudiée était d'utiliser un modèle CAT (modèle de catastrophe). Un modèle CAT est un outil informatique utilisé pour estimer les pertes potentielles causées par des événements catastrophiques tels que les cyclones, les tremblements de terre, et autres catastrophes. Ces modèles sont particulièrement utiles dans le secteur de l'assurance pour analyser les risques. Ils utilisent des données historiques et des simulations pour prédire les impacts financiers de ces événements. Ils prennent en compte plusieurs facteurs, notamment :

- Les caractéristiques des catastrophes : par exemple, la force d'un cyclone ou la magnitude d'un tremblement de terre.
- Les caractéristiques des biens assurés : comme le type de construction, l'emplacement géographique, et l'utilisation des bâtiments.
- Les conditions financières des polices d'assurance : telles que les limites de couverture et les franchises.

Cependant, leur mise en œuvre nécessite des ressources considérables et un temps de développement incompatible avec les contraintes de ce mémoire.

Face à ces limitations, une approche pragmatique a été privilégiée, avec un focus sur la fiabilisation des données, un travail long mais essentiel pour garantir des résultats pertinents. La qualité des données constitue en effet le socle de toute modélisation, selon le principe du « *garbage in, garbage out* » : des données fiables sont essentielles pour produire des estimations robustes. Ce choix méthodologique a permis de concevoir un modèle CAT simplifié, alliant efficacité et clarté, et adapté aux besoins opérationnels de Prudence Créole, tout en posant des bases solides pour un modèle CAT plus détaillé si le besoin d'en fait ressentir.

Prudence Créole continuera d'utiliser la méthode des triangles pour fournir une évaluation initiale des sinistres à J+10, mais pourra également utiliser le modèle pour communiquer des premières estimations à J+3 et avoir une idée de l'ordre de grandeur de l'événement lors de la construction du triangle, ce qui sera beaucoup plus rassurant que de le faire sans filet de sécurité.

C'est dans ce contexte de faible niveau de granularité des données qu'émerge l'élaboration d'un modèle CAT simplifié, conçu pour répondre efficacement aux exigences tout en tenant compte des contraintes identifiées.

Par ailleurs, les réassureurs et les membres des instances (comité exécutif, etc.) demandent des réponses rapides et facilement compréhensibles, ce qui limitent le recours à des seuls modèles jugés trop complexes, souvent perçus comme des « boîtes noires ». Un modèle CAT, même simplifié, pourrait être perçu comme complexe à comprendre pour des personnes n'ayant pas l'habitude de tels outils. Dans ce contexte de besoin de vulgarisation, en amont du modèle CAT simplifié, des indicateurs clairs et simples, tels que le pourcentage d'assurés potentiellement impactés ou les seuils météorologiques seront également élaborés. Cela permettra d'identifier rapidement si un événement sera couvert par la réassurance. Ces informations compléteront l'expérience terrain de Prudence Créole et serviront de base à une communication compréhensible avant de présenter le modèle CAT simplifié.

3.2. Fonctionnement et application d'un modèle « CAT » simplifié

Cette section décrit le fonctionnement du modèle CAT simplifié développé. Ce modèle repose sur quatre modules distincts : aléa, vulnérabilité, dommages bruts et dommages nets de réassurance CCR. Chacun des modules utilise les données déjà décrites dans les sections précédentes – météorologiques, de sinistralité historique et d'exposition – pour fournir une estimation rapide, fiable et compréhensible des pertes potentielles.

Le graphique ci-dessous offre une vue d'ensemble du processus de modélisation.

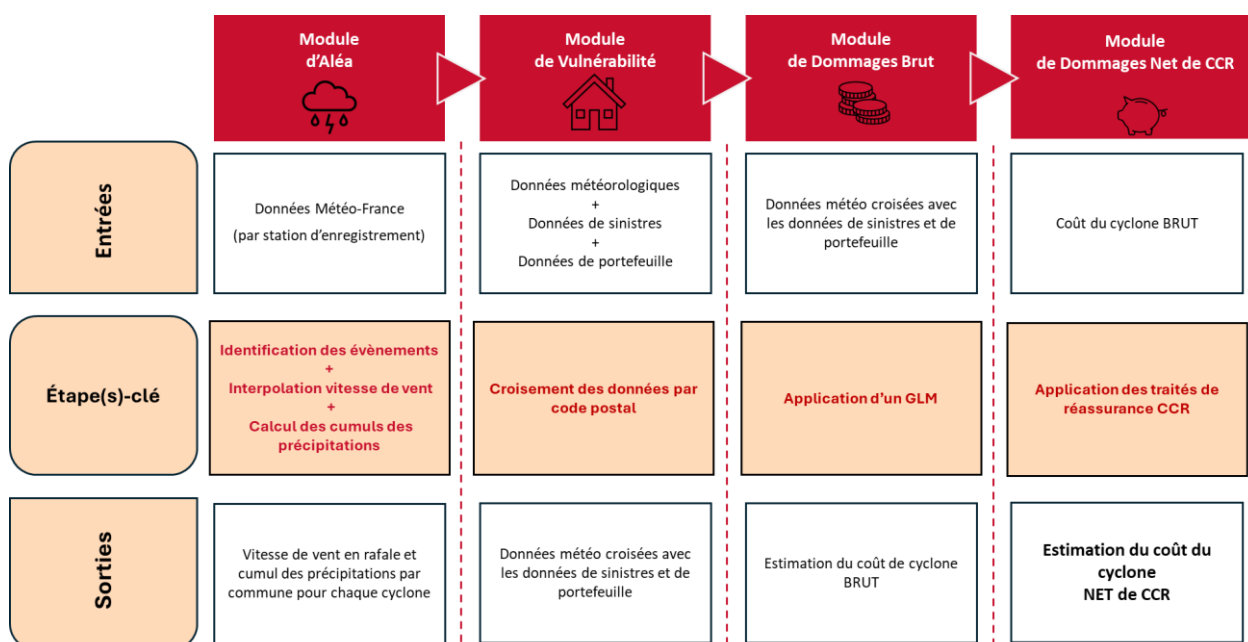


Figure 25 : Schéma du modèle CAT simplifié pour l'estimation du coût d'un cyclone. Réalisé avec PowerPoint.

Les quatre modules sont expliqués plus en détail dans les sous-sections suivantes.

Module d'Aléa

Le premier module, dit "d'aléa", repose sur des données historiques météorologiques fournies par Météo France, collectées via des stations d'enregistrement. Ce module est consacré à l'analyse des paramètres météorologiques associés à un cyclone, notamment la vitesse des vents et le cumul des précipitations. Les étapes essentielles sont les suivantes :

1. Identification des événements : Détection des cyclones basés sur la documentation marché et la sinistralité historique de Prudence Créole (nom des cyclones et dates associées). L'ensemble des cyclones majeurs de La Réunion depuis 2002 est bien listé dans les bases des sinistres de Prudence Créole.
2. Interpolation des vitesses de vent : Utilisation des données stationnaires pour modéliser la vitesse des vents à l'échelle communale.
3. Calcul des cumuls de précipitations : Agrégation des précipitations pour chaque commune touchée par le cyclone.

En cas de nouveau cyclone, les données utilisées seront disponibles dans la rubrique « observation³⁷ » de Météo France mais aussi dans leur publication sur les événements cycloniques quelques jours après l'événement.

³⁷ <https://meteofrance.re/fr/observations>

Les sorties de ce module incluent les vitesses de vent en rafales et les cumuls de précipitations estimés pour chaque cyclone à l'échelle communale, constituant ainsi les variables explicatives pour les modules suivants.

Module de Vulnérabilité

Le deuxième module s'intéresse à la vulnérabilité des biens assurés face aux aléas identifiés. Contrairement aux autres modules, il repose davantage sur la structuration des données que sur des concepts techniques complexes.

Ce module combine les données météorologiques, historiques de sinistralité et de portefeuille pour établir une base consolidée. Pour chaque cyclone, chaque commune est associée à des variables telles que la sinistralité historique, les caractéristiques des bâtiments et les conditions météorologiques spécifiques aux événements cycloniques. Ce croisement par les communes permet d'établir des relations entre l'intensité des aléas et les dommages observés, facilitant ainsi la modélisation des impacts futurs et permettant à Prudence Créole d'établir des indicateurs simples pour déterminer si un cyclone sera réassuré ou non (cf. Chapitre 1 : Identification des cyclones nécessitant un suivi par le service de réassurance).

En cas de nouveau cyclone, il faudra mettre à jour les données de portefeuille et de sinistralité si une mise à jour est disponible. Une visualisation de la base à cette étape est possible en Annexe 3.

Module de Dommages Bruts

Le troisième module a pour but, via une modélisation statistique, d'estimer les coûts bruts d'un cyclone. En s'appuyant sur les données issues des deux modules précédents, il prédit les charges financières associées à un événement en fonction des variables explicatives comme la vitesse des vents, le cumul des précipitations et les caractéristiques des biens assurés.

Le choix s'est porté sur un modèle linéaire généralisé (GLM). Ce type de modèle est adapté aux données disponibles et permet une estimation simple et rapide tout en restant suffisamment robuste pour répondre aux besoins des réassureurs³⁸.

Ce module fournit une estimation du coût brut d'un cyclone, qui correspond à la somme totale des dommages sans tenir compte de la réassurance.

Module de Dommages Nets de CCR

Enfin, le dernier module applique les cessions de réassurance CCR pour estimer le coût net supporté par Prudence Créole et Generali France consolidé. Ce module intègre les spécificités des traités de réassurance et utilise une méthodologie basée sur l'expérience métier pour estimer la reconnaissance CatNat. En effet, à La Réunion, le statut de petite île joue un rôle particulier dans la reconnaissance CatNat pour le vent. D'un point de vue politique et opérationnel, il paraît difficilement envisageable de reconnaître uniquement une partie de l'île en CatNat pour le vent. Par conséquent, deux scénarios distincts se dessinent : Non reconnaissance CatNat ou reconnaissance CatNat pour toute l'île.

Si très peu de stations dépassent le critère de 215 km/h en rafales, aucune reconnaissance CatNat pour le vent n'est accordée. Dans ce cas, le taux de cession appliqué repose sur le taux moyen historique observé (hors Dina), une période où aucun arrêté CatNat pour le vent n'avait été émis et où les arrêtés se limitaient principalement aux inondations, avec une sinistralité CatNat relativement faible.

³⁸ Voir la partie 3, Chapitre 2 : Estimation de la charge d'un cyclone réassuré.

Si une majorité de stations dépasse le critère des 215 km/h, une reconnaissance CatNat pour l'ensemble de l'île est accordée, l'ensemble des sinistres est considéré comme CatNat et éligible aux traités de réassurance CCR. Il est cependant important de noter que le modèle n'est pas calibré pour de tels événements où l'ensemble de l'île serait reconnu en CatNat, comme ce fut le cas pour le cyclone Dina. Pour ce type de sinistres extrêmes, il est recommandé d'effectuer une analyse spécifique en appliquant les taux de destruction observés lors de Dina. Le taux de destruction est un indicateur qui mesure la proportion des sinistres par rapport à la valeur totale assurée. Ce mémoire a d'ailleurs facilité cette tâche en rassemblant des données historiques détaillées sur cet événement, comblant une lacune importante puisque, avant ce travail, le seul autre événement majeur CatNat outre-mer correctement documenté chez Generali était l'ouragan Irma, dont les taux de destruction ne sont pas adaptés aux spécificités de La Réunion (cf. Partie 1 : Cadre de l'étude et contexte). Une fois ce taux appliqué, la charge serait rehaussée de 6,5% pour prendre en compte les effets de la réforme CatNat qui n'existaient pas encore en 2002 (prise en compte des frais de relogement d'urgence dans la garantie CatNat).

Ainsi, les sorties de ce module représentent les coûts nets de CCR supportés par Prudence Créole et par Generali France (si les coûts nets de CCR sont supérieurs à la priorité du traité de réassurance entre Prudence Créole et Generali France).

3.3. Limites et avantages d'un modèle CAT simplifié

Un modèle CAT simplifié offre des avantages notables en termes de rapidité et de simplicité, mais présente également certaines limites. Sa dépendance aux données historiques peut restreindre sa capacité à refléter les évolutions actuelles ou futures. Par exemple, des phénomènes comme le changement climatique, qui intensifie la sévérité des cyclones, ou des transformations socio-économiques, telles que l'urbanisation ou l'amélioration des normes de construction, ne sont pas pleinement intégrés. La mise en AS-IF des sinistres (cf. Partie 2, Chapitre 2 : Les données disponibles et leurs pré-traitements) répond partiellement à ces enjeux, notamment en tenant compte de l'évolution du nombre de risques assurés par Prudence Créole à La Réunion.

De plus, la sensibilité aux données historiques est particulièrement importante pour les événements extrêmes. Ces derniers, bien que rares, jouent un rôle clé dans le calcul des coûts de réassurance. Cependant, leur faible fréquence limite la précision de leurs projections, reflétant une contrainte inhérente à la modélisation des phénomènes rares en actuariat.

Contrairement aux modèles CAT classiques, qui utilisent des courbes de vulnérabilité pour établir un lien précis entre l'intensité des aléas et les dommages, le modèle simplifié construit s'appuie sur un GLM. Cette approche exploite directement les données de portefeuille, de sinistralité et météorologiques, garantissant des résultats pertinents tout en réduisant la complexité. Ce choix méthodologique a permis d'accélérer le développement du modèle, le rendant immédiatement opérationnel si un cyclone à La Réunion devait se produire pour la saison 2024-2025. Bien que les courbes de vulnérabilité soient absentes, elles pourraient être intégrées ultérieurement, le cadre posé dans ce mémoire offrant une base solide pour de futures évolutions.

Par ailleurs, le modèle ne repose pas sur des paramètres physiques complexes, tels que les équations de Navier-Stokes³⁹ ou l'effet Coriolis⁴⁰, fréquemment utilisés dans les modèles CAT classiques pour simuler les trajectoires cycloniques et les flux atmosphériques.

³⁹ « L'objectif des équations de Navier-Stokes est de décrire le mouvement des fluides. Il est possible de l'utiliser par exemple pour comprendre les mouvements des courants dans les océans, ainsi que ceux des grandes masses d'air dans l'atmosphère ». Source : Science Etonnante.

⁴⁰ « Force produite par l'accélération complémentaire due à la rotation terrestre et qui s'exerce sur tous les corps en mouvement à la surface de la Terre. (Elle détermine la direction générale des vents et des courants océaniques, les déviant vers la droite dans l'hémisphère Nord, vers la gauche dans l'hémisphère Sud) ». Source Larousse.

Malgré ces simplifications, le modèle répond efficacement aux besoins opérationnels grâce à sa rapidité, sa transparence et sa flexibilité. Sa conception modulaire, intégrant les aléas, les sinistres et les paramètres de réassurance, assure une estimation cohérente et fiable des coûts cycloniques. En comparaison, les modèles CAT classiques, bien que plus robustes, nécessitent des bases de données détaillées, des simulations stochastiques complexes et des ressources techniques importantes.

Enfin, cette simplicité opérationnelle facilite la communication des résultats aux réassureurs et aux parties prenantes, tout en offrant une précision adaptée aux analyses stratégiques et opérationnelles.

Partie 3 : L'estimation des pertes associées à la survenance des cyclones à La Réunion

Pour rappel, ce mémoire s'inscrit dans une démarche visant à répondre aux deux questions suivantes :

- 1- Comment identifier rapidement, si un cyclone impactera Prudence Créole de sorte à déclencher un traité de réassurance ? Plus précisément, il s'agit de déterminer les critères permettant d'évaluer si le coût d'un événement dépassera 75 % de la priorité du traité de réassurance DAB (120 xs 5), ce seuil correspondant généralement à celui à partir duquel les événements sont déclarés aux réassureurs.
- 2- Comment estimer les coûts engendrés par ces événements pour Prudence Créole et Generali France ?

Chapitre 1 : Identification des cyclones nécessitant un suivi par le service de réassurance

L'objectif de ce chapitre est de déterminer si le coût d'un cyclone dépasse le seuil de déclaration auprès des réassureurs. La démarche se déroule en deux étapes : d'abord, vérifier si les données météorologiques indiquent une intensité élevée ; ensuite, analyser le pourcentage des sommes assurées affectées par des conditions météorologiques intenses. Si ce pourcentage est suffisamment élevé, l'événement dépasse le seuil de déclaration ; sinon, il ne le dépasse pas.

L'analyse doit être rapide et précise en s'appuyant sur les données disponibles dans les heures suivant le passage d'un cyclone. Pour cela, la première partie de la démarche s'appuie sur des données provenant des relevés de Météo France. Ces informations, compte tenu de la nature des sinistres cycloniques, concernent les vitesses de vent et les cumuls de précipitations mesurés par les stations météorologiques.

Pour exploiter ces informations, deux approches théoriques sont envisageables : une analyse conjointe ou une analyse distincte des deux aléas. Afin de garantir une application intuitive pour les acteurs métier, il a été décidé de définir des seuils séparés pour chaque aléa. Ce choix simplifie la complexité inhérente au risque cyclonique en la traduisant sous forme d'indicateurs pratiques et directement exploitables, tout en permettant une anticipation efficace des coûts associés aux événements majeurs.

L'analyse est basée sur les données météorologiques des cyclones historiques ayant impacté La Réunion, mis à part le cyclone Belal de 2024 qui est utilisé en tant que *back-test*⁴¹ pour valider les résultats. Les données de vent et de précipitations, traitées et préparées dans la partie 2, servent de base à la construction d'indicateurs fiables et cohérents, adaptés à l'objectif de cette étude.

⁴¹ Le *back-testing* est une technique utilisée pour évaluer la performance d'un modèle en l'appliquant à des données historiques.

1.1. Distribution théorique des excès et détection de seuil

Comme l'enjeu est de définir un seuil à partir duquel les cumuls de précipitations et les vitesses de rafales peuvent être considérés comme représentatifs de phénomènes cycloniques intenses, l'approche *Peaks-Over-Threshold (POT)* de la Théorie des Valeurs Extrêmes a été utilisée. La méthode POT ou méthode des excès au-delà d'un seuil, consiste à observer les valeurs des observations qui dépassent un certain seuil afin de modéliser les excès associés aux événements les plus marquants. Cette approche repose sur l'idée que ces observations au-delà d'un seuil suivent une loi de Pareto généralisée (*GPD*). Dans ce mémoire, l'analyse portera uniquement sur la sélection de ce seuil critique, la modélisation du coût d'un cyclone étant abordée à l'aide d'un modèle linéaire généralisé dans le chapitre 2.

Dans cette section, les principes théoriques exposés s'appuient sur l'ouvrage « *An Introduction to Statistical Modeling of Extreme Values* » de Stuart Coles (2001) et le cours « *Extreme Value Theory* » dispensé par Christian Y. Robert en dernière année de l'ENSAE (2024).

Soit X une variable aléatoire modélisant le phénomène étudié (par exemple, les rafales de vent ou les cumuls de précipitations). X suit une loi de probabilité inconnue, définie par une fonction de répartition F . Afin d'étudier les propriétés de X , un échantillon de taille n , constitué de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) est notées $X_1, X_2, \dots, X_n$, où chaque X_i est une réalisation de la variable aléatoire X .

Pour se concentrer sur les observations extrêmes, un seuil p est fixé tel que $p < x_F$, où x_F est la borne supérieure des données. Le nombre d'observation dépassant le seuil p est noté N_p .

Dans le contexte de l'approche POT, les excès sont définis comme $X - p$. Cela signifie qu'au lieu d'analyser les valeurs X_i , seules les parties des X_i qui dépassent p sont prises en compte. Cela permet de se focaliser uniquement sur les comportements extrêmes, sans tenir compte des valeurs inférieures au seuil.

Définition : Distribution conditionnelle des excès

La distribution conditionnelle des excès au-delà de p est définie par la fonction de répartition de $X - p \mid X > p$, notée F_p :

$$F_p(x) = P(X - p < x \mid X > p) = \frac{F(p + x) - F(p)}{1 - F(p)}, \quad 0 \leq x < \infty$$

La fonction de répartition conditionnelle des excès permet de décrire la distribution des valeurs dépassant un seuil p , en isolant les observations extrêmes.

Toutefois, pour modéliser précisément ces excès, il est nécessaire de s'appuyer sur une distribution adaptée aux queues de distribution. C'est dans ce contexte que la loi de Pareto généralisée intervient.

Définition : Loi de Pareto généralisée (GPD)

La fonction de répartition de la loi de Pareto généralisée de paramètres de forme ξ et d'échelle β est caractérisée par :

$$G_{\xi, \beta} = \begin{cases} 1 - \left[1 + \xi \left(\frac{x}{\beta} \right) \right]_+^{-\frac{1}{\xi}} & \text{if } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{if } \xi = 0. \end{cases}$$

Où

- $x \geq 0$ si $\xi \geq 0$
- $0 \leq x \leq -\frac{\beta}{\xi}$ pour $\xi < 0$

La distribution de Pareto généralisée repose sur deux paramètres clés : le paramètre d'échelle (β), qui mesure la dispersion, et l'indice de queue (ξ), qui détermine la forme et la décroissance des queues. Selon la valeur de ξ , les queues peuvent être bornées ($\xi < 0$), décroître exponentiellement ($\xi = 0$), ou être épaisses⁴² ($\xi > 0$). Dans le contexte des événements météorologiques extrêmes, le paramètre de forme (ξ) attire particulièrement l'attention.

La relation entre les excès au-delà d'un seuil dans une distribution et leur ajustement à une loi de Pareto généralisée est formalisé par le théorème de Pickands-Balkema-de Haan. Le théorème est énoncé comme suit.

Théorème de Pickands-Balkema-de Haan (1975)

Soit F la fonction de répartition d'une variable aléatoire X , définie sur un intervalle borné par x_F . Si F appartient à l'un des trois domaines d'attraction des extrêmes (Fréchet, Gumbel ou Weibull⁴³), alors pour un seuil p suffisamment élevé, la fonction de répartition conditionnelle des excès au-delà de p , notée $F_p(x)$, converge vers une loi de Pareto généralisée $G_{\xi, \beta(p)}$, où ξ est le paramètre de forme et $\beta(p)$ le paramètre d'échelle. Cette convergence s'exprime comme suit :

$$\sup_{0 \leq x \leq x_F} |F_p(x) - G_{\xi, \beta(p)}(x)| \rightarrow 0 \quad \text{lorsque } p \rightarrow \infty.$$

Où :

- $F_p(x)$ est la fonction de répartition des excès au-delà du seuil p .
- $G_{\xi, \beta(p)}$ est la fonction de répartition de la loi de Pareto généralisée.

Ainsi, pour p suffisamment grand, la loi des excès au-delà de p peut être approximée par une loi de Pareto généralisée :

$$F_p(x) \approx G_{\xi, \beta(p)}(x)$$

Ce résultat théorique justifie l'utilisation de la GPD pour modéliser les valeurs extrêmes et guide la sélection d'un seuil cohérent avec la méthodologie *Peaks-Over-Threshold*. Toutefois, la détermination précise de ce seuil nécessite des outils supplémentaires qui s'appuient sur une propriété essentielle de la GPD : la propriété de stabilité.

Propriété de stabilité

Si une variable aléatoire X suit une loi $GPD(\xi, \beta)$, alors pour tout seuil $p > 0$, la variable conditionnelle $X - p \mid X > p$, notée $[X - p \mid X > p]$, suit également une loi $GPD(\xi, \beta + \xi p)$.

La propriété de stabilité est importante pour justifier que, peu importe le seuil choisi (tant qu'il est suffisamment élevé), la nature des excès reste stable. Cela renforce l'idée que le seuil peut être ajusté sans compromettre la cohérence des données extrêmes.

En lien direct avec cette propriété, des outils graphiques tels que la fonction moyenne des excès, le *Hill-Plot* et le *Pickands-Plot* permettent d'explorer la stabilité et de déterminer un seuil à partir duquel les observations peuvent être considérées comme extrêmes.

⁴² Contrairement aux lois classiques comme la normale, les queues lourdes (ou queues épaisses) diminuent plus lentement, reflétant ainsi une fréquence accrue d'événements rares et importants. Source : Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer.

⁴³ Définitions en annexe 5.

Fonction Moyenne des Excès (FME)

Considérons $X_1, X_2, \dots, X_n$, des variables aléatoires indépendantes et identiquement distribuées, suivant une fonction de répartition F . La fonction moyenne des excès est définie comme étant la moyenne des écarts au-dessus d'un seuil élevé, divisée par le nombre d'observations excédant ce seuil. Son expression empirique s'écrit comme suit :

$$e_n(p) = \frac{1}{N_p} \sum (X_i - p), \quad p > 0$$

Où N_p représente le nombre de valeurs dépassant le seuil p .

Pour une variable aléatoire X suivant une distribution de Pareto généralisée (ξ, β) , l'espérance conditionnelle de l'excès au-delà du seuil p est donnée par :

$$e_n(p) = E(X - p | X > p) = \frac{\beta + \xi p}{1 - \xi}$$

Où

- $0 \leq p < \infty$
- $0 \leq \xi < 1$. Si $\xi < 0$, alors $X - p \leq -\beta / \xi$.

La courbe associée à cette moyenne des excès est définie pour $p \in [0, X_{(n,n)}]$ comme suit :

$$\left((p, e_n(p)), X_{(n,n)} < p < X_{(1:n)} \right),$$

Où $X_{(1:n)}$ et $X_{(n,n)}$ désignent respectivement les maximums et minimums des valeurs de l'échantillon.

L'approche graphique consiste à utiliser la version empirique de cette fonction, en traçant les points $(p, \widehat{e}_n(p))$ et en sélectionnant un seuil p de manière que $\widehat{e}_n(p)$ soit approximativement linéaire après p . Visuellement, cela se manifeste par un changement de la pente de la courbe qui reste ensuite stable.

De plus, l'interprétation de la pente de la FME révèle la nature des excès :

- Une pente positive ($\xi > 0$) indique une loi de Pareto généralisée associée à des queues épaisses.
- Une pente nulle ($\xi = 0$) correspond à une loi exponentielle.
- Une pente négative ($\xi < 0$) reflète une distribution à queue fine⁴⁴, où les valeurs extrêmes sont limitées.

Bien que la fonction moyenne des excès constitue un outil graphique puissant pour identifier un seuil, son interprétation peut parfois être délicate. Les valeurs extrêmes de l'échantillon, souvent instables, compliquent l'identification précise d'un point critique. Pour surmonter cette difficulté, il est courant de s'appuyer sur d'autres analyses visuelles telles que les graphes de Hill et de Pickands.

Les graphes des estimateurs de queue

L'idée principale consiste à représenter graphiquement les estimateurs de ξ en fonction des différents seuils p , ou de manière équivalente, en fonction du nombre N_p d'excès retenus.

1- Estimateur de Hill

L'estimateur de Hill est calculé comme suit :

⁴⁴ Les distributions à queue fine se caractérisent par une décroissance rapide de la probabilité des valeurs extrêmes. Cela signifie que les événements très éloignés de la moyenne, ou extrêmes, sont beaucoup moins probables par rapport à une distribution à queue lourde.

$$\widehat{\xi_{n,N_p}^{\text{Hill}}} = \frac{1}{N_p} \sum_{i=1}^{N_p} \ln(X_{n-i+1:n}) - \ln(X_{n-N_p:n})$$

Où

- N_p représente le nombre d'observations excédant le seuil p ,
- $X_{n-i+1:n}$ sont les i -ème plus grandes valeurs triées.

2- Estimateur de Pickands

L'estimateur de Pickands, également adapté pour un seuil p , est donné par :

$$\widehat{\xi_{n,N_p}^{\text{Pickands}}} = \frac{1}{\ln 2} \ln \left(\frac{X_{n-N_p:n} - X_{n-2N_p:n}}{X_{n-2N_p:n} - X_{n-4N_p:n}} \right)$$

Où les valeurs $X_{n-N_p:n}, X_{n-2N_p:n}, X_{n-4N_p:n}$ représentent les plus grandes observations triées.

Ces outils graphiques permettent de visualiser les variations de l'estimation de ξ en fonction des paramètres N_p ou p . En pratique, pour ces deux méthodes, il faut identifier une zone où l'indice de queue ξ reste relativement constant, appelée zone de stabilité (palier horizontal). Le seuil optimal correspond alors au début de cette zone, et le nombre d'excès conservés est déterminé en conséquence.

Il est important de noter que ces deux estimateurs sont principalement adaptés aux distributions à queues épaisses, c'est-à-dire celles dont l'indice de queue $\xi > 0$. Cela limite leur applicabilité aux distributions telles que les lois normales ou à queues fines ($\xi < 0$), où ces estimateurs ne sont pas pertinents.

Le graphe de stabilité des paramètres

Cet outil graphique est utilisé pour déterminer un seuil optimal dans le cadre de l'ajustement des données à une loi de Pareto généralisée. Ce graphe repose sur la représentation des paramètres estimés, le paramètre de forme (ξ) et le paramètre d'échelle (β), en fonction de différents seuils p . L'objectif est d'observer si ces paramètres présentent une stabilité dans une certaine plage de seuils.

Cependant, il est important de noter que l'efficacité du graphe de stabilité dépend fortement des données. Dans certains cas, notamment avec des échantillons très limités ou des distributions particulières, les graphes de stabilité peuvent ne pas fournir d'informations concluantes. Néanmoins, lorsqu'ils sont pertinents, ils constituent un complément précieux aux autres méthodes graphiques comme la fonction moyenne des excès ou le *Hill-Plot*.

Après avoir utilisé les graphes des estimateurs de queue pour identifier un seuil optimal, il est possible d'envisager une validation statistique avec le test de Villasenor-Alva et Gonzalez-Estrada.

Test de Villasenor-Alva et Gonzalez-Estrada

Cette méthode statistique évalue l'ajustement des excès au-delà d'un seuil p à une loi GPD. Contrairement au test de Kolmogorov-Smirnov, qui évalue l'adéquation globale d'un échantillon à une loi, ce test se concentre spécifiquement sur les queues de distribution, ce qui le rend particulièrement adapté à l'étude des valeurs extrêmes. Le test repose sur les deux hypothèses suivantes :

- H_0 : L'échantillon suit une GPD avec un paramètre de forme $\xi \geq 0$ (queues épaisses).
- H_1 : L'échantillon suit une GPD avec un paramètre de forme $\xi < 0$ (queues fines).

Implémenté sous le langage R via la fonction *gpd.test* de la librairie *gPdtest*, ce test repose sur le théorème de Pickands-Balkema-de Haan, qui établit que la convergence vers une GPD est

asymptotiquement associée au seuil distinguant les conditions météorologiques ordinaires de celles intenses. Il utilise un bootstrap paramétrique⁴⁵, qui consiste à rééchantillonner les données pour estimer le paramètre de forme ξ sous chaque hypothèse. Cette approche permet de comparer les données observées à des distributions générées sous la GPD, en testant si les excès au-delà du seuil p respectent bien les propriétés attendues.

En pratique, le test peut être appliqué sur des échantillons correspondant à différents seuils candidats p . Le seuil optimal est alors défini comme le plus petit seuil p pour lequel le test reste significatif. Cette approche garantit que l'échantillon au-delà du seuil choisi est bien représentatif d'une distribution de Pareto généralisée.

En réalité la sélection du seuil est une tâche complexe. Un seuil trop bas augmente le risque d'inclure des valeurs non extrêmes, ce qui se traduit par une plus grande variance et une moindre précision dans l'ajustement des excès. À l'inverse, un seuil trop élevé réduit la taille de l'échantillon exploitable, ce qui introduit un biais en diminuant la représentativité statistique des excès réellement significatifs.

Pour résoudre cette problématique, plusieurs intervalles de valeurs possibles sont généralement identifiés pour chaque estimateur. La sélection du seuil final résulte alors d'un croisement entre ces intervalles, les contraintes métiers et les objectifs spécifiques de l'étude. Des seuils métiers, identifiés à dire d'expert, peuvent également aider pour la sélection du seuil final.

Pour évaluer la pertinence du seuil sélectionné, une approche complémentaire par QQ-Plot est utilisée.

Évaluation de l'adéquation entre données empiriques et loi théorique avec un QQ-Plot

Le QQ-Plot est un outil graphique permettant d'évaluer visuellement la qualité d'ajustement des excès à une GPD. Il trace les quantiles théoriques de la distribution ajustée en fonction des quantiles empiriques des données. La représentation graphique s'appuie sur les points suivants :

$$\left(F^{-1}\left(\frac{n-k+1}{n+1}\right), x_{(k)} \right), \quad k = 1, 2, \dots, n.$$

Où

- F^{-1} est la fonction de répartition inverse de la loi ajustée,
- $x_{(k)}$ représente les k -ième observations classées par ordre croissant.

Plus les points sont proches de la droite, plus l'ajustement est bon. Dans les queues de distribution, les points peuvent s'écarter en raison d'un manque de données, ce qui limite la précision du QQ-Plot. Ainsi, le nombre de points dans les queues influence directement la fiabilité de l'évaluation graphique.

1.2. Détermination de seuils météorologiques intenses

Le but de cette section est d'utiliser les concepts théoriques présentés précédemment pour déterminer deux seuils distincts en s'appuyant sur 384 observations de rafales et de cumuls de précipitations, préparées dans la Partie 2 : Analyse des données et méthodologie de la modélisation du risque cyclonique à La Réunion Chapitre 2 : Les données disponibles et leurs pré-traitements :

- 1- Un seuil de cumuls de précipitations intenses
- 2- Un seuil de rafales de vent intense

⁴⁵ L'idée principale du bootstrap paramétrique consiste à générer de nouveaux échantillons à partir d'un échantillon initial en supposant que ce dernier suit une loi paramétrique prédéfinie.

Dans la pratique, la détermination de seuils météorologiques intenses, constitue une tâche complexe qui ne repose pas uniquement sur des analyses statistiques rigoureuses, mais intègre également des considérations métier et des impératifs de communication. Cette approche permet de garantir que les seuils retenus soient à la fois cohérents avec les pratiques opérationnelles et compatibles avec les spécificités des données étudiées. Dans certains cas, comme pour les vitesses de vent, des seuils métiers existent, mais leur pertinence doit être validée statistiquement afin d'assurer leur adéquation avec les caractéristiques des événements analysés. Pour les cumuls de précipitations, l'absence de référence impose une démarche davantage exploratoire, fondée sur des analyses graphiques et des approches théoriques.

Seuil de cumuls de précipitations intenses

Dans le cadre de l'analyse des cyclones à La Réunion, les cumuls de précipitations observés lors d'événements majeurs tels que Dina atteignent des niveaux particulièrement élevés, justifiant l'application de la théorie des valeurs extrêmes pour modéliser ces phénomènes rares. De plus, il n'existe pas de seuil métier pouvant guider l'analyse.

1) Le Mean Excess Plot

L'analyse du *Mean Excess Plot* met en évidence une pente globalement croissante, caractéristique des distributions à queue épaisse. Cette observation est cohérente avec les records mondiaux de précipitations enregistrés sur l'île, renforçant l'idée que ces cumuls extrêmes relèvent des lois à queue lourde.

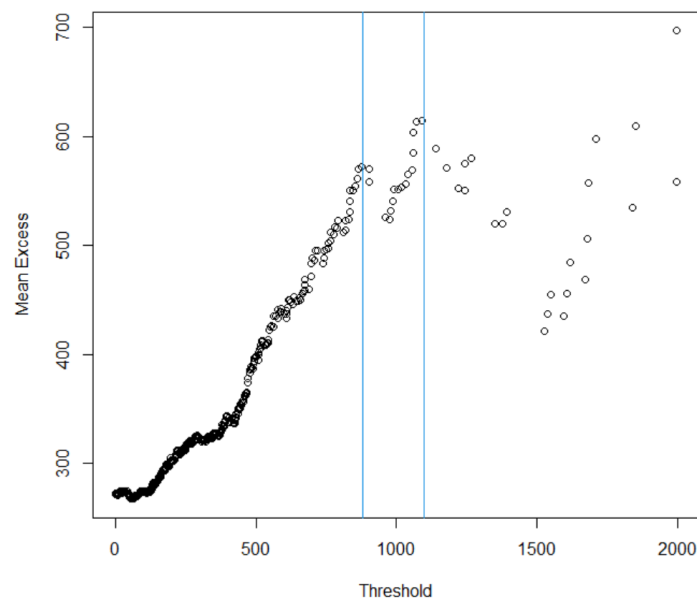


Figure 26 : Mean Excess Plot pour les cumuls de précipitations. Réalisé sous R.

Sur le *Mean Excess Plot*, deux seuils initiaux sont identifiables :

- 1100 mm (avec 27 observations au-dessus de ce seuil).
- 880 mm (avec 43 observations au-dessus de ce seuil).

Cependant, le seuil de 1100 mm a été écarté en raison du faible nombre d'observations disponibles (inférieur à 10%), rendant les estimations statistiquement fragiles. En revanche, le seuil de 880 mm a été retenu comme premier candidat pertinent, offrant un équilibre entre le nombre d'observations conservées et la stabilité des analyses.

2) Pickands

Le graphe de l'estimateur de Pickands met en évidence deux zones de stabilité :

- 1140 mm (trait rouge) : ce seuil est écarté car il ne laisse que 26 observations, un nombre insuffisant pour garantir la robustesse statistique des résultats.
- 585 mm (trait bleu) : ce seuil est jugé trop bas, avec 102 observations au-dessus, ce qui le rend incompatible avec les objectifs métier.

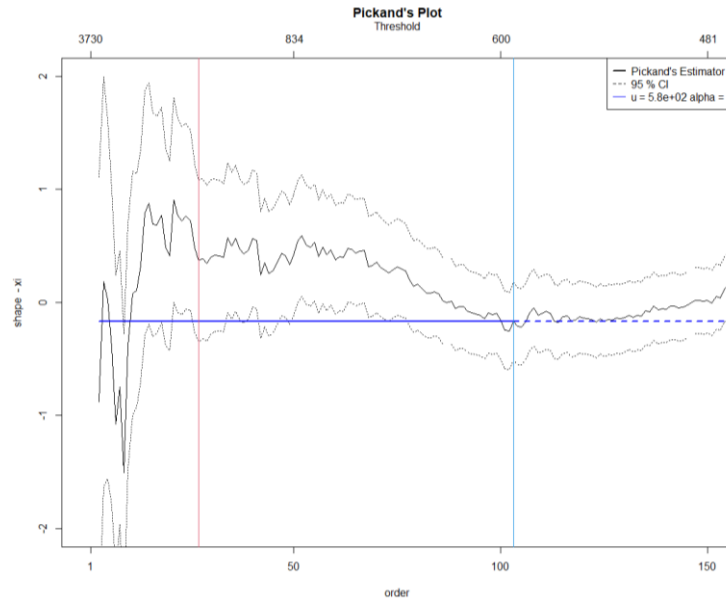


Figure 27 : Graphe de Pickand pour les cumuls de précipitations. Réalisé sous R.

3) Hills

Les résultats graphiques de l'estimateur de Hill montrent plusieurs seuils possibles :

- Un plateau horizontal est visible à partir des seuils de 902 mm ou 812 mm (traits rouges). Ces deux seuils laissent 42 et 54 observations au-dessus et sont jugés cohérents avec les contraintes métier.
- Une zone stable à partir de 585 mm (trait bleu). Ce seuil est écarté car il est trop faible (102 observations).

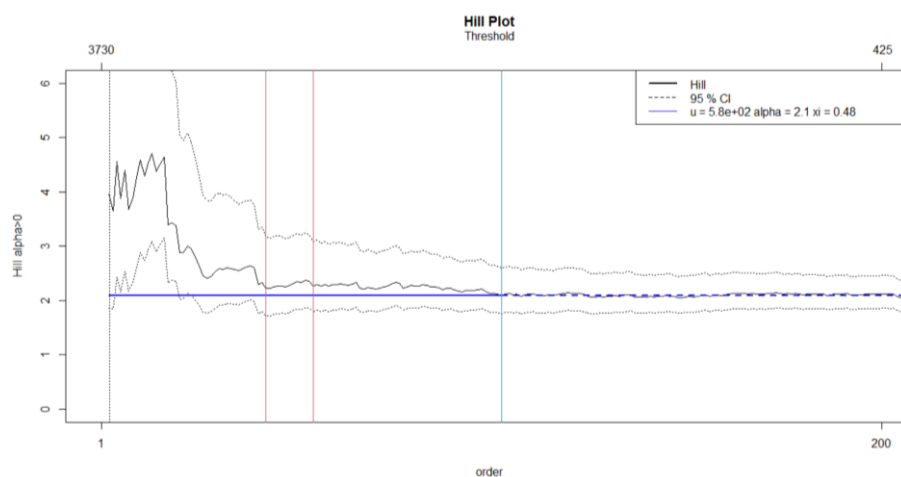


Figure 28 : Graphe de Hill pour les cumuls de précipitations. Réalisé sous R.

4) Stability Plot

Le *Stability Plot* identifie deux seuils intéressants :

- 850 mm (47 observations au-dessus) : seuil jugé robuste et cohérent.
- 1170 mm (26 observations au-dessus) : seuil exclu car insuffisant en termes d'observations.



Figure 29 : Graphique de stabilité pour les cumuls de précipitations. Réalisé sous R.

5) Villasenor-Alva et Gonzalez-Estrada

Dans le cadre du test de Villasenor-Alva et Gonzalez-Estrada, il est proposé de tracer les p-values du test en fonction des différents seuils de sévérité. Le seuil optimal peut alors être approximé comme étant le plus petit seuil pour lequel le test reste significatif, c'est-à-dire lorsque la p-value est supérieure à 5 %.

Pour nos données, la p-value dépasse définitivement 5 % à partir de 450 mm. Cependant, ce seuil est jugé trop bas (177 observations) pour capturer des valeurs réellement extrêmes et est donc écarté d'emblée.

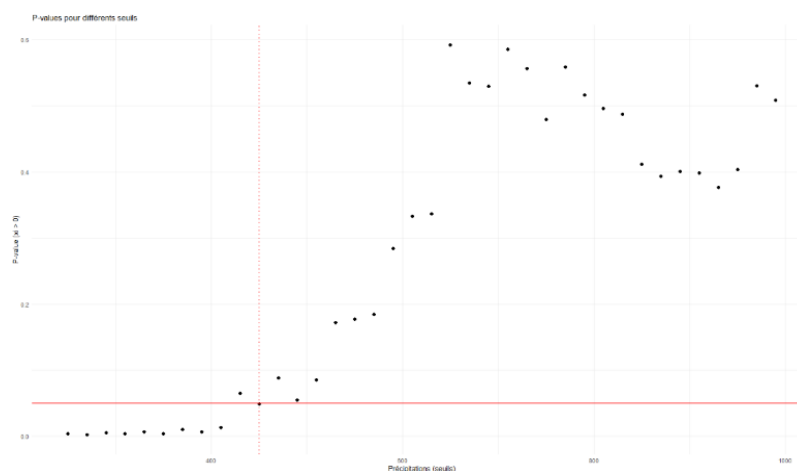


Figure 30 : Graphe des P-value du test de Villasenor-Alva et Gonzalez-Estrada en fonction des cumuls de précipitations. Réalisé sous R.

Le tableau ci-dessous récapitule les seuils candidats identifiés à partir des différentes méthodes graphiques et statistiques utilisées dans cette étude.

Techniques	Seuils candidats
MEPLOT	1100 mm → KO : trop haut 880 mm
PICKANDS	1140 mm → KO : trop haut 585 mm → KO : trop bas
HILL	902 mm 812 mm 585 mm → KO : trop bas
STABILITY PLOT	1170 mm → KO : trop haut 850 mm
VILLASENOR-ALVA et GONZALEZ-ESTRADA	450 mm → KO : trop bas

Tableau 11 : Tableau récapitulatif des seuils de cumuls de précipitations.

Le seuil de 812 mm, issu du croisement des différentes méthodes, est retenu comme le plus petit seuil mathématiquement valide, sans être jugé trop bas. Pour des raisons de communication et de vulgarisation, ce seuil est arrondi à 810 mm.

Pour valider le seuil retenu de 810 mm, un *QQ-Plot* est utilisé afin de vérifier l'adéquation des excès au-delà de ce seuil à une loi *GPD*.

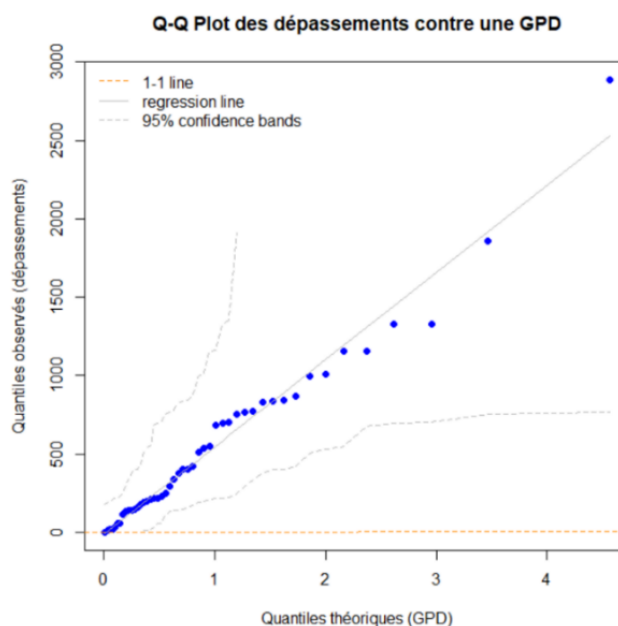


Figure 31 : QQ-Plot des dépassements de seuils (avec un seuil à 810mm) contre une GPD. Réalisé sous R.

La proximité des points avec la droite de référence confirme la qualité de l'ajustement et renforce la robustesse du seuil choisi.

Pour valider empiriquement le seuil de 810 mm, une analyse a été effectuée à partir des données historiques de sinistralité. Pour chaque événement, les données météorologiques de toutes les communes ont été examinées. Les communes ayant enregistré des précipitations supérieures à 810 mm ont été identifiées et leurs sommes assurées calculées. Un indicateur a ensuite été déterminé : le

taux des sommes assurées affectées par des précipitations supérieures à 810 mm pour chaque événement.

Il serait attendu que, pour les événements majeurs, le pourcentage de sommes assurées impactées soit plus élevé que pour les autres événements.

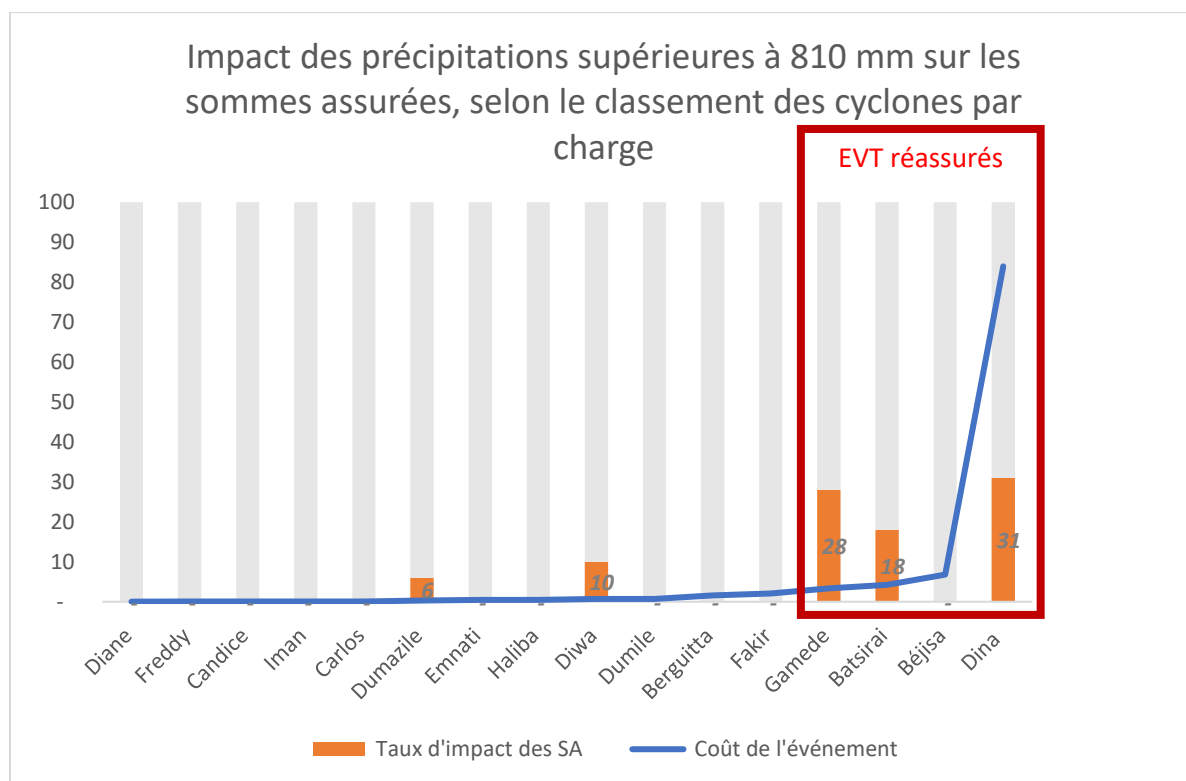


Figure 32 : Proportion des sommes assurées impactées par les précipitations supérieures à 810 mm lors des événements cycloniques. Réalisé avec Excel.

Les résultats montrent que les événements réassurés Gamède, Batsirai et Dina ont eu des effets plus significatifs en termes de taux de sommes assurées impactées que les cyclones non réassurés, confirmant ainsi la pertinence du seuil **810 mm** retenu. Toutefois, l'événement Béjisa, qui a été réassuré, présente une anomalie, avec un impact relativement faible sur les sommes assurés. Cette observation suggère que d'autres variables explicatives, telles que la vitesse des vents, doivent être intégrées pour expliquer ce comportement particulier.

Seuil de rafales de vents

L'analyse des rafales de vent à La Réunion, basée sur la théorie des valeurs extrêmes, montre des caractéristiques distinctes par rapport aux précipitations. Le ME-Plot affiche une pente décroissante, indiquant un paramètre de forme négatif ($\xi < 0$) et donc une distribution à queue fine. Les extrêmes sont donc moins nombreux que pour les cumuls de précipitation, ce qui est cohérent avec le fait que La Réunion soit plus connue pour ses cumuls de précipitations records (cf. Partie 1 : Cadre de l'étude et contexte et Annexe 6 : Graphique des périodes de retour des cyclones par pays selon les vitesses de vent), que pour ses vitesses de vents.

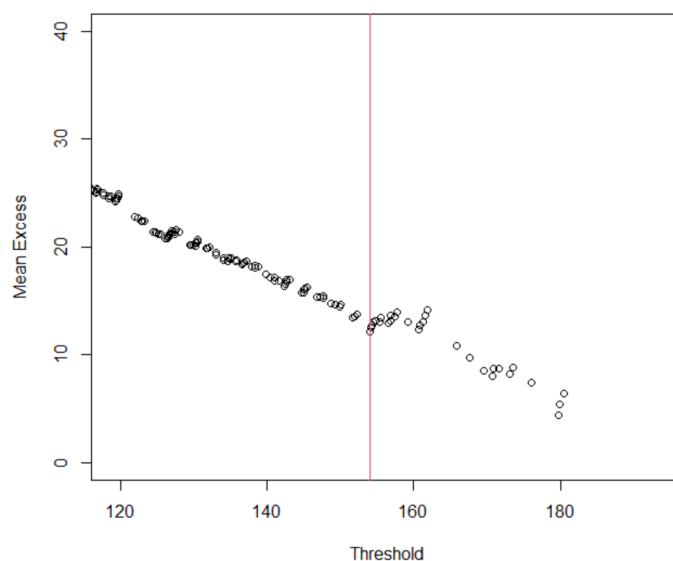


Figure 33 : Mean Excess Plot des rafales de vents. Réalisé sous R.

La recherche d'un seuil pertinent pour les rafales de vent s'appuie d'abord sur l'analyse de la fonction moyenne des excès (*ME-Plot*). Celle-ci met en évidence une rupture de la linéarité autour de 154 km/h (droite rouge sur le graphique).

Ce seuil est corroboré par des validations métiers. En effet, comme le montre la figure ci-dessous, 154 km/h correspond à la vitesse minimale définissant un cyclone de catégorie 2 sur l'échelle de Saffir-Simpson. À ce niveau, les risques de dommages structurels sur les habitations deviennent significatifs, avec une forte probabilité de destruction des toits et un risque accru d'effondrement des murs.

Description des types de dommages				Classe	Vitesse de vent en km/h (vent sur 1 min.)
Constructions légères	Maisons	Arbres	Récoltes		
Phénomène de faible ampleur (tempête ou dépression tropicale)				-1	< 118
Structures en bois ou en bardeaux peuvent être emportées.	Dommages aux toitures, pignons ou cheminées pour les maisons avec structures.	Les branches les plus importantes peuvent être cassées.	Dommages aux cannes à sucre et autres cultures	1	119-153
Forte probabilité d'être détruites.	Forte probabilité de destruction des toits. Murs peuvent s'effondrer.	Les arbres aux racines peu profondes peuvent être cassés ou déracinés.	Dommages aux grandes cultures (cannes, café, bananes) avec impacts sur les récoltes.	2	154-177
Effondrement de murs, haut niveau de destruction	Les maisons peuvent être détruites par l'enlèvement du toit ou seulement endommagées.	Les arbres cassés ou déracinés bloquent les voies de communications.	Destruction des cultures. Problème d'approvisionnement alimentaire possible.	3	178-208
Pourcentage de destruction important.	Importantes destructions. Toits arrachés et dégâts matériels importants.	Endommagement ou déracinement des arbres des poteaux télégraphiques.	Destruction sévère des cultures avec impacts sur l'économie. Problème de subsistance durable possible.	4	209-251
Destruction totale.	Important pourcentage de maisons détruites et effondrement de bâtiments en maçonnerie. Débris emportés par le vent.	Presque tous les arbres sont cassés ou déracinés. La destruction des poteaux électriques provoque interruption de l'alimentation.	Destruction complète des cultures. Crise de subsistance majeure et pertes économiques graves.	5	> 252

Figure 34 : Echelle Saffir-Simpson croisée par la NOAA avec une description des dommages potentiels pour différents types d'enjeux. Source : Taylor et al. (2010).

De plus, le seuil de 154 km/h coïncide avec le 90e quantile des données disponibles, renforçant sa pertinence. En actuariat, ce quantile est utile pour l'identification des valeurs extrêmes, car il capture les 10 % des événements les plus rares et intenses. Il offre également une bonne balance entre la robustesse statistique et la simplicité d'interprétation, en évitant les effets d'événements aberrants trop rares, tout en restant adapté à l'analyse des sinistres importants.

Ces croisements entre considérations statistiques et métier justifient pleinement l'adoption de 154 km/h comme seuil final pour les rafales dans ce modèle.

Pour crédibiliser ce seuil, plusieurs distributions théoriques ont été testées afin de modéliser les excès au-delà de ce seuil :

1. GPD avec les paramètres de formes de la distribution à partir du seuil 154 km/h ($\xi < 0$),
2. Exponentielle : Cas particulier de la GPD ($\xi = 0$),
3. Weibull : Cas particulier de la GPD ($\xi < 0$),
4. Log-Normale : Pour modéliser des données asymétriques,
5. Gamma : Appropriée pour des données positives.

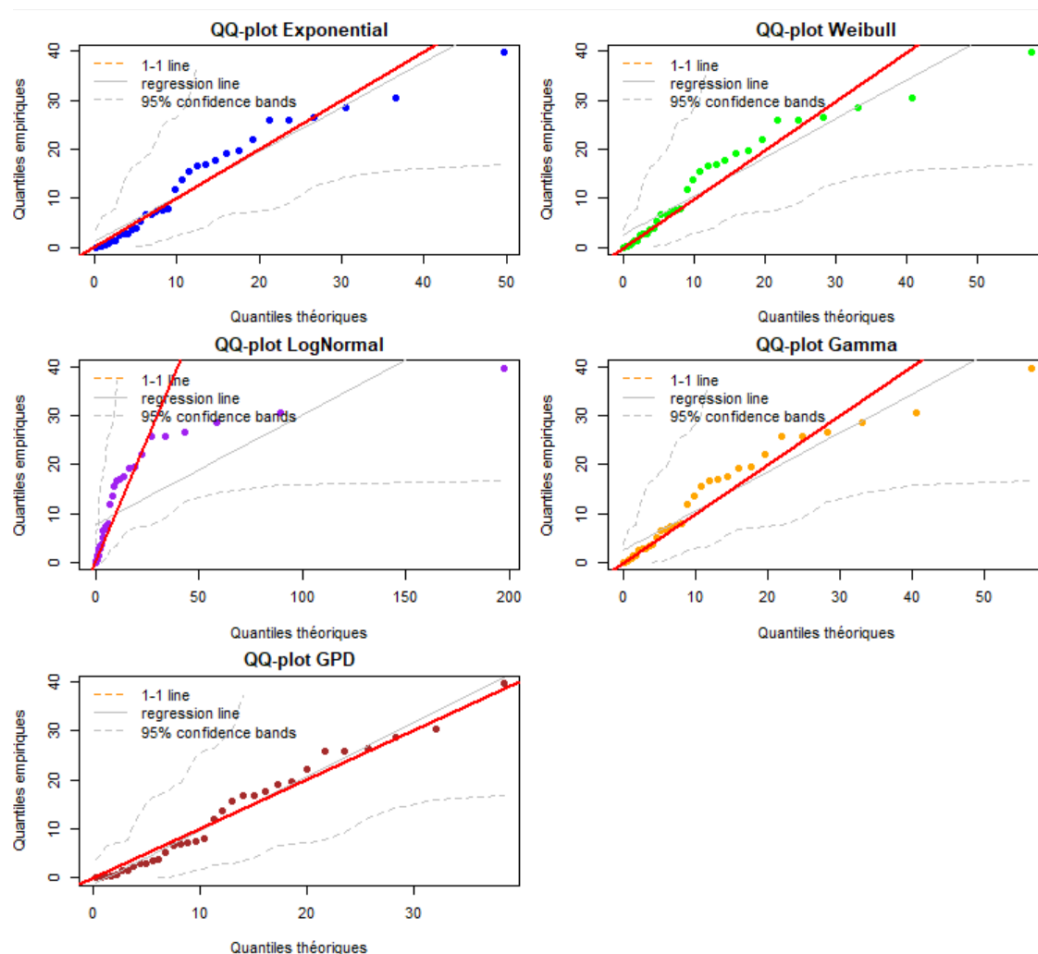


Figure 35 : QQ-plot des rafales de vents. Réalisé sous R.

Les QQ-Plots montrent que les excès au-delà de 154 km/h s'ajustent particulièrement bien à une loi de Pareto généralisée. Ce résultat, attendu compte tenu du choix initial du seuil, est néanmoins rassurant, car il valide statistiquement l'hypothèse retenue et renforce la confiance dans l'utilisation de la GPD pour cette analyse.

Une fois ce seuil fixé, la modélisation des excès au-delà de cette valeur a été réalisée à l'aide d'une loi de Pareto généralisée. Les paramètres de la GPD, à savoir le paramètre d'échelle (β) et le paramètre de forme (ξ), ont été estimés à l'aide de la méthode du maximum de vraisemblance, qui maximise la probabilité des données observées sous le modèle. Étant donné que l'estimation de ces paramètres ne

constitue pas l'objectif principal de ce mémoire, les lecteurs intéressés peuvent se référer à la section 2.6.3 de l'ouvrage de Stuart Coles, *An Introduction to Statistical Modeling of Extreme Values*.

Sur la base de ces estimations ($\beta = 15,22$ et $\xi = -0,27$), la borne supérieure théorique des rafales de vent a été calculée selon l'équation suivante :

$$\text{Borne supérieure théorique} = p + \frac{\beta}{|\xi|}$$

Où $p = 154$ est le seuil retenu.

En substituant les valeurs, une borne théorique à 211 km/h est obtenue.

L'observation maximale obtenue après interpolation et sélection des stations proches des zones assurées est de 194 km/h, ce qui reste en deçà de la borne supérieure théorique calculée à 211 km/h. Ce résultat confirme que la borne théorique fournit une estimation réaliste et cohérente avec les données disponibles.

En outre, seules 3 observations, parmi les stations supprimées car situées à plus de 1000 m d'altitude et donc non représentatives des zones assurées, dépassent cette borne.

Pour valider le seuil de rafales de 154 km/h, une analyse des données de sinistralité historique a été effectuée. Cette vérification repose sur le pourcentage de sommes assurées appartenant aux communes où des vitesses de vent supérieures à ce seuil ont été enregistrées lors des principaux événements cycloniques.

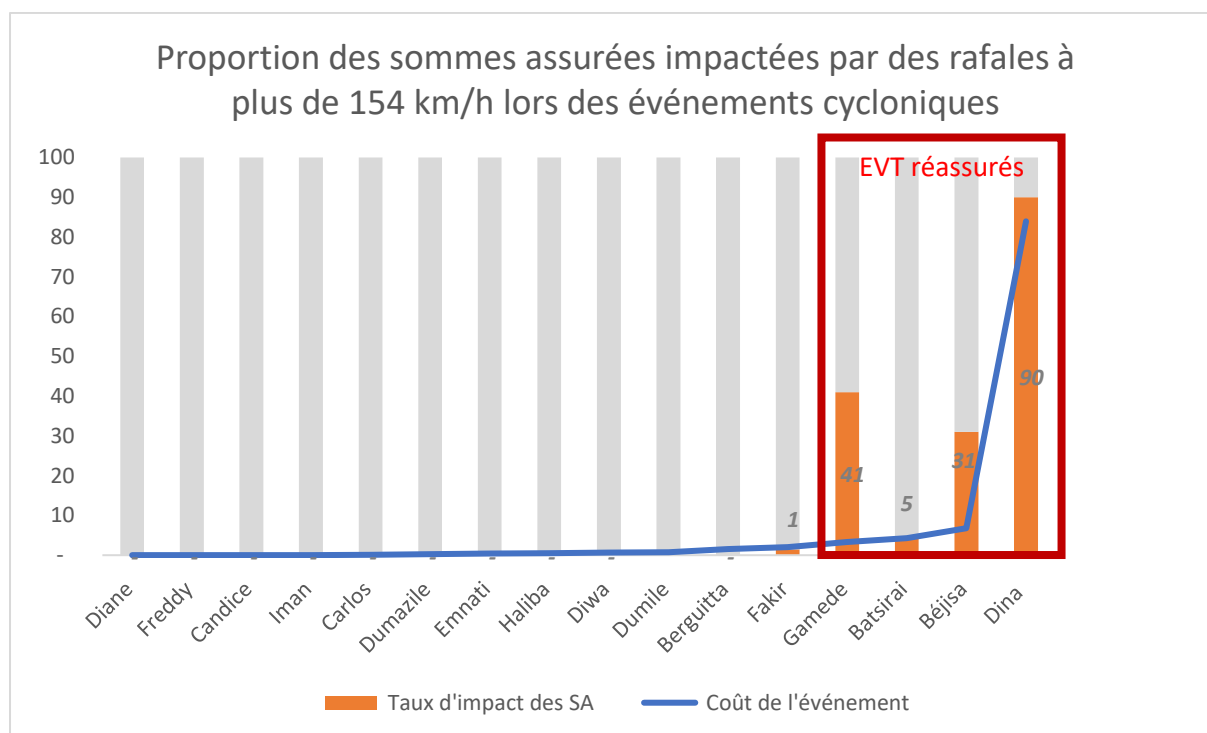


Figure 36 : Proportion des sommes assurées impactées par des rafales à plus de 154 km/h lors des événements cycloniques. Réalisé avec Excel.

Les résultats confirment que les événements Gamède, Béjisa et Dina ont entraîné des impacts significatifs sur les sommes assurées dans les communes dépassant ce seuil. Toutefois, le cas du cyclone Batsirai mérite une attention particulière. Avec un pourcentage d'impact faible des rafales de vent sur les sommes assurées (5 %), cet événement illustre les limites d'une analyse basée uniquement sur cet aléa. L'impact bien plus marqué des précipitations (18%), observé dans une analyse précédente,

souligne la nécessité d'étudier conjointement les deux aléas pour mieux appréhender les dommages causés par les cyclones.

L'analyse des sinistres historiques valide le seuil de **154 km/h**, en cohérence avec les impacts observés pour les événements majeurs. Ce seuil permet de capturer efficacement les dégâts causés par les rafales de vent dans les zones assurées.

Les seuils de 154 km/h pour les rafales de vent et de 810 mm pour les cumuls de précipitations permettent d'identifier les événements climatiques aux impacts significatifs. Dans la suite de cette analyse, l'accent sera mis sur l'étude des sommes assurées impactées par au moins un de ces deux aléas, afin de mieux comprendre leur contribution sur les sinistres enregistrés.

1.3. Critères de réassurance pour les cyclones : étude basée sur les sommes assurées impactées par des événements extrêmes

À ce stade de l'analyse, l'objectif est d'examiner les deux seuils précédemment définis pour les cumuls de précipitations (810 mm) et les rafales de vent (154 km/h), en évaluant leur impact conjoint sur les sommes assurées. Plus précisément il s'agit de déterminer le pourcentage de sommes assurées concernées par des précipitations et/ou des vents dépassant ces seuils à garder pour définir si un événement sera réassuré ou non. L'idée est que si un événement se produit dans une région où Prudence Créole est peu implantée, cet événement n'aura pas d'impact, quelle que soit l'intensité des vents ou des cumuls de précipitations.

Sur les données empiriques ci-dessous, un premier constat suggère que ce pourcentage se situe entre **10% et 18%** de sommes assurées impactées.

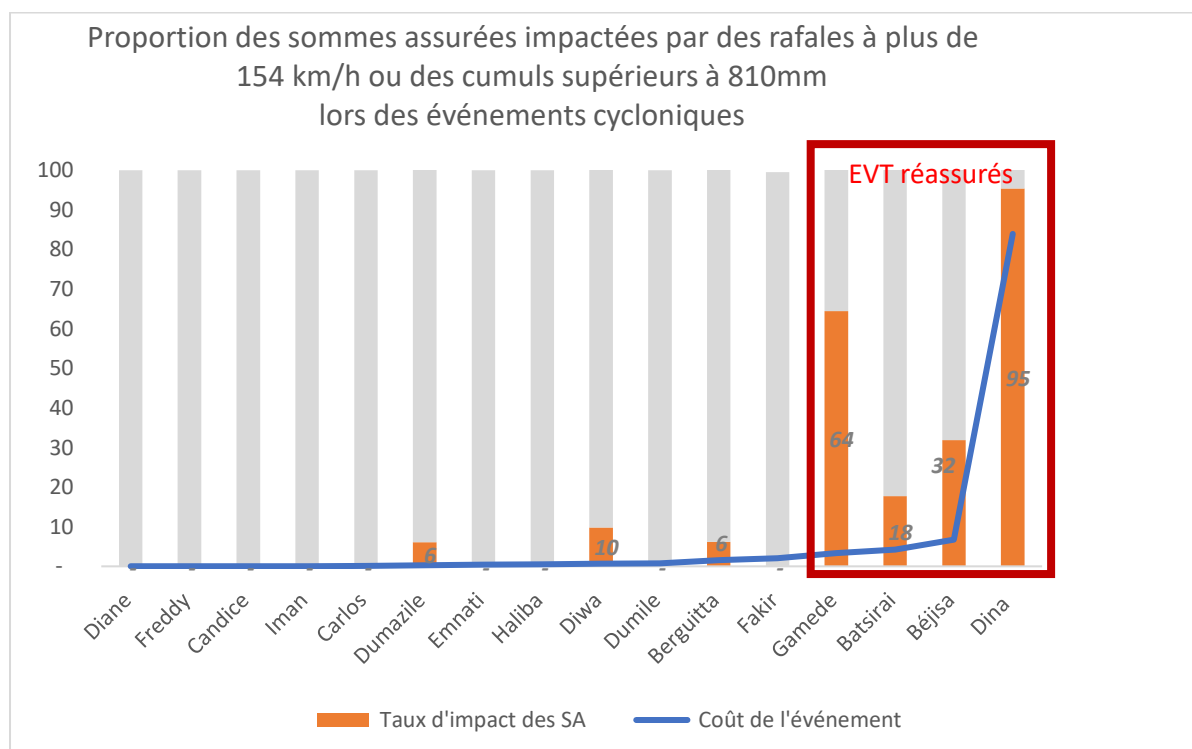


Figure 37 : Proportion des sommes assurées impactées par des rafales à plus de 154km/h ou des cumuls supérieurs à 810mm lors des événements cycloniques. Réalisé avec Excel.

Il reste à valider cette hypothèse afin de s'assurer qu'il ne s'agit pas d'un choix arbitraire. Pour ce faire, un arbre de décision via l'algorithme *CART* est utilisé. L'algorithme *CART* (*Classification and Regression Trees*), introduit par Breiman et al. (1984), est une méthode d'analyse exploratoire utilisée pour modéliser les relations entre une variable cible Y et un prédicteur X . Dans le cadre de ce mémoire, un arbre de décision est créé avec comme variable cible Y , une variable catégorielle (cyclone d'intérêt pour la réassurance ou non) et comme variable explicative X , le pourcentage de sommes assurées impactées par un aléa intense.

Cet algorithme repose sur un processus de division binaire des observations. Chaque division consiste à scinder un groupe d'observations, appelé nœud parent, en deux sous-groupes, appelés nœuds enfants. Ce processus commence au nœud racine, qui contient l'ensemble des données, et se poursuit jusqu'à ce que chaque nœud respecte un critère d'arrêt par exemple, lorsqu'un nœud est considéré comme pur ou lorsque que l'impureté d'un nœud est limitée.

Un nœud est pur lorsqu'il contient uniquement des observations appartenant à une seule catégorie de la variable cible Y (par exemple, cyclone d'intérêt "oui" ou "non"). À l'inverse, un nœud est impur lorsqu'il contient des observations appartenant à plusieurs catégories. L'objectif de l'algorithme est de réduire l'impureté à chaque étape de partitionnement pour arriver, autant que possible, à des nœuds terminaux purs.

Pour quantifier le degré d'impureté d'un nœud, l'indice de Gini est utilisé. Cet indice mesure la probabilité qu'une observation soit mal classée si elle était assignée aléatoirement à une catégorie en fonction de la distribution des classes dans le nœud. La formule mathématique de l'indice de Gini pour un nœud donné est la suivante :

$$I_G = 1 - \sum_{k=1}^K p_k^2$$

Où :

- K : nombre total de catégories de la variable cible Y (dans ce cas, $K = 2$, cyclone d'intérêt "oui" ou "non"),
- p_k : proportion d'observations appartenant à la catégorie k dans le nœud.

Lorsqu'un nœud est divisé en deux sous-nœuds enfants (gauche et droit), la réduction de l'impureté est calculée comme suit :

$$\Delta I_G = I_G^{parent} - (w_{gauche} \cdot I_G^{gauche} + w_{droit} \cdot I_G^{droit})$$

Où :

- I_G^{parent} : impureté du nœud parent,
- I_G^{gauche} et I_G^{droit} : indices de Gini des nœuds enfants,
- w_{gauche} et w_{droit} : proportions des observations dans les nœuds gauche et droit par rapport au nœud parent.

L'algorithme choisit la division qui maximise cette réduction d'impureté à chaque étape.

Lorsqu'un arbre *CART* est construit, il peut parfois se développer de manière excessive, capturant des détails inutiles ou du bruit dans les données, ce qui peut entraîner un surapprentissage.

Pour éviter ce problème, l'arbre est d'abord laissé se développer pleinement, jusqu'à ce que chaque nœud atteigne une impureté minimale. Ensuite, un processus d'élagage est appliqué pour réduire la taille de l'arbre. Ce processus repose sur le critère de complexité des coûts, introduit par Breiman et al. (1984), qui pénalise les arbres trop complexes. La mesure de complexité des coûts est définie par :

$$R_{\alpha}(T) = R(T) + \alpha|T|$$

Où :

- $R(T)$ est le risque de l'arbre, mesuré par le taux d'erreur de classification dans le cas des arbres de classification

$$\text{Erreur de classification} = \frac{\text{Nombre d'observations mal classées}}{\text{Nombre total d'observations}}$$

- α est un coefficient de pénalité qui contrôle la taille de l'arbre,
- $|T|$ est le nombre de nœuds terminaux.

L'objectif de l'élagage est de minimiser $R_{\alpha}(T)$, ce qui permet de trouver l'arbre optimal T_{α} pour chaque valeur de α . La sélection du meilleur α se fait par validation croisée, en choisissant celui qui minimise le risque estimé. L'arbre final ainsi obtenu offre un compromis optimal entre simplicité et précision.

Pour construire cet arbre de décision, nous avons utilisé l'algorithme CART avec la bibliothèque *rpart* sur R. Le modèle a été ajusté en utilisant :

- La variable cible Y : "1" si le coût d'un cyclone dépasse 3,75 M€ (cyclone d'intérêt car supérieur à 75% de la priorité du traité de réassurance DAB), et "0" sinon.
- La variable explicative X : le pourcentage de sommes assurées impactées par des rafales ou des précipitations.

L'objectif principal est de déterminer un seuil optimal à partir duquel les sommes assurées seraient impactées de manière à déclencher un suivi de réassurance.

L'arbre de décision généré a montré une première branche au seuil de 12 % pour la variable explicative X , divisant les cyclones en deux catégories : ceux ayant un pourcentage d'impact supérieur à ce seuil (potentiellement associés à un coût élevé) et ceux en dessous. L'arbre a ensuite continué à se développer, créant plusieurs branches correspondant à des divisions supplémentaires.

Un élagage a donc été réalisé à l'aide de la fonction *prune()*, en sélectionnant un paramètre de complexité correspondant à une seule division principale. Cette simplification a permis de réduire l'arbre à un nœud racine et deux nœuds terminaux, tout en maintenant une erreur relative de 0,2, soit une diminution de 80 % par rapport à l'arbre initial. De plus, l'erreur de validation obtenue avec cet arbre élagué est identique à celle de l'arbre complet, ce qui indique que des divisions supplémentaires n'améliorent pas la performance prédictive hors échantillon.

En conséquence, l'arbre élagué conserve toute sa pertinence en mettant en évidence le seuil clé de 12 %, identifié comme le critère optimal pour différencier les cyclones intéressants pour la réassurance de ceux qui ne le sont pas. Cette structure simple et robuste offre une règle décisionnelle claire, directement exploitable dans un contexte opérationnel.

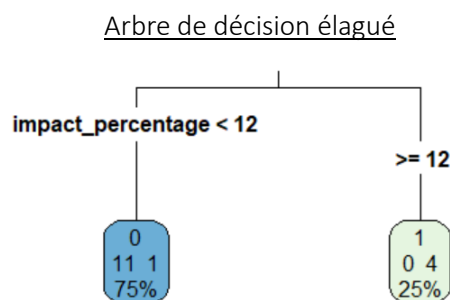


Figure 38 : Arbre de décision élagué pour déterminer le seuil de sommes assurées impactées nécessaire pour déclencher la réassurance. Réalisé sous R et retravaillé sous Paint pour ne pas communiquer le réel. Les résultats de l'arbre sont en revanche les bons.

L'arbre élagué ci-dessus met en évidence une règle décisionnelle simple et exploitable, basée sur le pourcentage de sommes assurées exploitées (SA) :

- À gauche ($SA < 12\%$) : un groupe incluant 12 cyclones, dont 11 qui ne font pas l'objet d'un suivi de réassurance (coût inférieur à 3,75 M€), représentant 75 % des observations.
Le cyclone Fakir est le seul dépassant le seuil de déclaration de 3,75 M€ mais de très peu, ce qui fait qu'il n'est pas réassuré.
- À droite ($SA \geq 12\%$) : un groupe regroupant 4 cyclones majeurs, tous avec un coût supérieur à 75% de la priorité, représentant 25 % des données.

Cet arbre élagué met en évidence le seuil de 12% comme critère clé pour identifier les cyclones significatifs dans un contexte opérationnel. Il offre une structure simple, robuste et directement exploitable, tout en soulignant des situations particulières, comme celle de Fakir, où la connaissance terrain de Prudence Créole est un atout afin de confirmer que l'événement ne sera pas réassuré.

Pour évaluer la robustesse des seuils établis, un *back-test* a été réalisé sur le cyclone Belal, qui n'a pas été inclus dans la construction des critères. Les résultats montrent qu'en prenant un seuil de rafales à 154 km/h et un seuil de précipitations à 810 mm, Belal aurait bien été identifié comme un événement à analyser dans le cadre d'une cession en réassurance. En effet, 65 % des SA ont été affectées par des rafales de vent au-dessus de 154 km/h et 4 % par des cumuls de précipitations au-dessus de 810 mm lors de l'événement. Ainsi, 67 % des SA ont été affectées par des vitesses de vent et/ou des cumuls de précipitation au-dessus des seuils lors cet événement, confirmant son caractère majeur. Ces résultats soulignent également la prépondérance des vents dans l'impact global du cyclone Belal, ce qui pourrait être un atout pour des études futures sur l'aléa à rattacher aux sinistres, la base de données de sinistralité ne fournissant pas cette information de manière exhaustive.

Ce *back-test*, bien qu'il repose sur un unique cyclone, offre des résultats encourageants sur la capacité du modèle à identifier efficacement les cyclones susceptibles d'être réassurés, tout en mettant en avant la complémentarité des deux aléas étudiés. Il sera également un bon complément au retour terrain de Prudence Créole.

Afin d'aller au-delà de la simple détection d'un événement d'ampleur susceptible de déclencher une couverture de réassurance l'étape suivante consiste à développer un modèle statistique : un modèle CAT simplifié pour permettre une estimation des coûts associés à chaque cyclone.

Chapitre 2 : Estimation de la charge d'un cyclone réassuré

À ce stade, les critères pour identifier un cyclone nécessitant un suivi par la réassurance ont été établis : si 12% des sommes assurées de Prudence Créole sont affectées par des rafales supérieures à 154 km/h ou des cumuls de précipitations dépassant 810 mm, alors le cyclone sera probablement réassuré. L'étape suivante consiste à estimer la charge financière associée à l'événement, en complément des triangles réalisés par Prudence Créole.

Dans cette partie, l'objectif principal est de construire et d'appliquer un modèle statistique adapté aux spécificités des données étudiées, en l'occurrence un modèle linéaire généralisé (GLM). Ce type de modèle permet de relier la variable cible : la charge des sinistres cycloniques, à plusieurs variables explicatives, tout en prenant en compte les particularités de la distribution des données, telles que leur asymétrie ou leur hétérogénéité.

La construction de ce modèle repose sur l'exploitation de données de sinistres historiques, des caractéristiques du portefeuille de Prudence Créole, ainsi que des données météorologiques, introduites dans la partie deux.

Chaque étape de cette section a pour but de démontrer la pertinence de l'approche adoptée, de garantir la robustesse du modèle final et d'illustrer pourquoi le GLM constitue une méthodologie adaptée à notre modélisation des charges de sinistres cycloniques.

2.1. Retraitements préalables à la modélisation

Le tableau ci-dessous énumère les variables à notre disposition pour construire notre modèle suite aux traitements détaillés dans la partie 2.

Variable	Type	Nature	Nombre de modalités
Communes	Qualitative	Nominale	24
Années de survenance	Temporelle	-	12
Vitesse de vent	Quantitative	Continue	-
Cumul de précipitation	Quantitative	Continue	-
Sommes assurées (SA)	Quantitative	Continue	-
SA en tout béton	Quantitative	Continue	-
SA en béton avec une toiture dans un autre matériel	Quantitative	Continue	-
SA en acier	Quantitative	Continue	-
SA appartement	Quantitative	Continue	-
SA commercial	Quantitative	Continue	-
SA immeuble	Quantitative	Continue	-
SA maison	Quantitative	Continue	-
SA industrielle	Quantitative	Continue	-

Tableau 12 : Tableau récapitulatif des variables disponibles pour la modélisation à la suite des traitements de la partie 2. Réalisé avec Excel.

Avant de construire le modèle, des retraitements préalables des données sont nécessaires afin d'assurer la robustesse et la pertinence de la modélisation. Ces retraitements incluent le regroupement des variables quantitatives en classes, la gestion des corrélations entre variables explicatives et « le plafonnement » des valeurs extrêmes, chaque étape étant expliquée théoriquement puis appliquée.

Regroupement en classe

Le regroupement en classes des variables quantitatives continues présente plusieurs intérêts essentiels pour la modélisation actuarielle.

Tout d'abord, il facilite l'établissement de relations linéaires entre les variables explicatives et la charge de sinistre, ce qui est indispensable pour leur intégration dans un modèle linéaire généralisé.

Ensuite, il réduit le risque de corrélations excessives entre les modalités des variables explicatives, contribuant ainsi à renforcer la stabilité des coefficients estimés et à améliorer l'interprétabilité des résultats.

Pour mettre en œuvre ce regroupement en classes, des techniques statistiques adaptées doivent être utilisées, tel que le data clustering utilisé pour diviser un ensemble de données en groupes homogènes selon des critères de proximité.

Dans cette étude, la méthode de data clustering retenue est celle des *K-means* en raison de son efficacité et de sa simplicité pour classer des variables quantitatives continues. Cette méthode a été appliquée aux variables quantitatives qui ne peuvent être directement utilisées dans un GLM en raison de leur nature. Le processus de classification par *K-means* est décrit ci-dessous.

Initialement, K centres sont choisis aléatoirement dans l'espace des observations. Chaque individu est ensuite affecté à la classe dont le centre est le plus proche, selon une mesure de distance. Une distance sur un ensemble E est une application d définie sur le produit $E^2 = E \times E$ à valeurs dans l'ensemble $\mathbb{R}^+$ des réels positifs, vérifiant les trois propriétés suivantes :

Symétrie : $\forall (x, y) \in E^2, \quad d(x, y) = d(y, x),$

Séparation : $\forall (x, y) \in E^2, \quad d(x, y) = 0 \Leftrightarrow x = y,$

Inégalité triangulaire : $\forall (x, y, z) \in E^3, \quad d(x, z) \leq d(x, y) + d(y, z).$

La distance euclidienne est couramment utilisée pour mesurer la proximité entre deux individus. Elle est définie comme suit :

$$d(x, y) = \sqrt{\sum_{j=1}^p (x_j - y_j)^2}$$

Où x et y sont des vecteurs représentant deux individus, et p est le nombre de dimensions.

Une fois cette étape réalisée, les centres des classes sont recalculés en prenant le barycentre⁴⁶ des individus leur appartenant, défini par :

$$g_k = \frac{1}{n_k} \sum_{i \in \pi_k} x_i$$

Où

- n_k est le nombre d'individus dans la classe k ,
- π_k désigne l'ensemble non vide des indices des individus assignés à cette classe.

⁴⁶ Dans le cadre d'un algorithme des *K-means*, un barycentre correspond au centre d'un cluster.

Ce processus d'affectation et de mise à jour des centres est répété jusqu'à l'obtention d'un critère de classification stable, garantissant ainsi des regroupements homogènes. Ces étapes sont illustrées ci-dessous.

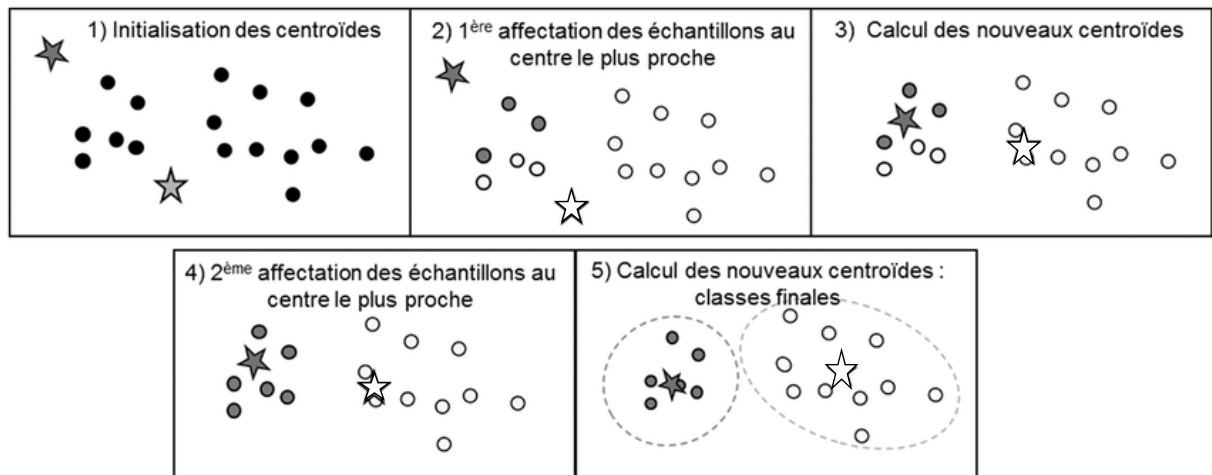


Figure 39 : Fonctionnement de l'algorithme des K-means dans le cas $k=2$. Source : researchgate.

Pour garantir une perte minimale d'information, le nombre de classes K est fixé de manière que le rapport entre la variance interclasse et la variance totale dépasse 95 % :

$$V = \frac{\text{variance interclasses}}{\text{variance totale}} > 95\%$$

Ce critère permet d'assurer une homogénéité suffisante des regroupements. La variance interclasse mesure à quel point les moyennes des classes sont éloignées de la moyenne globale des données. Elle est donnée par :

$$\text{Variance interclasses} = \sum_{k=1}^K n_k (\mu_k - \mu)^2$$

Où

- n_k est le nombre d'observations dans la classe k ,
- μ_k la moyenne des observations dans la classe k ,
- μ la moyenne globale des données.

Et la variance totale mesure la dispersion globale des observations par rapport à leur moyenne. Elle est donnée par :

$$\text{Variance totale} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

Où

- n est le nombre total d'observations,
- x_i une observation individuelle,
- μ la moyenne globale des données.

Cependant, il arrive fréquemment que les classes obtenues ne soient pas directement exploitables, par exemple en raison de tailles très hétérogènes entre les classes. Etant donné la taille limitée de nos données disponibles, une fois le nombre optimal de classes identifié, les données seront redistribuées de manière équitable entre les classes pour assurer leur pertinence. Ce processus est illustré dans l'application ci-dessous.

Dans notre cas, pour la variable vent, la méthode des *K-means* a été utilisée en suivant deux étapes principales. Tout d'abord, le nombre optimal de classes a été déterminé en maximisant le ratio entre la variance interclasses et la variance totale. Une séparation en six classes a été retenue, ce seuil étant le premier à dépasser les 95 %, comme le montre le tableau suivant. Bien que sept classes auraient permis un ratio supérieur (96,5 %), la faible augmentation ne justifie pas la complexité supplémentaire, rendant six classes plus adaptées pour garantir un bon équilibre entre précision et simplicité.

Nombre de classe de vitesses de vents	V
4	90,3%
5	93,4%
6	95,2%
7	96,5%

Tableau 13 : Tableau du nombre de classe de vitesse de vents. Réalisé avec Excel.

Néanmoins, en raison du faible nombre d'observations (308), les classes obtenues initialement par l'algorithme des *K-means*, présentaient des tailles trop hétérogènes. Pour remédier à cela, une répartition équitable des données entre les six classes, tout en conservant l'homogénéité au sein de chaque classe a été effectué. La répartition finale est la suivante avec en gras la modalité de référence, fixée comme celle ayant le plus grand effectif.

Classe de vitesse de vent en km/h	Effectif	Effectif en %
26 - 75,44	55	18%
75,44 - 93,6	55	18%
93,6 – 110,98	56	18%
110,98 - 128,75	53	17%
128,75 - 149,79	51	17%
149,79 - 193,67	38	12%

Tableau 14 : Tableau des classes de vitesse de vents. Réalisé avec Excel.

Ce processus garantit des regroupements pertinents et adaptés à l'intégration dans le modèle GLM. Le graphique ci-dessus illustre cette partition, mettant en évidence les différentes intensités de vent réparties de manière cohérente entre les six classes.

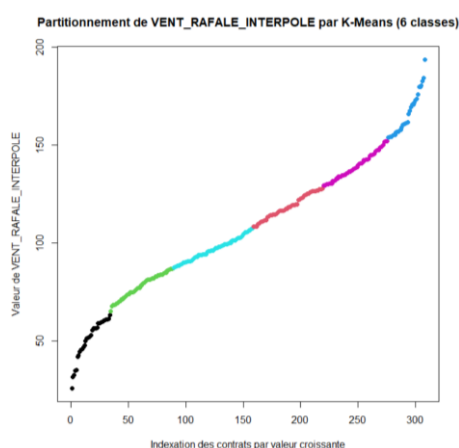


Figure 40 : Partitionnement des rafales de vent en 6 classes par K-means. Réalisé sous R.

Remarque : une classification des vitesses de vent avec des seuils métier a été testée (échelle Saffir-Simpson et classification des vitesses avec les seuils définissant les cyclones dans l’Océan Indien). Le nombre de groupes était cohérent avec celui proposé par la méthode *K-means*, mais les classifications n’ont pas été retenues car les groupes créés étaient trop hétérogènes. Les tableaux sont disponibles en Annexe 4 : Tableaux de regroupement en classes (variable rafales de vents et variables non utilisées).

Le même protocole a été suivi pour obtenir sept classes de cumul de précipitation et 4 classes d’engagements. Dans les deux cas, la modalité de référence a été désignée comme la classe contenant le plus d’observations (en gras dans les tableaux).

Classe de cumul de pluie sur la durée de l'événement en mm	Effectif	Effectif en %
0 - 58,6	46	17%
58,6 - 116,54	45	16%
116,54 - 194,5	46	17%
194,5 - 297,49	44	16%
297,49 - 437,77	41	15%
437,77 - 634,53	42	15%
634,53 – 3730	44	16%

Tableau 15 : Tableau du nombre de classe de cumul de pluie. Réalisé avec Excel.

Pour des raisons de confidentialité, les montants précis des sommes assurées ne peuvent pas être affichés. Par conséquent, ils ont été remplacés par des classes anonymisées, nommées “Classe de Somme Assurée 1”, “Classe de Somme Assurée 2”, etc.

Classe de SA en €	Effectif	Effectif en %
Classe SA 1	77	21%
Classe SA 2	80	21%
Classe SA 3	85	23%
Classe SA 4	66	18%

Tableau 16 : Tableau des classes de sommes assurées. Réalisé avec Excel.

Les autres variables quantitatives (SA béton, etc.) seront écartées par la suite au cours du choix des variables pour le GCL. Leur nombre de classes est disponible en Annexe 4 : Tableaux de regroupement en classes (variable rafales de vents et variables non utilisées).

Pour ce qui est des communes, pour simplifier la modélisation et limiter le nombre initial de modalités, qui était de 24, l’algorithme des *K-means* recommandait 3 classes qui ont été définies en fonction de la proportion de bâtiments assurés en béton dans chaque commune. Ce critère a été choisi pour éviter un regroupement arbitraire et garantir une meilleure homogénéité. Le type de construction influence directement la vulnérabilité aux cyclones, rendant ce choix particulièrement adapté pour refléter les spécificités des communes dans la modélisation. La carte ci-dessous présente ces trois classes.

Saint-Denis, en raison de la forte implantation de Prudence Créole, a été retenue comme commune de référence.

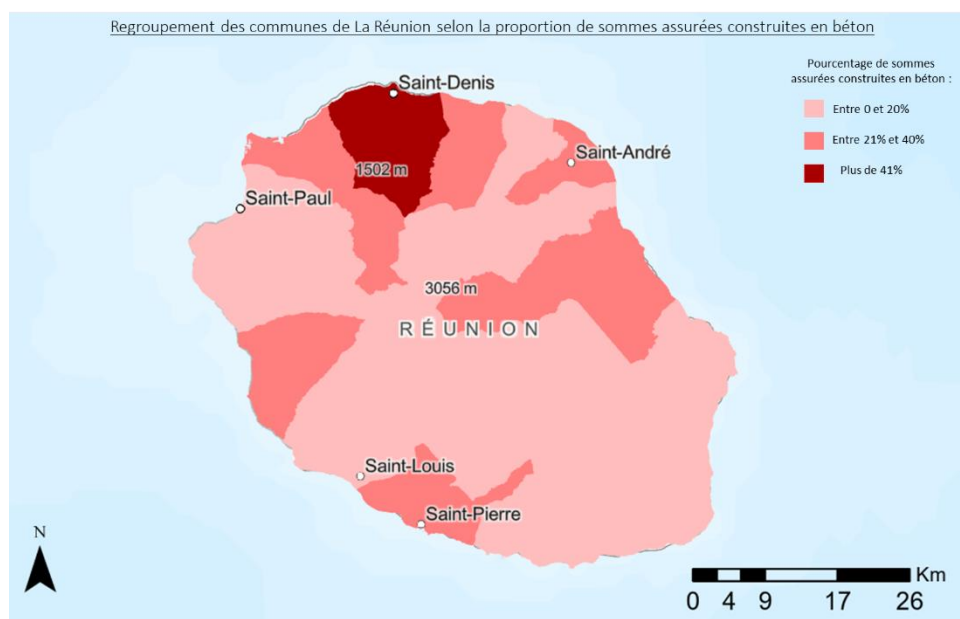


Figure 41 : Regroupement des communes de La Réunion selon la proportion de sommes assurées construites en béton. Réalisé avec ArcGIS.

Les années, quant à elles, ne sont pas regroupées pour l'instant, mais elles feront l'objet d'un regroupement après le test de significativité des variables dans le modèle GLM.

Corrélation entre les variables explicatives : implications et méthodes de détection

Lorsque deux variables explicatives sont fortement corrélées, cela signifie qu'elles apportent presque la même quantité d'information. Cette redondance peut poser des problèmes lors de l'estimation des coefficients de régression, car elle affecte la fiabilité et l'interprétation des résultats.

Une première conséquence serait que les coefficients pourraient contredire les connaissances métier. Par exemple, dans notre modèle, un coefficient négatif pour la vitesse des rafales suggérerait qu'une augmentation des vents diminue la charge de sinistre, ce qui est illogique.

De plus, une corrélation élevée peut augmenter les variances des coefficients, les rendant non significatifs selon le test de Wald⁴⁷, ce qui peut entraîner l'élimination injustifiée de variables essentielles.

Enfin, l'ajout ou le retrait de variables corrélées peut provoquer d'importantes variations dans les coefficients, compromettant ainsi la fiabilité des résultats. Ces problèmes soulignent l'importance de limiter les corrélations dans la construction des modèles.

Après transformation, toutes les variables explicatives qui pourraient expliquer la sinistralité sont désormais qualitatives. Parmi elles, les variables relatives aux sommes assurées (SA), telles que les classes de SA totales, béton, acier, ou encore les classes correspondant aux différents usages de bâtiments, sont vraisemblablement corrélées.

Pour analyser ces relations, le test d'indépendance du χ^2 est souvent utilisé. Cependant, une limitation de cette méthode réside dans son incapacité à mesurer l'intensité de la relation entre deux variables qualitatives. En effet, ce test permet uniquement de détecter si une dépendance existe ou non, mais il ne donne aucune indication sur la force ou la nature de cette relation. Par exemple, dans le cas où une association significative est détectée entre la classe des SA béton et les classes de SA Industriel, le χ^2 ne permet pas de déterminer si cette dépendance est forte, modérée ou faible.

⁴⁷ Le test de Wald permet d'évaluer la significativité individuelle des coefficients d'un modèle statistique en comparant leur estimation à leur erreur standard, sous l'hypothèse que le coefficient est nul. Source : cours de traitement statistique de l'information économique de Marie Bessec dispensé en licence 3 à l'université Paris Dauphine (2021).

Pour pallier cette limite, l'indice de V de Cramer, défini par la formule ci-dessous, est privilégié :

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(l - 1, c - 1)}}$$

Où :

- χ^2 représente la statistique du test du χ^2 ,
- n est la taille totale de l'échantillon,
- l et c sont respectivement le nombre de modalités des variables en ligne et en colonne.

Une valeur de V égale à 0 indique une absence totale de relation, tandis qu'une valeur proche de 1 traduit une association forte.

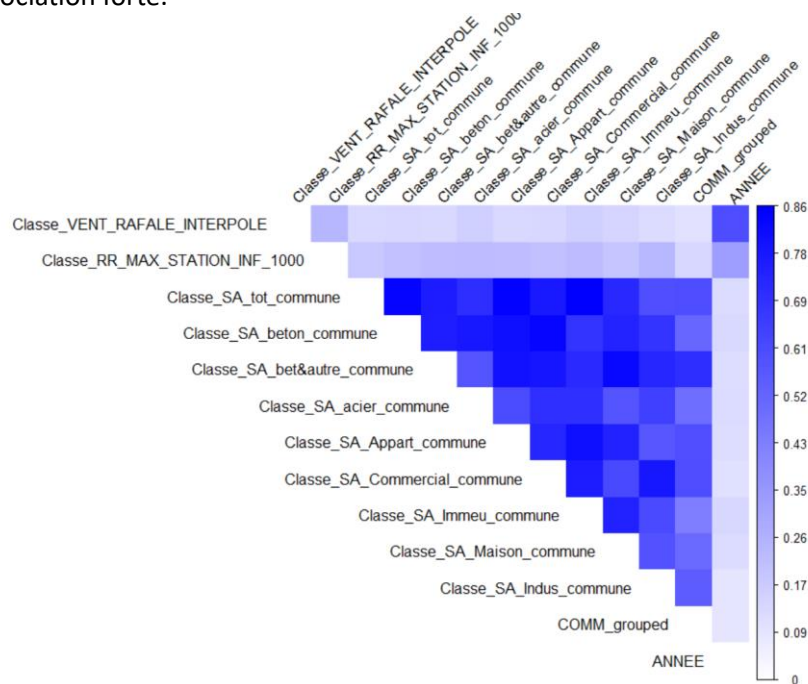


Figure 42 : Graphe des indices de Cramer pour les variables catégorielles. Réalisé sous R.

Les indices de Cramer relèvent que les variables météorologiques ou temporelles présentent des corrélations relativement faibles avec les autres variables. En revanche, des relations plus fortes entre les variables de sommes assurées suggèrent des redondances potentielles.

Pour éviter toute redondance et simplifier l'intégration dans le modèle, une seule variable parmi celles relatives aux sommes assurées sera retenue. Afin de guider ce choix, le test de Kruskal-Wallis est utilisé. Ce test statistique non paramétrique permet d'évaluer l'influence d'une variable explicative catégorielle (comme les SA) sur une variable continue, ici la charge de sinistre. Contrairement au test de l'ANOVA⁴⁸, ce test ne suppose pas que les données suivent une distribution normale, ce qui le rend particulièrement utile dans des contextes où les distributions sont asymétriques.

L'objectif du test est de comparer les rangs médians d'une variable continue Y entre différents groupes formés par les modalités d'une variable catégorielle X . Sous l'hypothèse nulle (H_0), les différents groupes sont similaires et ont une même tendance centrale. En d'autres termes, X n'a pas d'influence sur Y .

⁴⁸ Le test ANOVA (Analysis of Variance) permet de comparer les moyennes de plusieurs groupes pour déterminer si au moins une d'entre elles est significativement différente, en analysant la variance intra- et intergroupes. Source: Fisher, R. A. (1925). Statistical Methods for Research Workers. Edinburgh: Oliver & Boyd.

La statistique de Kruskal-Wallis (KW) mesure l'écart entre les rangs moyens de chaque groupe et la moyenne globale des rangs. Elle est calculée à l'aide de la formule suivante :

$$KW = \frac{12}{n(n+1)} \sum_{i=1}^k n_i \left(\frac{R_i}{n_i} - \frac{n+1}{2} \right)^2$$

$$\Leftrightarrow KW = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1)$$

Où :

- n_i : est l'effectif de la population i , où $i \in [1, k]$ et $n = \sum_{i=1}^k n_i$ représente la somme totale des effectifs.
- X_{ij} : la j -ième observation de l'échantillon i , avec $j \in [1, n_i]$.
- R_{ij} : rang de l'observation X_{ij} parmi les n observations et,
- $R_i = \sum_{j=1}^{n_i} R_{ij}$ est la somme des rangs associée à chaque groupe i .

Sous (H_0) , la statistique Kruskal-Wallis suit une loi du chi-deux avec $k - 1$ degrés de liberté. La p-value associée $P(X \leq \chi_{k-1}^2)$ permet de décider si l'hypothèse nulle peut être rejetée. Si la p-value est inférieure à un seuil de significativité (5 %), cela signifie que les groupes diffèrent significativement.

Toutes les p-values étant inférieures au seuil de 5 %, l'hypothèse nulle est rejetée. Cela signifie que ces variables exercent un effet significatif sur la charge de sinistre.

Le choix de la variable représentant les sommes assurées repose sur deux critères principaux. L'indice de V de Cramer montre que la variable des classes de sommes assurées totales sont les moins corrélées aux variables météorologiques et temporelles, réduisant ainsi le risque de redondance. Par ailleurs, le test de Kruskal-Wallis indique que cette variable distingue le mieux les variations de charges de sinistres, grâce à une statistique de test élevée. Les **sommes assurées totales** par commune seront alors conservées.

Lorsque toutes les données à l'échelle du bâtiment seront disponibles, il sera possible d'intégrer directement ces variables qualitatives sans passer par les sommes assurées. Une telle évolution permettrait d'affiner la modélisation et d'améliorer encore la précision des estimations.

Ces étapes de sélection permettent de réduire les redondances tout en conservant des variables explicatives pertinentes. Le modèle final sera donc construit à partir des classes de vents, des cumuls de précipitations, de sommes assurées, des communes et des années.

Plafonnement des sinistres extrêmes

La sinistralité cyclonique de Prudence Créole se caractérise par une forte volatilité avec seulement un cyclone si extrême, qu'il est le seul à avoir été reconnu CatNat pour le vent. Il s'agit du cyclone Dina, survenu il y a plus de 20 ans, et dont l'envergure reste unique dans l'histoire récente de La Réunion. Dans ce contexte, le traitement des données extrêmes nécessite une attention particulière.

En théorie, lorsqu'un ensemble de données présente des valeurs extrêmes, une solution consiste à séparer la distribution en deux parties : le "bas" de la distribution, modélisé par un modèle linéaire généralisé pour son comportement pouvant suivre une loi de la famille exponentielle⁴⁹, et le "haut" de la distribution, modélisé par une loi de Pareto généralisée (GPD) dans le cadre de la théorie des valeurs extrêmes. Cependant, cette approche est inadaptée dans notre cas, car un modèle GPD ne permet pas

⁴⁹ Cette notion sera détaillée dans la partie 2.2 Modèle linéaire généralisé

d'incorporer des variables explicatives telles que les rafales de vent ou les précipitations, qui sont essentielles pour comprendre la sinistralité cyclonique⁵⁰.

Pour contourner cette limitation, le choix a été fait de recourir au plafonnement. Cette méthode consiste à fixer un montant maximal pour les sinistres, réduisant ainsi l'impact des valeurs extrêmes sur la modélisation tout en permettant l'utilisation d'un GLM sur l'ensemble de la distribution. Le plafonnement garantit ainsi la possibilité d'intégrer des variables explicatives pertinentes dans le modèle tout en maîtrisant la volatilité des données et en respectant la fréquence des sinistres, aucun sinistre n'étant supprimé de la base.

Le choix du seuil de plafonnement a été réalisé en s'appuyant sur la théorie des valeurs extrêmes, en particulier la méthode du Peak Over Threshold (dont la théorie a été détaillée dans le 1.1 de cette partie). Le Mean Excess Plot a permis d'identifier un seuil à *1,5 millions d'euros* (représenté par la ligne rouge), déterminé visuellement en observant la cassure de linéarité limitant le plus possible le nombre d'observations à plafonner vu la taille de l'échantillon. Les légendes des axes des abscisses et des ordonnées ont été retirées pour des raisons de confidentialité.

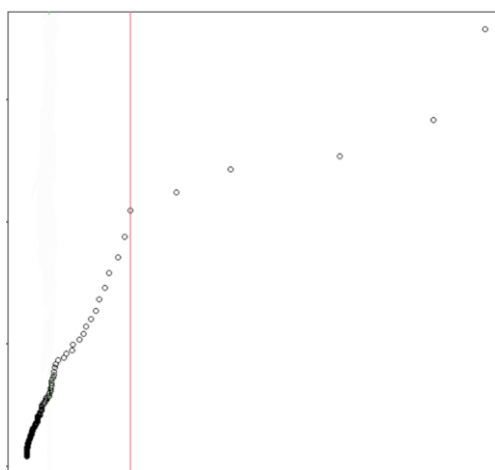


Figure 43 : ME plot des charges de sinistres sur l'échantillon d'entraînement. Réalisé sous R.

Ce choix a conduit au plafonnement des sinistres de huit communes dans la base d'entraînement, pour un total cumulé de *86 millions d'euros* de dépassements plafonnés. Cette démarche vise à limiter l'impact des sinistres exceptionnels sur la modélisation, tout en préservant la cohérence des données.

De plus, ce plafonnement est justifié, même dans le cadre d'une modélisation des extrêmes, car la base de données utilisée se concentre exclusivement sur les cyclones, qui sont déjà des événements rares et intenses parmi les sinistres climatiques. Inclure des observations excessivement extrêmes, comme certaines communes lors du cyclone Dina risquerait de surévaluer les paramètres du modèle, surtout compte tenu de la taille limitée de l'échantillon. Ces observations, par leur rareté, ne sont pas représentatives et ne peuvent pas être modélisées efficacement dans un cadre de modèle CAT simplifié. D'ailleurs comme évoqué dans la partie 2, un cyclone aussi destructeur que Dina serait évalué en appliquant les taux de destruction de Dina directement, plutôt que d'utiliser le modèle.

Toutefois, Dina n'a pas été entièrement exclu, car il constitue une référence incontournable pour les cyclones majeurs ayant impacté La Réunion, permettant de conserver une certaine représentativité

⁵⁰ Les modèles GPD traditionnels se concentrent uniquement sur les paramètres de la distribution des excès (ξ et σ) et ne permettent pas d'inclure des variables explicatives comme les rafales de vent ou les précipitations. Des approches récentes, telles que celle de Le Carrer et Gaetan (2022) (Distributional regression models for Extended Generalized Pareto distributions), proposent d'intégrer ces covariables, ce qui pourrait constituer une piste d'amélioration intéressante. Toutefois, ces méthodes restent complexes et hors du périmètre de ce mémoire, qui privilégie une solution simple et opérationnelle avec un GLM.

des événements rares dans le modèle. De plus, pour éviter de sous-estimer les résultats, un ajustement des paramètres est réalisé en appliquant un produit en croix, détaillé dans les sections suivantes. Ce processus permet de recalibrer les projections tout en conservant la cohérence et la pertinence du modèle.

Pour compléter cette analyse, l'Annexe 7 fournit un aperçu du modèle d'estimation des charges de sinistres sans plafonnement. Toutefois, il est recommandé au lecteur de consulter cette annexe à la fin du Chapitre 2 : Estimation de la charge d'un cyclone réassuré, afin de mieux comprendre les impacts du plafonnement dans le cadre global de la modélisation.

Après avoir appliqué les traitements nécessaires aux données, la base de données est désormais prête à être utilisée pour la modélisation.

La prochaine section se concentre sur la présentation des concepts théoriques des modèles linéaires généralisés, qui serviront à établir une relation statistique robuste entre les coûts des sinistres cycloniques et les variables explicatives retenues.

2.2. Modèle linéaire généralisé

Les modèles linéaires généralisés constituent une extension des modèles linéaires classiques, permettant de traiter des données dont les distributions ne sont pas nécessairement normales⁵¹. Ils sont particulièrement adaptés pour modéliser des relations entre une variable réponse Y et un ensemble de variables explicatives $X_0, X_1, \dots, X_p$, lorsque Y suit une distribution de la famille exponentielle.

Dans ce mémoire, Y correspond au coût total des sinistres associés à un événement cyclonique. Ce coût est obtenu en sommant les charges observées pour l'ensemble des communes affectées par le cyclone. Soit S_e la charge totale pour l'événement e , définie comme :

$$S_e = \sum_j S_{j,e}$$

où $S_{j,e}$ représente la charge observée pour la commune j , j variant de 1 à n , lors de l'événement e .

L'objectif est de modéliser l'espérance conditionnelle $E[S_{j,e} \mid X_{j,e} = x_{j,e}]$ pour chaque commune, c'est-à-dire prédire la charge de sinistre attendue en fonction des variables explicatives.

Ces prédictions permettent ensuite d'estimer la charge totale S_e pour l'événement, en sommant les prédictions sur toutes les communes.

Cette démarche permet de relier les caractéristiques des communes et des cyclones aux charges observées, tout en fournissant des estimations fiables pour des événements futurs.

En effet, les caractéristiques explicatives, regroupées dans le vecteur $X_{j,e}$, incluent des éléments spécifiques aux communes (les sommes assurées par exemple) ainsi que des caractéristiques propres au cyclone, comme des données météorologiques.

Les données explicatives sont regroupées dans une matrice X de dimension $n \times (p + 1)$, où chaque ligne correspond à une commune et chaque colonne à une variable explicative tel que :

$$X = (X_0, X_1, \dots, X_p) \in R^{n \times (p+1)}$$

⁵¹ La théorie développée dans cette section s'appuie sur les enseignements dispensés par Monsieur Baradel dans le cadre du cours d'actuariat Non-Vie de dernière année de l'ENSAE.

Où pour un e fixé,

- $X_0 = 1_n$ représente l'intercept (une colonne de 1),
- $X_j = (x_{1j}, \dots, x_{nj})'$ désigne la j -ème variable explicative (par exemple la vitesse maximale du vent), avec $j \in \{1, \dots, p\}$,
- $(x_{ij})_{0 \leq j \leq p}$ correspond à la valeur de la j -ème variable explicative pour la commune i (i allant de 1 à n).

Ainsi, chaque observation $S_{j,e}$ (la charge observée pour une commune j et un événement e) est associée à un vecteur explicatif $X_{j,e}$, regroupant les variables explicatives pertinentes. Elles permettent de capturer les caractéristiques spécifiques à chaque commune et aux conditions cycloniques liées à l'événement e .

Dans le cadre de ce mémoire, l'hypothèse d'indépendance entre les observations est adoptée pour garantir la simplicité et l'efficacité de la modélisation via un GLM. Bien que cette hypothèse ne reflète pas entièrement la réalité, notamment en raison de possibles corrélations spatiales entre communes voisines, elle constitue une approximation robuste permettant de capturer les dynamiques principales des sinistres. Cette approche offre une base solide pour la modélisation tout en ouvrant des perspectives d'amélioration pour intégrer ces dépendances dans de futurs travaux.

Le modèle GLM repose sur trois composantes principales, détaillées ci-dessous, permettant d'établir une relation entre les variables explicatives et la variable cible.

Les trois composantes d'un modèle linéaire généralisé

1. Une distribution de la variable réponse

Dans un GLM, la variable réponse $S_{j,e}$ suit une loi appartenant à la famille des lois exponentielles. Cette famille est définie par une fonction de densité donnée par :

$$f(S_{j,e}; \theta_{j,e}, \phi) = \exp\left(\frac{S_{j,e}\theta_{j,e} - b(\theta_{j,e})}{a(\phi)} + c(S_{j,e}, \phi)\right)$$

Où :

- $\theta_{j,e}$ est le paramètre canonique (inconnu),
- ϕ représente le paramètre de dispersion, supposé connu,
- $a(\cdot)$ est une fonction non nulle dérivable sur $\mathbb{R}$,
- $b(\cdot)$ est une fonction trois fois dérivable sur $\mathbb{R}$ et dont la dérivée première est inversible,
- $c(\cdot, \cdot)$ est une fonction dérivable sur $\mathbb{R}^2$ qui ne dépend pas du paramètre canonique θ .

Lorsque que $S_{j,e}$ suit une loi appartenant à la famille exponentielle et si b est une fonction C^2 alors la moyenne et la variance de $S_{j,e}$ sont reliées à $\theta_{j,e}$ par les formules suivantes :

$$E[S_{j,e}] = b'(\theta_{j,e}) \text{ et } \text{Var}(S_{j,e}) = b''(\theta_{j,e})a(\phi)$$

Où

- $b'(\cdot)$ est la dérivée de $b(\cdot)$, correspondant à l'espérance,
- $b''(\cdot)$, sa dérivée seconde, donne une relation directe avec la variance.

Les lois classiques comme la normale, Bernoulli, binomiale, Poisson, et gamma appartiennent à cette famille exponentielle.

Dans le cas de charges de sinistres, caractérisées par une forte asymétrie et des queues lourdes, les distributions adaptées sont la loi gamma, la loi de Weibull, la loi de Cauchy ou encore la loi de Pareto. Ces lois permettent de capturer la variabilité importante observée dans les charges de sinistres, notamment pour les événements extrêmes, en fournissant une meilleure adéquation à leurs caractéristiques spécifiques. Il est important de noter que seule la loi gamma fait partie de la famille exponentielle. Si une des trois autres lois s'adapte mieux aux données, il faudra envisager une modélisation alternative au GLM.

2. Une fonction de lien canonique

En supposant que $b': R \rightarrow A$ où $A \subset R$ est bijective, l'équivalence suivante est obtenue :

$$E[S_{j,e}] = \mu = b'(\theta_{j,e}) \Leftrightarrow \theta_{j,e} = b'^{(-1)}(\mu)$$

La fonction b'^{-1} est la fonction de lien canonique et comme la relation est bijective, si $\theta_{j,e} = g(\mu)$ avec une fonction g , alors $g = b'^{(-1)}$ est la fonction de lien canonique.

Ainsi, la fonction de lien g relie l'espérance μ à une combinaison linéaire des variables explicatives :

$$b'^{(-1)}(\mu) = \theta_{j,e} = g(\mu)$$

Le choix de la fonction de lien n'est pas strictement contraint, mais il peut être privilégié pour simplifier les calculs. Toutefois, dans le cas d'un modèle additif ou multiplicatif, la fonction de lien devient obligatoire : une fonction identité est utilisée pour un modèle additif, tandis qu'une fonction logarithmique est requise pour un modèle multiplicatif.

Dans le cadre de l'estimation des charges de sinistre après le passage d'un cyclone, l'utilisation d'un modèle multiplicatif s'avère particulièrement adaptée pour plusieurs raisons fondamentales.

Tout d'abord, les charges de sinistre ne peuvent être négatives, et un modèle multiplicatif garantit que les prédictions respectent cette contrainte naturelle, contrairement à un modèle additif qui pourrait générer des valeurs négatives non réalistes.

De plus, ce type de modèle est particulièrement performant pour représenter les sinistres de grande ampleur. En amplifiant naturellement les effets des facteurs explicatifs sur les charges élevées, il offre une représentation fidèle des conséquences financières des événements cycloniques majeurs.

Par ailleurs, il garantit également une stabilité dans la modélisation des sinistres de faible ampleur, en limitant les biais associés aux petites valeurs.

Ainsi, le modèle multiplicatif avec une fonction de lien logarithmique s'impose comme une approche plus réaliste et cohérente pour estimer les charges de sinistre, en respectant les contraintes structurelles et les interactions entre les variables explicatives.

3. Une composante déterministe

Dans un GLM, l'objectif est d'estimer les coefficients $(\beta_0, \beta) \in R^{p+1}$ qui relient les variables explicatives aux charges des sinistres.

Il est alors supposé qu'il existe une fonction bijective g (la fonction de lien canonique) reliant l'espérance conditionnelle $E[S_{j,e} | X_{j,e}]$ à une combinaison linéaire des variables explicatives. Cette relation peut s'écrire comme suit :

$$\begin{aligned} \theta_{j,e}(x) &= g(\mu(x)) = g(E[S_{j,e} | X_{j,e}]) = \beta_0 + X'_{j,e}\beta \\ &\Leftrightarrow E[S_{j,e} | X_{j,e}] = g^{-1}(\beta_0 + X'_{j,e}\beta) \end{aligned}$$

Où

- β_0 est l'intercept
- $X_{j,e}$ le vecteur des variables explicatives associé à la commune j pour l'événement cyclonique e .

Dans le cas d'une fonction de lien logarithmique, $g(\mu) = \log(\mu)$, cela devient :

$$E[S_{j,e} | X_{j,e}] = \exp(\beta_0 + X'_{j,e}\beta)$$

L'objectif est d'estimer les paramètres $(\beta_0, \beta) \in R^{p+1}$ de manière à maximiser l'adéquation entre les prédictions du modèle et les charges observées. Cette estimation repose sur la méthode du maximum de vraisemblance, qui sera détaillée ci-dessous.

Estimateur du maximum de vraisemblance

Dans le contexte des modèles linéaires généralisés, il est supposé que les observations $S|X$, suivent une distribution issue de la famille exponentielle.

En particulier, dans le cas des charges observées $S_{j,e}$ pour n communes affectées par un cyclone et des variables explicatives $X_{j,e}$, ces données sont supposées indépendantes et identiquement distribuées. La log-vraisemblance, dans le cas d'une fonction de lien canonique et pour une événement donné, s'écrit :

$$l(S_{j,e}, \theta_{j,e}) = \sum_{j=1}^n \frac{S_{j,e}\theta_{j,e} - b(\theta_{j,e})}{a(\phi)} + c(S_{j,e}, \phi).$$

$$\Leftrightarrow l(S_{j,e}, \theta_{j,e}) = \sum_{j=1}^n \frac{S_{j,e}(\beta_0 + X'_{j,e}\beta) - b(\beta_0 + X'_{j,e}\beta)}{a(\phi)} + c(S_{j,e}, \phi)$$

Où

- $S_{j,e}$ représente la charge observée pour la commune j lors de l'événement e ,
- $X_{j,e}$ est le vecteur des variables explicatives pour cette même commune et cet événement,
- $a(\phi)$ et $b(\cdot)$ sont des fonctions spécifiques à la famille exponentielle.

L'estimation des paramètres par maximum de vraisemblance est indépendante du paramètre de dispersion ϕ , qui peut être déterminé ultérieurement si nécessaire. La démarche consiste à maximiser la log-vraisemblance afin d'obtenir les meilleurs estimateurs pour β_0, β , garantissant une modélisation efficace des charges de sinistres.

L'estimateur des paramètres par maximum de vraisemblance, $(\widehat{\beta}_0, \widehat{\beta})$, s'obtient en résolvant l'équation suivante :

$$(\widehat{\beta}_0, \widehat{\beta}) = \arg \max_{\beta_0, \beta} l(S_{j,e}, \theta_{j,e}).$$

$$\Leftrightarrow (\widehat{\beta}_0, \widehat{\beta}) = \arg \max_{\beta_0, \beta} \left[\sum_{j=1}^n (S_{j,e}(\beta_0 + X'_{j,e}\beta) - b(\beta_0 + X'_{j,e}\beta)) \right]$$

En résolvant les équations du premiers ordres suivants :

$$\frac{\partial l(S_{j,e}, \theta_{j,e})}{\partial \beta} = 0 \text{ et } \frac{\partial l(S_{j,e}, \theta_{j,e})}{\partial \beta_0} = 0$$

Pour plus de détail sur l'estimation par maximum de vraisemblance possibilité de se rendre en Annexe 8 : Compléments sur les modèles linéaires généralisés.

Une fois les paramètres estimés à l'aide du maximum de vraisemblance, il est nécessaire d'évaluer la qualité des modèles proposés afin de déterminer celui qui représente le mieux les données observées.

Comparaison de modèles

Le critère d'information bayésien

Le Critère d'Information Bayésien (BIC, pour Bayesian Information Criterion) est un outil d'évaluation des modèles dérivé du critère d'information d'Akaike (AIC)⁵². Le BIC introduit une pénalisation plus importante qui prend en compte non seulement le nombre de paramètres dans le modèle, mais également la taille des données utilisées. Il est défini par la relation suivante :

$$BIC = -2 \ln(l) + \ln(n) k$$

Où

- l représente la log-vraisemblance maximale du modèle,
- n est la taille de l'échantillon,
- k correspond au nombre de paramètres estimés dans le modèle.

Le BIC privilégie les modèles offrant un compromis optimal entre la qualité d'ajustement et la simplicité du modèle. Plus précisément, un modèle ayant le plus faible BIC est préféré, car cela reflète un équilibre entre une bonne représentation des données et une complexité modérée. Cela permet d'éviter un surajustement, notamment dans les cas où la taille de l'échantillon est limitée.

Erreur quadratique moyenne

Une deuxième mesure de la qualité d'ajustement est l'erreur quadratique moyenne, ou Root Mean Squared Error (RMSE). Elle est définie comme la racine carrée de la moyenne des carrés des écarts entre les valeurs observées et les valeurs prédites par un modèle. Elle sert à quantifier la dispersion des résidus, c'est-à-dire la proximité entre les données observées et la ligne de régression ou d'ajustement du modèle. Une RMSE faible reflète une meilleure précision du modèle, les prédictions étant plus proches des observations réelles. La RMSE est exprimée dans la même unité que la variable cible, ce qui la rend facile à interpréter.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (\widehat{S}_{j,e} - S_{j,e})^2}$$

Où

- $\widehat{S}_{j,e}$ désigne la charge prédite pour la commune j lors de l'événement e par le modèle,
- $S_{j,e}$ représente la charge réellement observée,
- n correspond au nombre total de communes considérées.

Test de Kolmogorov-Smirnov

Le test de Kolmogorov-Smirnov a pour hypothèse nulle que l'échantillon suit une distribution théorique donnée. La statistique de ce test est exprimée comme suit :

⁵² Le critère d'information d'Akaike (AIC) mesure la qualité relative d'un modèle statistique en équilibrant l'ajustement aux données et la complexité du modèle, en minimisant $AIC = -2 \ln(L) + 2k$, où L est la log-vraisemblance maximale et k le nombre de paramètres.

$$D = \max |F_n(x) - F(x)|$$

Où $F_n(x)$ désigne la fonction de répartition empirique et $F(x)$ la fonction de répartition théorique.

Une p -value est associée à ce test, permettant de juger de la validité de l'hypothèse nulle. Si cette p -value est supérieure à un seuil critique, 5%, l'hypothèse selon laquelle les données suivent la distribution théorique choisie n'est pas rejetée. En revanche, une p -value inférieure au seuil indique que les données s'écartent significativement de cette distribution.

Qualité d'ajustement d'un modèle

Test de significativité des variables explicatives

Pour évaluer la significativité des variables explicatives incluses dans le modèle, il est essentiel d'examiner les p -values associées aux coefficients estimés. Ces p -values permettent de tester si chaque coefficient est significativement différent de zéro, indiquant ainsi l'importance de la variable correspondante dans l'explication de la variable réponse.

Pour rappel, le modèle peut être exprimé sous la forme suivante :

$$g(E[S_{j,e} | X_{j,e}]) = \beta_0 + X'_{j,e}\beta$$

Sous l'hypothèse nulle H_0 , le test de significativité vérifie si le coefficient associé à une variable explicative β_k est significativement différent de zéro. La statistique de test utilisée est donnée par :

$$t = \frac{\widehat{\beta}_k}{SE(\widehat{\beta}_k)}$$

Où

- $\widehat{\beta}_k$ est l'estimation du coefficient et
- $SE(\widehat{\beta}_k)$ est son erreur standard.

Cette statistique suit approximativement une loi t de Student⁵³ avec $n - p - 1$ degrés de liberté, où n est le nombre d'observations et p est le nombre de variables explicatives.

Enfin, la p -value est définie comme :

$$p\text{-value} = 2 \cdot P(T > |t|)$$

Où $T \sim t_{n-p-1}$.

Une p -value inférieure à 5% permet de rejeter H_0 , suggérant que la variable X_k apporte une contribution significative au modèle.

En revanche, une p -value supérieure à ce seuil indique que la variable n'a probablement pas un effet significatif et pourrait, le cas échéant, être retirée ou regroupée avec d'autres modalités similaires.

Cette analyse est cruciale pour s'assurer que chaque variable retenue dans le modèle contribue efficacement à son pouvoir explicatif, tout en évitant une complexité inutile qui pourrait nuire à l'interprétation et à la robustesse du modèle.

⁵³ La loi de Student, ou distribution t , est utilisée pour évaluer si une estimation est significativement différente d'une valeur hypothétique (souvent zéro).

Analyse des résidus

Pour s'assurer de la validité du modèle choisi, il est essentiel d'examiner les résidus, qui permettent de vérifier si les hypothèses liées aux termes d'erreur $(\varepsilon_1, \dots, \varepsilon_n)$ sont respectées. Ces hypothèses incluent principalement deux points :

- L'indépendance des résidus : $\varepsilon_{j,e} \perp \varepsilon_{k,e}$, pour $j \neq k$,
- Une variance constante : $\text{Var}(\varepsilon_{j,e}) = \sigma^2$, aussi appelée homoscédasticité.

Les résidus bruts, définis pour chaque observation, sont calculés comme suit :

$$\xi_{j,e} = S_{j,e} - \widehat{S}_{j,e}$$

Où $S_{j,e}$ est la charge observée et $\widehat{S}_{j,e}$ la charge prédite.

Ces résidus permettent d'évaluer la qualité de l'ajustement pour chaque observation. Toutefois, leur interprétation peut être limitée puisqu'ils ne tiennent pas compte des variations de la variance entre les observations.

Pour surmonter cette limitation, les résidus standardisés de Pearson sont utilisés, et s'expriment par :

$$r_{j,e} = \frac{S_{j,e} - \widehat{S}_{j,e}}{\sqrt{\text{Var}(\widehat{S}_{j,e})(1 - h_{j,e})}}$$

Où

- $h_{j,e}$ est le terme diagonal de la matrice « chapeau » : $H = W^{1/2}X(X^TWX)^{-1}X^TW^{1/2}$,
- W est une matrice diagonale dont le j -ème terme est donné par : $w_{jj} = \frac{1}{\text{Var}(\widehat{S}_{j,e})}$

L'analyse des résidus repose sur plusieurs outils graphiques permettant de vérifier les hypothèses du modèle.

Un nuage de points est d'abord utilisé pour représenter les résidus en fonction des valeurs prédites. Une répartition aléatoire des résidus autour de zéro, sans structure apparente, est attendue afin de valider l'indépendance des résidus et l'homoscédasticité.

Ensuite, un QQ-plot est construit pour comparer les quantiles des résidus standardisés à ceux d'une distribution normale. Un alignement des points sur la diagonale indique que les résidus suivent une distribution normale.

Enfin, un histogramme permet de visualiser la distribution des résidus. Une forme en cloche de l'histogramme suggère une normalité des résidus.

En complément de ces analyses graphiques, le test de Shapiro-Wilk peut être réalisé pour évaluer la normalité des résidus de manière statistique. Ce test repose sur l'hypothèse nulle selon laquelle les résidus $(r_{j,e})_{j=1}^n$ suivent une distribution normale. La statistique de test W est définie comme suit :

$$W = \frac{(\sum_{j=1}^n a_j r_{(j,e)})^2}{\sum_{j=1}^n (r_{j,e} - \bar{r})^2}$$

Où :

- $r_{(j,e)}$ représente le j -ème résidu pour l'événement e , trié par ordre croissant (statistique d'ordre),
- $\bar{r} = \frac{1}{n} \sum_{j=1}^n r_{j,e}$ est la moyenne des résidus,
- a_j est une constante donnée par : $a_j = \frac{m^T V^{-1} e_j}{\sqrt{m^T V^{-1} m}}$ où $m = (m_1, \dots, m_n)$ est le vecteur des espérances des statistiques d'ordre d'un échantillon issu d'une loi normale standard, V est la

matrice de variance-covariance des statistiques d'ordre, et e_j est le vecteur canonique (composé de 1 à la j -ème position et de 0 ailleurs).

Une p-value supérieure à 5 % confirme cette hypothèse, renforçant ainsi la validité du modèle.

Courbe de Lorentz et indice de Gini

La courbe de Lorentz est utilisée pour analyser la répartition des charges de sinistres entre les communes après un cyclone. Pour la tracer, les communes sont classées par ordre croissant de charges prédites. Ensuite, la proportion cumulée des charges observées (axe des ordonnées) est mise en relation avec la proportion cumulée des communes nécessaires pour atteindre ces charges (axe des abscisses).

La courbe commence toujours au point $(0,0)$, représentant 0 % des communes et 0 % des charges, et se termine au point $(1,1)$, correspondant à 100 % des communes et 100 % des charges. Entre ces deux points, la courbe est convexe, car les charges cumulées augmentent avec les communes classées.

Si toutes les charges étaient réparties de manière égale entre les communes, la courbe suivrait la diagonale $x=y$. En revanche, si une seule commune concentrait toutes les charges, la courbe passerait par $(0,0)$, $(1,0)$, et $(1,1)$, indiquant une forte concentration.

L'indice de Gini, calculé à partir de l'aire entre la courbe de Lorentz et la diagonale multipliée par 2, permet de quantifier cette concentration. Un indice proche de 1 reflète une forte inégalité, où quelques communes concentrent l'essentiel des charges, tandis qu'un indice proche de 0 indique une répartition plus uniforme. Cet outil est essentiel pour évaluer la capacité du modèle à bien prédire la répartition des charges entre les communes.

Après avoir présenté les outils théoriques de comparaison des modèles, il est désormais possible d'appliquer ces notions pour évaluer les performances des modèles et estimer le coût d'un cyclone.

2.3. Estimation du coût d'un cyclone après son passage proche ou sur La Réunion

Etape préliminaire : séparation des données en échantillons d'entraînement et échantillon de test

La séparation des données en échantillons d'entraînement et de test est une étape clé pour tester la robustesse du modèle une fois qu'il sera construit. L'échantillon d'entraînement regroupe l'ensemble des cyclones à l'exception de Belal, survenu cette année, qui est réservé à l'échantillon de test. Ce choix permet de valider les performances prédictives du modèle sur un événement indépendant.

Première étape : choix de la loi

L'objectif initial est d'identifier la loi de probabilité offrant la meilleure adéquation avec les données. Les distributions retenues pour cette analyse incluent les lois Gamma, Weibull, Pareto et Cauchy, reconnues pour leur pertinence dans la modélisation de charges élevées, notamment celles associées aux sinistres cycloniques.

Loi	Log-vraisemblance	AIC	BIC
Log-Gamma	-3 680	7 363	7 370
Weibull	-3 729	7 462	7 469
Pareto	-3 690	7 384	7 391
Cauchy	-3 988	7 981	7 988

Tableau 17 : Log-vraisemblance, Score AIC et BIC des 4 lois testées pour modéliser les charges des sinistres cycloniques.

À partir du tableau ci-après, la loi log-Gamma est identifiée comme la plus appropriée pour modéliser les charges. Elle présente la plus grande log-vraisemblance et les critères AIC et BIC les plus faibles, indiquant une meilleure adéquation avec la distribution des sinistres.

Cette conclusion est renforcée par l'analyse graphique. Le QQ-plot et le PP-plot montrent que la loi Log-Gamma offre un bon ajustement aux données. Les points sont alignés sur la diagonale, ce qui indique une correspondance satisfaisante entre les quantiles observés et théoriques ainsi qu'entre les probabilités cumulées.

La loi Pareto, bien qu'elle ne soit pas un mauvais candidat, s'ajuste moins bien que la Log-Gamma, notamment sur le QQ-plot, où des écarts plus importants sont observés dans les quantiles extrêmes.

En revanche, les lois de Cauchy et de Weibull présentent des ajustements insuffisants, comme le montrent des écarts significatifs dans les QQ-Plots et PP-Plots, ainsi que des inadéquations dans la fonction de répartition cumulée empirique et théorique. Ces éléments confirment la supériorité de la loi Log-Gamma pour modéliser les charges de sinistres.

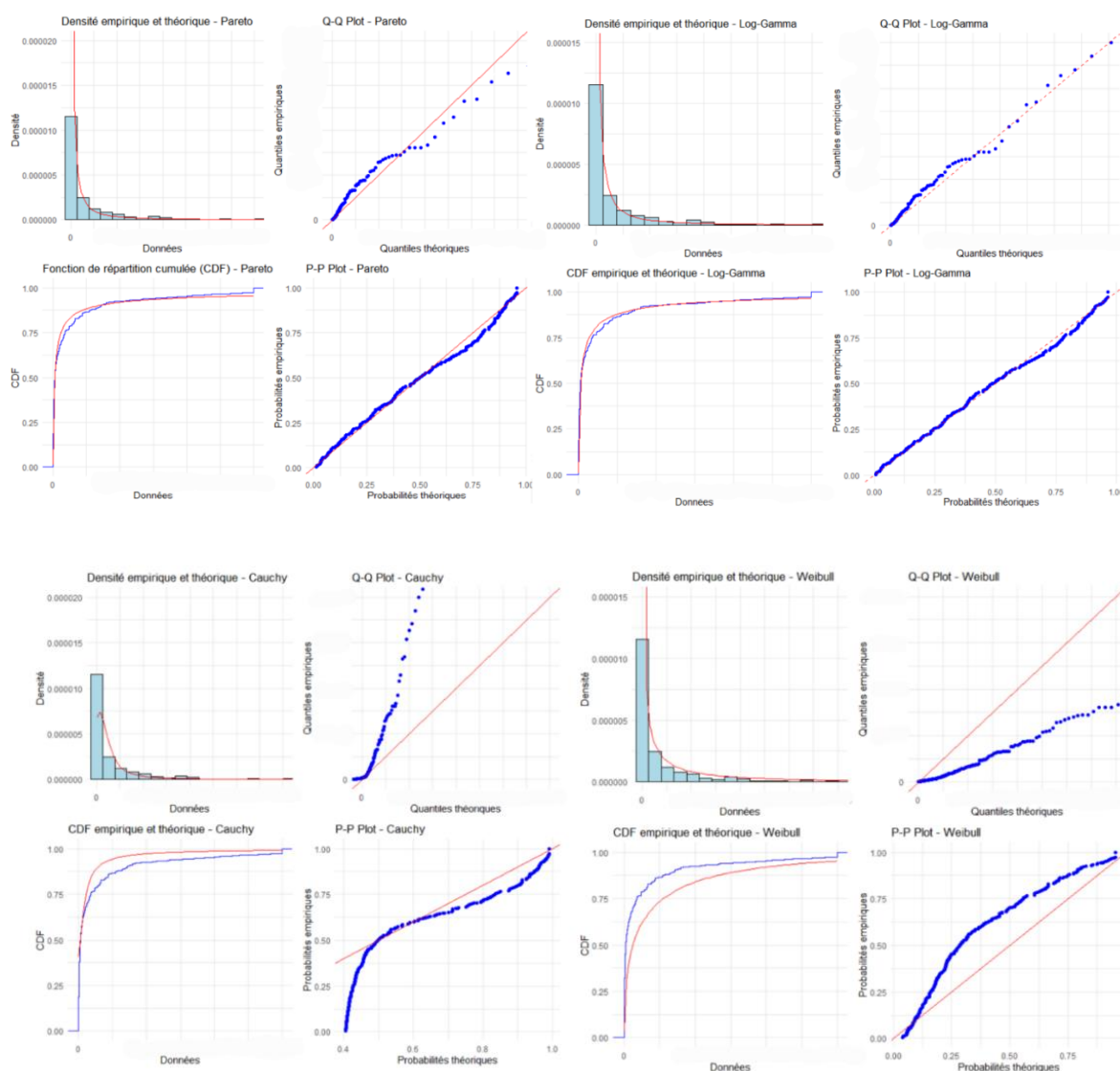


Figure 44 : Ensemble de graphes de comparaison des ajustements empiriques et théoriques des distributions Pareto, Log-gamma, Cauchy et Weibull. Certaines légendes d'axes ont été retirées pour des raisons de confidentialité. Réalisé avec R.

De plus, la p-value du test de Kolmogorov-Smirnov est supérieure à 5 %, indiquant que l'hypothèse nulle d'adéquation entre la distribution théorique (log-Gamma) et les données observées ne peut être

rejetée. Cela valide la pertinence de la loi log-Gamma pour représenter les charges de sinistres après le passage d'un cyclone.

Ainsi, la loi log-Gamma, faisant partie de la famille exponentielle, est retenue pour la modélisation des charges des sinistres et confirme que l'utilisation d'un modèle linéaire généralisé est possible.

Deuxième étape : construction d'un modèle linéaire généralisé

À la suite des étapes d'épure des variables fortement corrélées, de plafonnement des valeurs extrêmes et de regroupement des variables continues en classes abordées dans la section précédente, les variables explicatives retenues pour la modélisation sont présentées dans le tableau ci-dessous.

Variables	Nombre de classe	Classe de référence
Année	-	2022
Communes	3	Saint Denis
Vitesse de vent	6	93,6 – 110,98
Cumul de précipitation	7	0 – 58,6
SA	4	Classe SA 3

Tableau 18 : Tableau rappel des variables utilisées pour la modélisation. Réalisé avec Excel.

Pour les variables relatives aux années, 2022 a été choisie comme année de référence, car la sinistralité observée cette année-là, après mise en as-if, correspond à l'enveloppe climatique annuelle définie pour Prudence Créole. Cette enveloppe représente une estimation de la sinistralité climatique impactant le résultat de Generali, pour son propre compte et en tant que réassureur de ses filiales, sur l'horizon 2025-2027, élaborée pour anticiper les effets des risques climatiques à moyen terme.

Encadré 5 – Méthodologie utilisée pour déterminer de l'enveloppe climatique de Prudence Créole

La forte volatilité de la sinistralité climatique, en particulier à La Réunion, complique l'élaboration de l'enveloppe climatique, car elle rend difficile l'anticipation précise des événements futurs. De ce fait, elle repose sur des hypothèses simples et facilement explicables, rendant leur utilisation accessible en instance. Les principales hypothèses sont les suivantes :

1. La sinistralité climatique est principalement portée par les cyclones. Les autres périls naturels (inondations, glissements de terrain, etc.) ne sont pas pris en compte dans l'enveloppe, car leur impact financier est jugé négligeable au regard des cyclones.

2. Deux types de cyclones sont envisagés : ceux d'ampleur suffisante pour activer la réassurance et ceux qui ne le sont pas, ces derniers n'ayant pas d'impact sur le résultat de Generali.

3. Base de données limitée : L'enveloppe repose sur les sinistres cycloniques enregistrés entre 2013 et 2024. Cette période a été retenue, car les sinistres antérieurs n'étaient ni consolidés ni pleinement intégrés au moment de la mise en place de l'enveloppe.

4. Fréquence des événements : Une hypothèse forte est posée, selon laquelle un cyclone réassuré survient en moyenne tous les trois ans.

5. La sinistralité climatique est calibrée sur le coût du cyclone ayant une période de retour de trois ans dans la base (2013-2024), en l'occurrence le cyclone Batsirai (2022). Ce coût a été arrondi au million d'euros près pour simplifier son utilisation.

Ces hypothèses permettent de définir une enveloppe qui, bien que simplifiée, offre une approximation de la sinistralité attendue.

Ces travaux, réalisés dans le cadre du stage, ont permis de mieux comprendre la base des sinistres de Prudence Créole et d'en extraire des enseignements directement exploitables pour la modélisation. Ils pourront également renforcer la base de données utilisée pour le plan 2026-2028, en y ajoutant dix années supplémentaires d'études.

Les autres années sont donc utilisées uniquement pour calibrer le modèle. Lors de l'application du modèle en *back-test* ou en conditions réelles pour un futur cyclone, toutes les observations sont alignées sur l'année 2022. Autrement dit, les prédictions sont effectuées comme si toutes les données correspondaient aux conditions de 2022, indépendamment de l'année réelle des observations. Cette approche reflète les hypothèses du plan climatique, en s'affranchissant des variations interannuelles qui ne sont pas exploitées dans le modèle⁵⁴.

La fonction *gamlss* sous Rstudio permet d'obtenir une première estimation des coefficients via la méthode du maximum de vraisemblance⁵⁵.

Toutefois, pour les années de survenance, certains coefficients obtenus ne se révèlent pas statistiquement significatifs. Des regroupements ont donc été réalisés pour améliorer la pertinence du modèle. La méthode de regroupement repose sur l'identification des p-values les plus élevées. Chaque modalité associée à une p-value non significative est fusionnée avec la modalité la plus proche en termes de valeur du coefficient. Ainsi, pour n'avoir que des modalités significatives, les variables sont regroupées de la manière suivante :

Variables	Nombre de classe	Classe de référence
Année	6	2022
Communes	3	Saint Denis
Vitesse de vent	5	93,6 – 110,98
Cumul de précipitation	6	0 – 58,6
SA	4	Classe SA 3

Tableau 19 : Tableau final des variables utilisées pour la modélisation après regroupements. Réalisé avec Excel.

Cette approche permet d'optimiser la significativité des variables tout en conservant la cohérence métier et statistique des regroupements.

Les coefficients estimés pour chaque classe des variables explicatives après application des regroupements sont présentés ci-dessous :

	Classe	Coefficients du GLM
SA	Classe SA 1	-1,24
	Classe SA 2	-0,45
	Classe SA 4	0,83
Année	2013	-0,95
	2021	-1,41
	2006 - 2015 - 2024 - 2007	-0,51
	2023 - 2020 - 2017	-1,61
	2014 - 2002 - 2018	0,42
Commune	Saint Leu, Saint André, Saint Pierre, Sainte Marie, Le Port, La Possession, Saint Benoit	-0,62
	Salazie, La Plaine des Palmistes, L'Étang Salé, Saint Philippe, Les Trois Bassins, Saint Joseph, Les Avirons, Entre Deux, Petite Île, Sainte Rose, Le Tampon, Sainte Suzanne, Bras Panon, Saint Louis, Cilaos, Saint Paul	-0,66

⁵⁴ Il est difficile de prendre en compte les variations interannuelles en raison de leur imprévisibilité, liée à des fluctuations importantes en fréquence et en intensité d'une année à l'autre, ce qui rend leur modélisation peu fiable et non représentative des tendances à long terme.

⁵⁵ Les calculs sont détaillés en annexe 8.

Vent	[26 - 93,6]	-0,54
	[110,98 - 128,75]	0,61
	[128,75 - 148,79]	1,54
	[149,79 - 193,67]	2,24
Précipitation	[58,6 - 194,5]	0,37
	[194,5 - 297,49]	0,51
	[297,49 - 437,77]	0,88
	[437,77 - 634,53]	1,1
	[634,53 - 3730]	1,17
Intercept		10,5

Tableau 20 : Coefficients du GLM. Réalisé avec Excel.

Les coefficients des modalités de référence ne figurent pas dans le tableau, car ils sont fixés à zéro⁵⁶. Les autres coefficients doivent être interprétés par rapport à ces modalités de référence, ce qui permet d'évaluer l'impact relatif de chaque classe sur la variable cible.

Pour les sommes assurées, les résultats mettent en évidence une relation logique et attendue : plus la somme assurée est élevée, plus les dégâts occasionnés par un cyclone sont importants. Cette tendance s'explique par l'exposition accrue des biens à de fortes valeurs.

Concernant les années, les coefficients montrent que 2002, 2014 et 2018 ont été particulièrement destructrices. Ces résultats sont cohérents avec l'historique des cyclones majeurs survenus ces années-là, comme le cyclone Dina en 2002. En revanche, les années 2017, 2020 et 2023 présentent des coefficients faibles, traduisant une absence d'événements cycloniques significatifs. Pour rappel, le cyclone Belal, survenu en 2024, a été exclu de la base de données pour être utilisé dans l'échantillon de test lors du *back-test*. La même année, le cyclone Candice a eu un impact minime, c'est pour cela que 2024 ne semble pas marquée par une destructivité particulière jusqu'à présent. Il faut également garder à l'esprit que les années servent uniquement à construire et calibrer le modèle. Lors de l'estimation de la charge d'un nouveau cyclone, toutes les données seront alignées sur l'année 2022, comme expliqué précédemment.

Les coefficients associés aux différentes classes de vents montrent une progression nette : plus la vitesse des rafales augmente, plus les dégâts sont conséquents. Cette relation confirme le rôle primordial des vents violents dans les dommages matériels.

Enfin, les précipitations suivent une logique similaire. Les coefficients croissants des différentes classes montrent que des cumuls de précipitations importants amplifient significativement les dégâts. Ces résultats reflètent l'impact des pluies intenses sur des événements tels que les inondations ou les infiltrations d'eau dans les bâtiments.

Enfin, le modèle explicatif est comparé au modèle "null" pour renforcer l'évaluation de la pertinence des variables explicatives. Le modèle "null", qui repose uniquement sur la moyenne des données sans variables explicatives, sert de point de référence. Dans cette étude, le modèle explicatif montre une nette amélioration par rapport au modèle "null", avec une réduction du BIC de 7 472 à 7 115. Ces résultats mettent en évidence l'apport des variables explicatives pour mieux représenter la réalité des sinistres, tout en évitant la simplicité excessive d'un modèle sans prédicteurs.

⁵⁶ Dans un modèle linéaire généralisé, chaque coefficient représente l'écart entre une modalité donnée et la modalité de référence. Fixer cette dernière à zéro permet de simplifier l'interprétation des résultats en comparant directement chaque modalité à cette référence. Cela ne signifie pas que la modalité de référence n'a aucun impact, mais qu'elle sert de point de départ pour mesurer les variations relatives des autres modalités.

	Modèle « null »	Modèle construit
BIC	7 472	7 115

Tableau 21 : Score BIC du modèle Null et du modèle construit pour le mémoire. Réalisé avec Excel.

Ainsi, l'ensemble des coefficients estimés et la comparaison avec le modèle "null" confirment la cohérence avec les mécanismes physiques et financiers sous-jacents aux dégâts cycloniques, renforçant la validité du modèle construit.

Troisième étape : analyse des résidus

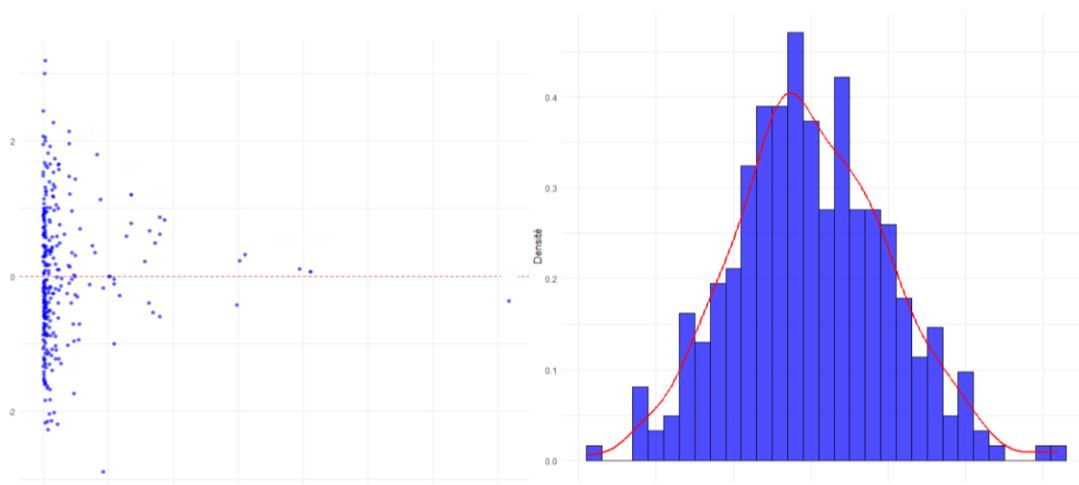


Figure 45 : Résidus en fonction des valeurs prédites à gauche et histogramme des résidus à droite. Réalisé sous R.

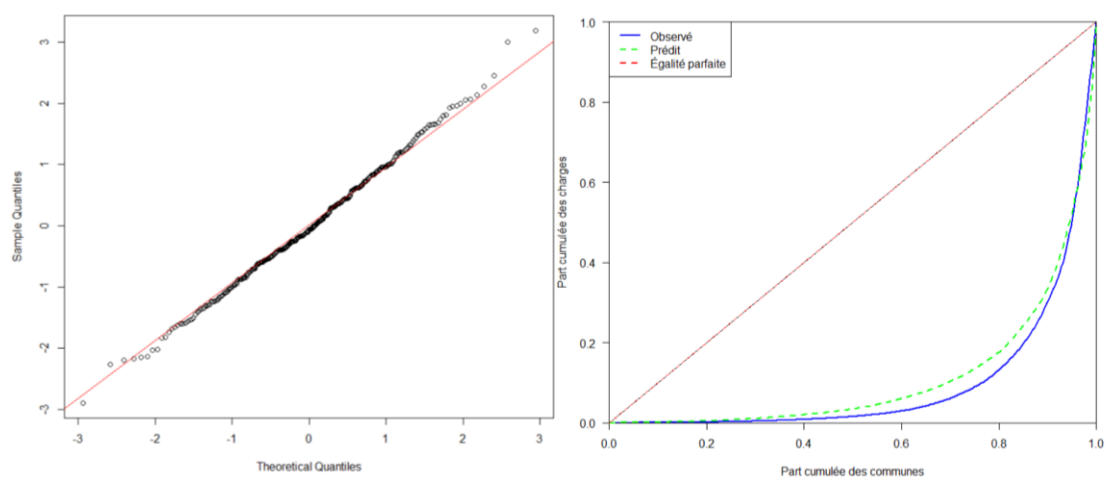


Figure 46 : QQ-plot des résidus à gauche et courbe de Lorentz à sur l'observé et le prédit à droite. Réalisé sous R.

L'analyse des résidus montre que les hypothèses du modèle sont globalement respectées.

L'histogramme des résidus standardisés, superposé à une courbe de densité normale, indique une distribution symétrique et centrée autour de zéro, cohérente avec l'hypothèse de normalité.

Le QQ-plot confirme cette observation, avec des points alignés sur la diagonale, bien que de légers écarts soient visibles aux extrémités.

Ces résultats sont corroborés par le test de normalité de Shapiro-Wilk dont la p-value supérieure à 5% permet de ne pas rejeter l'hypothèse de normalité.

Le graphique des résidus en fonction des valeurs prédites révèle une répartition homogène autour de zéro, sans structure apparente, suggérant l'absence de biais systématique. Une dispersion légèrement plus importante pour les valeurs prédites les plus élevées est observée, mais elle reste cohérente avec la nature des données assurantielles et ne remet pas en cause la validité globale du modèle.

L'analyse des résidus a confirmé la validité globale du modèle, avec des erreurs bien réparties. Pour compléter cette évaluation, l'analyse des courbes de Lorenz et des indices de Gini permet d'examiner la capacité du modèle à reproduire la concentration des charges entre les communes, en mettant en lumière sa performance dans la répartition des sinistres observés et prédits.

La courbe de Lorenz des observations montre une forte inégalité dans la répartition des charges, avec une majorité des charges concentrée sur une minorité de communes, comme en témoigne l'indice de Gini sur l'observé de 0,81.

La courbe de Lorenz des prédictions, légèrement en dessous de celle des observations, reflète une concentration légèrement moindre, avec un indice de Gini des prédictions de 0,78. Cette différence suggère que le modèle capte correctement les grandes tendances de concentration des charges, mais tend à légèrement sous-estimer les inégalités extrêmes.

Ces résultats confirment toutefois la capacité du modèle à représenter de manière fidèle la distribution des charges entre les communes, tout en laissant une marge d'amélioration potentielle pour mieux saisir les disparités les plus marquées.

Quatrième étape : back test

Le *back-test* constitue une étape clé pour évaluer la capacité du modèle à estimer les charges de sinistres pour un événement indépendant, en l'occurrence le cyclone Belal survenu en 2024. Cette étape met en application le modèle construit précédemment, permettant de prédire les charges associées à Belal en exploitant les nouvelles données disponibles, à savoir le cumul de précipitations, les rafales de vent, les sommes assurées et les communes impactées. Bien que le modèle fournisse une estimation de la charge pour chaque commune impactée, l'objectif principal, dans une optique de réassurance, est d'évaluer la coût totale associée à l'événement. Ce focus global répond à l'enjeu stratégique de cette section : développer un modèle adapté aux réalités de La Réunion et capable de compléter le retour terrain de Prudence Créole pour répondre de manière rapide, simple et fiable aux sollicitations des réassureurs.

L'année 2022 a été retenue comme paramètre dans la variable "année", car elle reflète la sinistralité climatique annuelle attendue. Ce choix permet d'obtenir des projections robustes tout en alignant les résultats du modèle sur l'enveloppe climatique.

Une particularité de cette évaluation réside dans le traitement du plafonnement. Les données extrêmes, qui avaient été écartées lors de la phase de modélisation pour éviter de rehausser artificiellement les résultats et de compromettre leur stabilité, doivent être prises en compte après coup pour ne pas sous-estimer les charges. Une fois la modélisation terminée, le résultat est donc recalibré pour réintégrer l'effet des sinistres plafonnés.

Pour garantir une cohérence entre les charges observées et celles prédites par le modèle pour l'année 2022, une méthode de recalibrage du paramètre β_0 a été appliquée. Ce recalibrage consiste à ajuster β_0 de manière que la charge estimée par le modèle pour 2022 soit égale à la charge observée pour cette même année. La formule suivante a été utilisée :

$$\beta_0^{\text{calibré}} = \frac{\text{Charge observée 2022} \times \beta_0}{\text{Charge estimée 2022}}$$

Une fois $\beta_0^{calibré}$ déterminé, ce dernier a été appliqué au modèle pour prédire la charge associée au cyclone Belal.

Pour éclairer davantage l'impact du plafonnement, l'Annexe 7 présente les principaux résultats obtenus avec un modèle construit sans plafonnement. Cette comparaison met en évidence les différences clés entre les deux approches, notamment en termes de stabilité et de précision des projections lors des back-tests.

Les résultats du *back-test* mettent en évidence une performance satisfaisante du modèle pour le cyclone Belal. La charge prédite par le modèle, après calibration, présente un écart de 0,7 % par rapport à la charge observée. Cette faible différence reflète la bonne capacité du modèle à estimer les sinistres pour un événement. Il s'agit d'une nette amélioration par rapport à la méthode des triangles, dont les estimations présentaient un delta de l'ordre de 15 % à 20 % (cf. Partie 2). Ainsi, l'objectif de fiabilisation des estimations semble avoir été pleinement atteint.

La RMSE, est de 235 556 euros. Cette valeur indique l'écart type moyen entre les charges observées et celles prédites par le modèle. Elle est relativement faible comparée aux charges totales observées et prédites pour le cyclone Belal, d'autant plus que les cyclones sont communiqués au million d'euro près lors des estimations.

Une fois la charge brute estimée, une analyse est réalisée pour déterminer ce qui pourrait être cédé à la CCR. Dans le cas de Belal, seul un enregistrement de rafale supérieur à 215 km/h a été observé sur une station située à plus de 2000 mètres d'altitude. Cette zone, caractérisée par une faible densité d'assurés, n'a donc pas été reconnue comme CatNat pour les vents. Dans ce contexte, le taux de cession moyen historique hors Dina a été appliqué conformément à la méthodologie du module dommages nets CCR expliqué en partie 2, bien que, pour des raisons de confidentialité, il ne puisse être divulgué. Tout comme dans la réalité, les cessions CCR estimées étaient négligeables.

Enfin, si les charges nettes de CCR excèdent *5 millions d'euros*, il est possible de calculer ce qui sera cédé en réassurance à Generali France, voire à Generali Italie.

Ainsi, l'application du modèle sur l'échantillon de test met en lumière la capacité du modèle à non seulement estimer les charges, mais aussi à orienter les décisions stratégiques en matière de gestion des cessions et des réassurances, consolidant ainsi sa pertinence opérationnelle.

Chapitre 3 : Limites et fiabilisation du modèle

3.1. Limites du modèle

Bien que le modèle CAT simplifié développé dans ce mémoire présente des avantages opérationnels indéniables, il comporte certaines limites qui doivent être prises en compte pour interpréter correctement ses résultats.

La première partie du modèle repose sur un pourcentage de sommes assurées impactées par un événement cyclonique sans intégrer la distribution spatiale précise des dommages ou la résilience différenciée des zones géographiques. Cette simplification peut masquer des variations importantes entre les communes, notamment en ce qui concerne la vulnérabilité des biens assurés ou les spécificités locales telles que les infrastructures ou la densité urbaine. L'intégration future de ces paramètres pourrait enrichir les analyses, tout en gardant à l'esprit que l'objectif principal de cette partie du modèle est de fiabiliser les retours terrains de Prudence Créole en s'appuyant sur des indicateurs complémentaires clairs et facilement explicables. Il est donc crucial de préserver la simplicité et la vulgarisation qui font la force de cette méthode, tout en explorant des pistes d'amélioration ciblées.

Les seuils de précipitations et de vitesses de vent obtenus via la théorie des valeurs extrêmes sont fortement dépendants de la qualité et du volume des données disponibles. Avec seulement 16 cyclones historiques dans la base d'apprentissage, il peut être difficile de garantir une robustesse suffisante des seuils retenus. Un échantillon plus large permettrait de réduire les incertitudes et d'affiner ces seuils. Il est cependant à noter qu'après le *back-testing* de Belal, les données météorologiques du cyclone ont été intégrées pour mettre à jour le modèle et que les seuils n'ont pas évolué, atténuant la limite identifiée.

En travaillant à la maille communale, le modèle ne prend pas en compte les spécificités topographiques de La Réunion, telles que les zones montagneuses ou les vallées, qui influencent pourtant significativement l'intensité des rafales et des précipitations. Construire un indice de topographie par commune aurait nécessité des traitements longs et potentiellement approximatifs, mais cela pourrait améliorer les estimations des impacts.

Bien que le modèle ait été conçu pour identifier les cyclones significatifs, son application sur des événements extrêmes reste incertaine. Ces derniers, qui représentent souvent une part disproportionnée des coûts totaux, dans notre cas uniquement le cyclone Dina, sont mal capturés en raison du faible nombre d'observations disponibles. D'où la préconisation, pour de tels cas, d'appliquer le taux de destruction du cyclone Dina en attendant de trouver une solution plus robuste. Cette alternative représente une avancée pour l'estimation des cyclones extrêmes car avant les travaux initiés par le mémoire, le seul taux de destruction extrême pour les vents tropicaux à disposition était celui d'Irma, bien trop élevé pour La Réunion.

Par ailleurs, le modèle suppose également que les relations entre les variables explicatives et les charges de sinistres restent constantes dans le temps, ce qui peut devenir problématique dans un contexte de changement climatique ou d'évolution rapide des constructions et des normes. Cette hypothèse peut introduire des biais dans les projections futures.

La modélisation, basée sur un GLM, a permis d'obtenir des résultats cohérents avec la réalité observée. Cependant, par manque de temps, d'autres approches, telles que des régressions pénalisées ou d'autres techniques de *machine learning*, n'ont pas pu être testées pour challenger la méthode actuelle. Une comparaison future avec ces alternatives pourrait renforcer encore la robustesse du modèle.

Enfin, le faible nombre d'événements disponibles dans la base rend impossible une validation croisée classique sur plusieurs événements cycloniques. Cela limite la capacité à tester la généralisabilité du modèle, bien que le *back-test* sur Belal ait montré des résultats prometteurs. L'enrichissement progressif de la base d'événements dans le futur devrait permettre de combler cette lacune.

3.2. Fiabilisation et perspectives d'amélioration

Malgré ces limites, le modèle présente un potentiel d'amélioration continue grâce à une alimentation régulière des données. En intégrant chaque année les caractéristiques des nouveaux cyclones, il serait possible de suivre avec davantage de précision l'évolution des phénomènes et d'ajuster les seuils et les relations explicatives pour les rendre plus représentatifs.

De plus, avec l'enrichissement progressif de la base de données, il deviendrait envisageable d'effectuer des validations croisées et ainsi d'améliorer encore sa fiabilité.

Avec le temps, l'augmentation de la sinistralité documentée ouvrira également la voie à une étude à partir de 2018, année marquant la possibilité de rattacher directement les sinistres aux données de portefeuille. En associant les sinistres à un numéro de client, il deviendrait possible de relier ces informations aux coordonnées géographiques des biens assurés. Cela permettrait de mieux intégrer les spécificités locales, telles que la topographie, la densité urbaine ou la résilience des infrastructures. Ce mémoire constitue une première étape pour fiabiliser une variable clé du portefeuille de cette future modélisation : la structure du bâti. Par ailleurs, les travaux d'interpolation initiés sur les vitesses de vent en rafale permettront de pouvoir étudier la vitesse de vent à une maille plus fine que la maille communale. Ces travaux seront généralisables aux cumuls de précipitation.

A terme, disposer de mesures plus nombreuses et homogènes des vitesses de vent permettrait de limiter la dépendance aux techniques d'interpolation, améliorant ainsi la précision des estimations. Par ailleurs, l'intégration de nouvelles métriques telles que la direction des vents, la pression atmosphérique ou encore la température de l'océan enrichirait l'analyse des phénomènes cycloniques, tout en prenant mieux en compte les spécificités topographiques propres à La Réunion.

De plus, première partie du modèle permet indirectement de détecter quel est le péril principal du cyclone dans son ensemble, ce qui pourra être une bonne base de travail si les périls sont mieux renseignés à l'avenir afin de fiabiliser le contenu de cette nouvelle variable.

Au-delà des données géolocalisées, des améliorations techniques pourraient également être apportées sur les données à la maille communale. L'intégration de courbes de vulnérabilité spécifiques à La Réunion, reliant directement l'intensité des aléas aux dommages observés, offrirait une représentation plus fine des impacts, particulièrement pour les sinistres extrêmes. De même, l'incorporation de paramètres physiques tels que la topographie, les caractéristiques des infrastructures ou les risques d'inondation apporterait une dimension supplémentaire au modèle, en affinant les projections des dommages. Ces ajustements, bien que nécessitant des traitements supplémentaires, renforceraient la capacité du modèle à capturer les spécificités locales des cyclones.

L'intégration du changement climatique, notamment par l'étude des scénarios futurs, pourrait par ailleurs enrichir le modèle en anticipant les évolutions à long terme de la fréquence et de l'intensité des cyclones. Cependant, cette approche nécessite des données spécifiques et des hypothèses robustes qui ne sont pas toujours disponibles à l'échelle régionale, notamment pour La Réunion. De plus, la complexité méthodologique et temporelle liée à cette intégration dépasse le cadre de ce

mémoire. Ces limites n'enlèvent rien à l'importance de cette perspective, qui pourrait être explorée dans des travaux futurs.

Enfin, Prudence Créole fournissant déjà des estimations à J+10, il sera aisé de confronter les projections du modèle à chaque nouveau cyclone et d'évaluer rapidement sa pertinence et son efficacité.

Conclusion

À La Réunion, les cyclones représentent une menace récurrente, causant des dommages plus ou moins importants aux biens et aux infrastructures, tout en perturbant la vie des habitants. Ces impacts majeurs nécessitent une gestion efficace et rapide de la part des assureurs, car, dès le passage d'un cyclone, de nombreuses parties prenantes – qu'il s'agisse des sinistrés, des autorités ou des réassureurs – attendent des réponses précises dans des délais très courts. Cette réactivité est essentielle pour estimer les coûts des sinistres. Cette estimation permet par la suite de déterminer les éventuels besoins en réassurance et ainsi de garantir la pérennité financière des compagnies d'assurance (si le cyclone est tel qu'une demande d'avance de fonds est nécessaire auprès des réassureurs par exemple).

En l'absence de modèle marché spécifiquement adapté aux caractéristiques uniques de La Réunion, Prudence Créole, filiale de Generali France, s'appuie actuellement sur des jugements d'experts et des triangles de développement basés sur les appels de déclaration de sinistres. Bien que cette approche permette d'obtenir des estimations globalement fiables, avec une marge d'erreur de l'ordre de 15 à 20 % pour le cyclone Belal qui a eu lieu en 2024, elle présente des limites en termes de robustesse et de rapidité.

L'objectif de ce mémoire était donc de développer un outil complémentaire et opérationnel pour affiner ces projections. Ce modèle visait à répondre à deux défis majeurs. D'une part, identifier, à partir des données météorologiques et des sommes assurées, les conditions dans lesquelles un événement peut être considéré comme intense afin de déterminer s'il nécessite un suivi par le service de réassurance de Generali, réassureur de sa filiale Prudence Créole. D'autre part, estimer avec précision la charge financière associée aux sinistres causés par ces événements. Ces estimations rapides et fiables sont essentielles pour renforcer la crédibilité de Prudence Créole et Generali France auprès des récessionnaires, souvent réticents à couvrir les risques des outre-mer, tout en consolidant leur position en tant qu'acteurs crédibles dans la gestion des risques cycloniques.

Les travaux réalisés se sont appuyés sur une collecte et une préparation approfondie des données disponibles. Concernant la sinistralité historique, un important travail d'agrégation a été effectué pour réunir des données issues de multiples bases, en prenant en compte les évolutions du portefeuille et l'inflation. Ce processus a permis de reconstruire les historiques de sinistres à la maille communale, maille la plus fine disponible pour l'ensemble des sinistres de la base et donc niveau d'analyse retenu pour la modélisation. Pour les données d'exposition, un effort particulier a été consacré à caractériser les types de construction des biens assurés, afin de mieux refléter les spécificités de chaque commune dans le modèle.

En ce qui concerne les données météorologiques, les relevés terrestres de Météo France des vitesses de vent en rafale et de cumul de précipitation ont été choisis pour leur disponibilité rapide, à peine quelques jours après le passage d'un cyclone, essentielle pour une modélisation rapide. Après une phase d'appropriation des données météorologiques historiques, il est apparu nécessaire d'écarter certaines stations, notamment celles situées en altitude ou dans des zones éloignées des biens assurés, afin d'assurer une meilleure représentativité des conditions réelles. La contrainte de travailler à la maille communale a ensuite imposé un traitement particulier pour adapter ces données météorologiques au modèle. Pour les précipitations, les enregistrements étaient particulièrement fournis grâce au réseau dense de stations météorologiques. Après consultation avec un expert en construction à La Réunion, il a été décidé de conserver les cumuls maximaux pour chaque commune, cette approche permettant de mieux refléter la réalité des impacts subis localement. Tandis que, pour les vents, en raison du faible nombre de stations d'enregistrement, il a été nécessaire de recourir à une interpolation. Cela a permis de garantir une continuité spatiale et de produire des estimations cohérentes avec les conditions locales, en prenant en compte les spécificités géographiques et climatiques des communes.

Les données préparées ont permis de définir des seuils clairs pour identifier les événements qui pourraient être réassurés. Grâce à la théorie des valeurs extrêmes, il a été établi qu'un événement est considéré comme déclarable auprès des réassureurs lorsque 12 % des sommes assurées du portefeuille dommages aux biens sont touchées par des rafales de vent dépassant 154 km/h et/ou des précipitations cumulées supérieures à 810 mm. Ces seuils offrent une méthode claire pour décider rapidement de la nécessité de la réassurance et ainsi communiquer les bons signaux aux différentes parties prenantes.

Puis, un modèle CAT simplifié a été construit pour estimer la charge de sinistres associée à un cyclone. Ce modèle, basé sur un modèle linéaire généralisé utilisant une loi Gamma et une fonction de lien logarithmique, a été optimisé par des regroupements en classes, des plafonnements et un recalibrage *post*-plafonnement, permettant de réduire l'influence des données extrêmes sans pour autant les négliger. Après des traitements rigoureux, la charge de sinistre cyclonique a pu être expliquée de manière cohérente en fonction des vitesses de vent, des précipitations, des sommes assurées et des communes. Le modèle a ensuite été back-testé sur le cyclone Belal survenu en 2024. La charge prédite par le modèle, après calibration, présentait un écart de 0,7 % par rapport à la charge observée. Cela démontre la capacité du modèle à estimer les sinistres avec précision, une nette amélioration par rapport à la méthode des triangles, qui avait un écart de 15 % à 20 %. Ces résultats soulignent la pertinence de l'approche adoptée et offrent à Prudence Créole un outil précieux pour affiner ses estimations de coûts, tout en renforçant sa crédibilité auprès de ces interlocuteurs.

Bien que les résultats obtenus soient très prometteurs, certaines pistes d'amélioration méritent d'être explorées pour enrichir les travaux réalisés.

Tout d'abord, les données météorologiques, bien que cruciales pour ce modèle, pourraient être enrichies par l'accès à d'autres sources de données ou à des relevés plus détaillés. Par exemple, disposer de mesures plus nombreuses et homogènes des vitesses de vent permettrait de réduire la dépendance à l'interpolation. De même, intégrer des métriques supplémentaires telles que la direction des vents, la pression atmosphérique ou la température de l'océan offrirait une vision plus complète des phénomènes cycloniques, tout en prenant mieux en compte les spécificités topographiques de La Réunion.

Pour les données de sinistralité, il serait utile de mieux identifier les causes de chaque sinistre, qu'il s'agisse de vents, d'infiltrations, d'inondations ou encore de coulées de boue. Cela permettrait d'affiner l'analyse des impacts. En outre, relier les sinistres aux biens assurés concernés permettrait de travailler à une échelle plus fine pour l'ensemble de l'étude, en passant de la maille communale à une localisation exacte.

Pour les données d'exposition, leur utilisation à la maille adresse ouvrirait des perspectives particulièrement intéressantes. Cette granularité permettrait de mieux exploiter les informations détaillées sur les biens assurés, telles que les matériaux de construction, les types d'usage (résidentiel, commercial, industriel) ... Ces données offrent une opportunité unique de différencier les vulnérabilités spécifiques de chaque bâtiment et d'affiner les estimations de risques. Par ailleurs, intégrer des caractéristiques supplémentaires, comme l'altitude ou la proximité de zones sensibles (côtes, rivières, pentes), enrichirait encore davantage l'analyse. Une telle précision permettrait de mieux refléter les spécificités locales et de capturer les facteurs contribuant aux variations de vulnérabilité d'un bien à l'autre.

Sur le plan méthodologique, le passage à des courbes de vulnérabilité par type de péril pourrait représenter une avancée majeure. Ces courbes permettraient de relier plus directement l'exposition aux sinistres et faciliteraient une approche stochastique. Cela reste toutefois conditionné à la disponibilité de données suffisamment détaillées, notamment des sinistres géolocalisés et des périls identifiés. Enfin, l'intégration de concepts théoriques plus avancés pourrait être envisagée, comme les modèles de dépendance ou les approches bayésiennes. Ces méthodes sophistiquées, bien qu'exigeantes en termes de données et d'interprétation, pourraient apporter une profondeur supplémentaire à l'analyse. Cependant, elles pourraient également rendre l'outil moins accessible pour un usage opérationnel, ce qui constituerait un compromis à évaluer.

Ces perspectives s'inscrivent dans une démarche d'amélioration continue et soulignent les opportunités de valoriser davantage ce travail. D'autant que, même si les sinistres disponibles aujourd'hui sont limités à l'échelle communale, ce mémoire a déjà permis de fiabiliser les données d'exposition et météorologiques, offrant ainsi un socle robuste pour de futures améliorations. Parmi les avancées clés, la préparation des informations sur la structure des bâtiments et l'interpolation des vents en rafale pourront être pleinement exploités si, à l'avenir, les sinistres venaient à être localisés à l'adresse.

Au-delà des aspects techniques et des outils de modélisation développés dans ce mémoire, la prévention constitue un levier essentiel pour réduire les dommages causés par les cyclones. Ce travail s'inscrit dans une démarche plus large portée par Generali France et Prudence Créole, qui vise à renforcer la résilience des assurés et à promouvoir une meilleure préparation face aux risques climatiques.

Le projet Ensemble Face aux Risques, développé par le *Generali Climate Lab*, incarne cette ambition. En 2024, l'équipe a déployé le site Ensemble Face aux Risques chez Prudence Créole. Grâce à cet outil numérique, les particuliers et les entreprises peuvent évaluer leur exposition aux risques naturels à La Réunion et Mayotte, dont le vent et les inondations, en renseignant simplement leur adresse. En retour, ils accèdent à une fiche de prévention personnalisée, leur fournissant des recommandations claires et adaptées. Ces fiches permettent d'identifier les mesures concrètes à mettre en œuvre, comme le renforcement des structures, l'aménagement de zones à risque ou l'adoption de bonnes pratiques, afin de réduire efficacement leur vulnérabilité face aux événements climatiques. Le fonctionnement d'Ensemble Face aux Risques Prudence Créole est illustré en Annexe 9. L'élaboration du site, notamment les cartes d'inondations et de vent, a bénéficié de contributions significatives réalisées dans le cadre du stage au sein du Generali Climate Lab.

Bibliographie

- 1- 001333_ *Cyclone DINA survenu à la Réunion les 22 et 23 janvier 2002* - Catnat. (s. d.). Catnat. https://catastrophes-naturelles.ccr.fr/-/001333_cyclone-dina-survenu-a-la-reunion-les-22-et-23-janvier-2002
- 2- 001889_ *Cyclone Bejisa à la Réunion en janvier 2014* - Catnat. (s. d.). Catnat. https://catastrophes-naturelles.ccr.fr/-/001889_cyclone-bejisa-a-la-reunion-en-janvier-2014
- 3- 002078_ *CYCLONE IRMA* - Catnat. (s. d.). Catnat. https://catastrophes-naturelles.ccr.fr/-/002078_cyclone-irma
- 4- 002131_ *CycloneBatsirai_LaReunion* - Catnat. (2022). Catnat. https://catastrophes-naturelles.ccr.fr/-/002131_cyclonebatsirai_lareunion
- 5- 002152_ *CycloneBelal* - Catnat. (2024). Catnat. https://catastrophes-naturelles.ccr.fr/-/002152_cyclonebelal
- 6- Aly, S. (2024). Responsable du provisionnement des cyclones chez Prudence Créole. Entretien personnel réalisé en septembre 2024.
- 7- Atia, R. (2016). *Mémoire d'actuariat : Mise en place de modèles de tarification alternatifs face à la suppression réglementaire d'une variable tarifaire en automobile*. [https://www.ressources-actuarielles.net/EXT/ISFA/1226-02.nsf/0/8438d41317cc7cffc1257f8c004ab629/\\$FILE/ATIA.pdf](https://www.ressources-actuarielles.net/EXT/ISFA/1226-02.nsf/0/8438d41317cc7cffc1257f8c004ab629/$FILE/ATIA.pdf)
- 8- Banoun, S., Colrat, A., Jagorel, Q., & Mars, G. (2020). *Le phénomène de non-assurance dans les départements et collectivités d'outre-mer*. Récupéré de https://www.igf.finances.gouv.fr/files/live/sites/igf/files/contributed/Rapports%20de%20mission/2020/2019-M-056-03_Non_assurance_Dom-Com-.pdf
- 9- Baradel, N. (s.d.). *Cours d'assurance dommage*. Récupéré de https://nicolasbaradel.fr/enseignement/ressources/cours_assurance_dommage.pdf
- 10- *BELAL a touché La Réunion le 15 janvier* | Météo-France. (s. d.). <https://meteofrance.com/actualites-et-dossiers/actualites/belal-touche-la-reunion-le-15-janvier>
- 11- Boissier, L. (2024). Responsable du pôle recherche climatique chez Generali France.
- 12- Bresson, P. (2024). Responsable du modèle interne non-vie de Generali. Entretien personnel réalisé en juin 2024.
- 13- Carrer, M., & Gaetan, C. (2022). Incorporating explanatory variables in generalized Pareto distributions for extreme value modeling. *Journal of Statistical Computation and Simulation*, 92(3), 567-589.

- 14- Comité technique CRC - Données des désordres à la réunion. (2013a). Ministère de la Transition écologique. https://www.reunion.developpement-durable.gouv.fr/IMG/pdf/Comite_technique_CRC_20juin2013_donnees_desordres_Reunion_cle26fdfd.pdf
- 15- Comité technique CRC - Données des désordres à la réunion. (2013b). Ministère de la Transition écologique. https://www.reunion.developpement-durable.gouv.fr/IMG/pdf/Comite_technique_CRC_20juin2013_donnees_desordres_Reunion_cle26fdfd.pdf
- 16- Comité technique CRC - Données des ordres à La Réunion. (2013). Ministère de la Transition écologique. https://www.reunion.developpement-durable.gouv.fr/IMG/pdf/Comite_technique_CRC_20juin2013_donnees_desordres_Reunion_cle26fdfd.pdf
- 17- Cyclone Jenny 1962 La Reunion - CCR. (s. d.-a). CCR. <https://www.ccr.fr/-/cyclone-jenny-1962-la-reunion>
- 18- Cyclones et changement climatique | Météo-France. (s. d.-a). <https://meteofrance.com/le-changement-climatique/observer-le-changement-climatique/cyclones-et-changement-climatique>
- 19- Desarthe, J. (2024). Responsable des opérations à la Mission Risques Naturels. Entretien personnel réalisé en septembre 2024.
- 20- Desarthe, J., & Moncoulon, D. (2017). Quatre siècles de cyclones tropicaux dans les départements français d’outre-mer. *La Météorologie* N° 99 - Novembre 2017. https://tirex.univ-montp3.fr/publi/meteo_2017_99_52.pdf
- 21- ÉVOLUTION DU RISQUE CYCLONIQUE EN OUTRE-MER à HORIZON 2050. (s. d.). Dans *Météo France*. Caisse Centrale de Réassurance Département R&D, Modélisation Cat & Agriculture. <https://catastrophes-naturelles.ccr.fr/documents/148935/368905/Etude+Risque+cyclonique+en+Outre-mer.pdf/cc7c7693-f82b-a2d0-0123-c7c09e7aac49?t=1592920940614>
- 22- France Assureurs. (2024, 30 janvier). *Le coût du cyclone Belal s’élève à 100 millions d’euros* [Communiqué de presse]. https://www.franceassureurs.fr/wp-content/uploads/20240129_france_assureurs_cp_tempete-belal_final.pdf
- 23- *Il y a 20 ans : le cyclone Dina frappait la Réunion* | Météo-France. (2022, 21 janvier). <https://meteofrance.com/actualites-et-dossiers/magazine/il-y-20-ans-le-cyclone-dina-frappait-la-reunion>
- 24- Inondations causées par la tempête tropicale Berguitta à La Réunion en janvier 2018 - Catnat. (2023, 13 février). Catnat. <https://catastrophes-naturelles.ccr.fr/-/inondations-causees-par-la-tempete-tropicale-berguitta-a-la-reunion-en-janvier-2018>

- 25- JAGOREL, Q., COLRAT, A., MARS, G., & BANOUN, S. (2020). LE PHÉNOMÈNE DE NON-ASSURANCE DANS LES DÉPARTEMENTS ET COLLECTIVITÉS D'OUTRE-MER. Dans *igf.finances.gouv* (N° 2019-M-056-02).
https://www.igf.finances.gouv.fr/files/live/sites/igf/files/contributed/Rapports%20de%20mission/2020/2019-M-056-03_Non_assurance_Dom-Com-.pdf
- 26- KERRY, E. (2018). 100 years of progress in tropical Cyclone research : Tropical Cyclones and Global Climate Change : A Post-IPCC Assessment (1998). *Meteorological Monographs*, 59, 15.1–15.68. https://texmex.mit.edu/pub/emanuel/PAPERS/AMS_monograph.pdf
- 27- La 1ère, O. (2023, 25 janvier). Archives d'Outre-mer : le 29 janvier 1989, le cyclone Firinga déferle sur La Réunion. *Outre-mer la 1ère*. <https://la1ere.francetvinfo.fr/archives-d-outre-mer-le-29-janvier-1989-le-cyclone-firinga-deferle-sur-la-reunion-1360134.html>
- 28- Laffin, L. (2024). Expert sinistre cyclonique, société SARETEC. Entretien personnel réalisé en juillet 2024.
- 29- Laghzali, Y. (2018). *Mémoire d'actuariat : Modélisation de la charge sinistre en assurance deux roues : Impact des sinistres graves sous Solvabilité II et optimisation des traités de réassurance*. [https://www.ressources-actuarielles.net/EXT/ISFA/1226-02.nsf/0/df3217e4d61a1d51c1258261006c2c57/\\$FILE/Memoire_Actuaire_Yassine%20LAGHZALI.pdf](https://www.ressources-actuarielles.net/EXT/ISFA/1226-02.nsf/0/df3217e4d61a1d51c1258261006c2c57/$FILE/Memoire_Actuaire_Yassine%20LAGHZALI.pdf)
- 30- LA PRÉVENTION DES CATASTROPHES NATURELLES PAR LE FONDS DE PRÉVENTION DES RISQUES NATURELS MAJEURS. (2023). Ministère de la transition écologique et de la cohésion des territoires. https://catastrophes-naturelles.ccr.fr/documents/148935/1306878/20240424_Rapport-region-OUTRE-MER_HD.pdf/2b927834-c9e4-c334-5c31-06d33d7d0033?t=1715848810591
- 31- LANGRENEY, T., LE COZANNET, G., & MERAD, M. (2023). ADAPTER LE SYSTÈME ASSURANTIEL FRANÇAIS FACE À L'ÉVOLUTION DES RISQUES CLIMATIQUES. https://www.ecologie.gouv.fr/sites/default/files/documents/Rapport_final_Mission-assurance_climat.pdf?v=1712160459
- 32- *Le cyclone tropical BELAL*. (2024). Météo France. <https://meteofrance.re/fr/actualites/le-cyclone-tropical-belal>
- 33- *Le cyclone intense BATSIRAI*. (2022, 14 février). meteofrance.re. <https://meteofrance.re/fr/actualites/le-cyclone-intense-batsirai>
- 34- *Les dangers des cyclones*. (s. d.). http://www.meteo.fr/temps/domtom/La_Reunion/webcmrs9.0/francais/education/danger/danger.html
- 35- *Les RECORDS en matière de CYCLONES TROPICAUX*. (s. d.). http://www.meteo.fr/temps/domtom/antilles/public/cyclone/tout_cyclone/records.htm

- 36- Louapre, D. (2014, 3 mars). *La mystérieuse équation de Navier-Stokes – Science étonnante*. <https://scienceetonnante.com/2014/03/03/la-mysterieuse-equation-de-navier-stokes/>
- 37- *Meteo.data.gouv.fr*. (s. d.-b). <https://meteo.data.gouv.fr/form>
- 38- Meskani, Z. (2018). *Mémoire d'actuariat : Modélisation de la charge du risque cyber des entreprises à partir des données externes*. <https://www.institutdesactuaire.com/docs/mem/389301a7d3623aa74cec3465fbf85c3c.pdf>
- 39- *Multirisque habitation - Dispositions générales*. (2022). Prudence créole - GENERALI. https://prudencecreole.com/sites/default/files/2022-06/20220506_Prudence%20creole_DG_MRH_HD.pdf
- 40- OBSERVATOIRE DE LA QUALITÉ DE LA CONSTRUCTION. (2013). Sycodés régional ILE DE LA RÉUNION. Dans *reunion.developpement-durable.gouv.fr*. https://www.reunion.developpement-durable.gouv.fr/IMG/pdf/Comite_technique_CRC_20juin2013_donnees_desordres_Reunion_cle26fdfd.pdf
- 41- OpenAI. (2024). *ChatGPT* (version GPT-4.0) [Modèle d'intelligence artificielle générative assistée]. <https://chat.openai.com/> (pour la reformulation)
- 42- *Principe de l'algorithme des k-means*. (s. d.). ResearchGate. https://www.researchgate.net/figure/Principe-de-lalgorithme-des-k-means_fig114_346079582
- 43- *Projections macroéconomiques – Juin 2024 | Banque de France*. (s. d.). Banque de France. <https://www.banque-france.fr/fr/actualites/projections-macroeconomiques-juin-2024>
- 44- Prudence Créole. (2024). Rapport sur la solvabilité et la situation financière - exercice 2023. Dans *Generali.fr*. https://www.generali.fr/sites/default/files-d8/2024-04/RSSF_Prudence_Creole_2023.pdf
- 45- Robert, C. Y. (2024). *Extreme Value Theory* [Cours ENSAE, 3A 2024].
- 46- Wilson, K. M., Baldwin, J. W., & Young, R. M. (2022). Estimating Tropical Cyclone Vulnerability : A Review of Different Open-Source Approaches. *Hurricane Risk/Hurricane Risk*, 255-281. https://doi.org/10.1007/978-3-031-08568-0_11

Table des figures

Figure 1 : Structure d'un cyclone. Source : Météo France.	8
Figure 2 : Carte des Appellations Régionales des Cyclones. Carte réalisée à l'aide du logiciel Canva.	8
Figure 3 : Carte de l'océan Indien sud. Carte réalisée avec Canva.	9
Figure 4 : Chronique et sévérité des cyclones de La Réunion et Mayotte de 1900 à 2014. Source : CCR.	9
Figure 5 : Classification des phénomènes cycloniques en fonction de la zone géographique. Source : Météo France.	10
Figure 6 : Trajectoires des cyclones dans le bassin indien sud depuis 2002.	13
Figure 7 : Survenance d'un cyclone lors d'une année sans activation du Stop-Loss. Source : Formation réassurance de Generali France.	22
Figure 8 : Survenance d'un cyclone lors d'une année avec activation du Stop-Loss. Source : Formation réassurance de Generali France.	22
Figure 9 : Structure de réassurance simplifiée pour les cyclones de Prudence Créole. Source : documentation de la réassurance de Generali France.	23
Figure 10 : Sinistralité climatique 2024 de Generali France y compris acceptation en réassurance. Source : données internes.	23
Figure 11 : Structure de réassurance détaillée de Prudence Créole.	25
Figure 12 : Deux illustrations de failles de construction pouvant provoquer des infiltrations d'eau. Source : Ootravau, bricozone.	27
Figure 13 : Communication de Prudence Créole suite au passage du cyclone BELAL.	28
Figure 14 : Table de correspondance permettant d'obtenir la branche.	33
Figure 15 : Évolution de l'indice FFB (2002-2024). Réalisé sur Excel à partir des données ffbatiment.	34
Figure 16 : Évolution de l'indice Risque Industriel (2002-2024). Réalisé sur Excel à partir des données assurances-gesco.	34
Figure 17 : Evolution de la charge de sinistre une fois l'évolution du portefeuille, l'inflation et la réforme CatNat prises en compte. Réalisé avec Excel.	36
Figure 18 : Exemple d'agrégation des données par commune	41
Figure 19 : Carte des stations météorologiques retenues et exclues pour la région de Saint-Denis. Carte réalisée avec ArcGIS.	43
Figure 20 : Carte des stations météorologiques de La Réunion ayant enregistrées des vitesses de vents lors de cyclones. Réalisé sous python grâce au package ArcPy.	44
Figure 21 : Interpolation des rafales de vents lors du cyclone Bêjisa grâce à une interpolation IDW. Carte réalisée sur python à l'aide de la librairie ArcPy.	45
Figure 22 : Carte des stations météorologiques de La Réunion ayant enregistrées des cumuls de pluie lors de cyclones. Réalisé sous python grâce au package ArcPy.	46
Figure 23 : Comparaison des vitesses de rafales enregistrées lors du cyclone Batsirai : carte de Météo France (à gauche) et des données interpolées (à droite). Carte de droite réalisée avec ArcGIS.	47
Figure 24 : Comparaison des cumuls de précipitations enregistrées lors du cyclone Batsirai : carte de Météo France (à gauche) et données de notre base (à droite). Carte de droite réalisée avec ArcGIS.	47
Figure 25 : Schéma du modèle CAT simplifié pour l'estimation du coût d'un cyclone. Réalisé avec PowerPoint.	50
Figure 26 : Mean Excess Plot pour les cumuls de précipitations. Réalisé sous R.	60
Figure 27 : Graphe de Pickand pour les cumuls de précipitations. Réalisé sous R.	61
Figure 28 : Graphe de Hill pour les cumuls de précipitations. Réalisé sous R.	61
Figure 29 : Graphique de stabilité pour les cumuls de précipitations. Réalisé sous R.	62
Figure 30 : Graphe des P-value du test de Villasenor-Alva et Gonzalez-Estrada en fonction des cumuls de précipitations. Réalisé sous R.	62
Figure 31 : QQ-Plot des dépassements de seuils (avec un seuil à 810mm) contre une GPD. Réalisé sous R.	63
Figure 32 : Proportion des sommes assurées impactées par les précipitations supérieures à 810 mm lors des événements cycloniques. Réalisé avec Excel.	64
Figure 33 : Mean Excess Plot des rafales de vents. Réalisé sous R.	65

Figure 34 : Echelle Saffir-Simpson croisée par la Noaa avec une description des dommages potentiels pour différents types d'enjeux. Source : Taylor et al. (2010).	65
Figure 35 : QQ-plot des rafales de vents. Réalisé sous R.	66
Figure 36 : Proportion des sommes assurées impactées par des rafales à plus de 154 km/h lors des événements cycloniques. Réalisé avec Excel.	67
Figure 37 : Proportion des sommes assurées impactées par des rafales à plus de 154km/h ou des cumuls supérieurs à 810mm lors des événements cycloniques. Réalisé avec Excel.	68
Figure 38 : Arbre de décision élagué pour déterminer le seuil de sommes assurées impactées nécessaire pour déclencher la réassurance. Réalisé sous R et retravaillé sous Paint pour ne pas communiquer le réel. Les résultats de l'arbre sont en revanche les bons.	70
Figure 39 : Fonctionnement de l'algorithme des K-means dans le cas k=2. Source : researchgate.	74
Figure 40 : Partitionnement des rafales de vent en 6 classes par K-means. Réalisé sous R.	75
Figure 41 : Regroupement des communes de La Réunion selon la proportion de sommes assurées construites en béton. Réalisé avec ArcGIS.	77
Figure 42 : Graphe des indices de Cramer pour les variables catégorielles. Réalisé sous R.	78
Figure 43 : ME plot des charges de sinistres sur l'échantillon d'entraînement. Réalisé sous R.	80
Figure 44 : Ensemble de graphes de comparaison des ajustements empiriques et théoriques des distributions Pareto, Log-gamma, Cauchy et Weibull. Certaines légendes d'axes ont été retirées pour des raisons de confidentialité. Réalisé avec R.	89
Figure 45 : Résidus en fonction des valeurs prédites à gauche et histogramme des résidus à droite. Réalisé sous R.	93
Figure 46 : QQ-plot des résidus à gauche et courbe de Lorentz à sur l'observé et le prédit à droite. Réalisé sous R.	93
Figure 47 : Evolution de la charge de sinistre une fois l'évolution du portefeuille, l'inflation et la réforme CatNat prises en compte. Réalisé avec Excel.	111
Figure 48 : Vitesse des rafales de vents à gauche et cumuls des précipitations à droite lors du cyclone Batsirai en 2022. Réalisé avec ArcGIS.	113
Figure 49 : Schéma du modèle CAT simplifié pour l'estimation du coût d'un cyclone. Réalisé avec PowerPoint.	114
Figure 50 : Proportion des sommes assurées impactées par des rafales à plus de 154km/h ou des cumuls supérieurs à 810mm lors des événements cycloniques. Réalisé avec Excel.	115
Figure 51 : Tableau du nombre de classe optimale des variables non gardées dans le mémoire.	131
Figure 52 : période de retour des cyclones selon l'approche historique avant la survenance d'Irma et Maria. Source : CCR.	133
Figure 53: histogramme des résidus sans plafonnement à gauche et avec à droite. Réalisé sous R.	135
Figure 54 : courbe de Lorentz avec plafonnement à gauche et sans à droite. Réalisé sous R.	136

Table des tableaux

Tableau 1 : Records mondiaux de précipitations. Source : Meteo.fr.	11
Tableau 2 : Historiques de CoR de Prudence Créole depuis 2018. Source : Rapports RSSF de Prudence Créole.	20
Tableau 3 : Tableau de l'activité cyclonique depuis 2002 et impact sur le portefeuille de Prudence Créole.	24
Tableau 4 : Tableau de variables reçues pour la base de sinistralité historique. Source : Prudence Créole.	31
Tableau 5 : Tableau illustratif des saisies manuelles des communes. Source : base de données de Prudence Créole.	32
Tableau 6 : Coefficients de revalorisation AS-IF par année pour la branche MRH. Réalisé avec Excel.	35
Tableau 7 : Exemple d'analyse des valeurs manquantes pour les variables adresse et coassurance et code organisation. Source : données internes.	37
Tableau 8 : Tableau synthétique sur la qualité des données. Source : données internes.	38
Tableau 9 : Tableau de catégories de type de construction.	40
Tableau 10 : Exemple d'application de la nouvelle classification des types de construction.	40
Tableau 11 : Tableau récapitulatif des seuils de cumuls de précipitations.	63
Tableau 12 : Tableau récapitulatif des variables disponibles pour la modélisation à la suite des traitements de la partie 2. Réalisé avec Excel.	72
Tableau 13 : Tableau du nombre de classe de vitesse de vents. Réalisé avec Excel.	75
Tableau 14 : Tableau des classes de vitesse de vents. Réalisé avec Excel.	75
Tableau 15 : Tableau du nombre de classe de cumul de pluie. Réalisé avec Excel.	76
Tableau 16 : Tableau des classes de sommes assurées. Réalisé avec Excel.	76
Tableau 17 : Log-vraisemblance, Score AIC et BIC des 4 lois testées pour modéliser les charges des sinistres cycloniques.	88
Tableau 18 : Tableau rappel des variables utilisées pour la modélisation. Réalisé avec Excel.	90
Tableau 19 : Tableau final des variables utilisées pour la modélisation après regroupements. Réalisé avec Excel.	91
Tableau 20 : Coefficients du GLM. Réalisé avec Excel.	92
Tableau 21 : Score BIC du modèle Null et du modèle construit pour le mémoire. Réalisé avec Excel.	93
Tableau 22 : Tableau de catégories de type de construction.	112
Tableau 23 : Tableau de toutes les variables initialement fournies par Prudence Créole dans la base de données d'exposition. Source : données internes.	129
Tableau 24 : Base de données à l'issue du module Vulnérabilité du modèle CAT simplifié.	130
Tableau 25 : Tableaux des tentatives de classification de la variable rafale de vent.	131
Tableau 26 : Score BIC du modèle avec et sans plafonnement. Réalisé avec Excel.	134
Tableau 27 : Coefficients du GLM avec et sans plafonnement. Réalisé avec Excel.	135
Tableau 28 : Résultat du back-test avec et sans plafonnement. Réalisé sur Excel.	136

Glossaire des acronymes

BDM - Bris de machine
CatNat - Catastrophes naturelles
CCR - Caisse Centrale de Réassurance
CIE - Code compagnie (exemple : P pour Prudence Créole)
CN / HCN - Catastrophe Naturelle / Hors Catastrophe Naturelle
CoR - Ratio combiné net
CPI - Code produit identifié
CSD - Catégorie statistique détaillée
DAB - Dommages Aux Biens
EVT - Événement
FME - Fonction Moyenne des excès
GLM - Modèle linéaire généralisé
GMRH - Generali Multirisque Habitation
INSEE - Institut National de la Statistique et des Études Économiques (Code de commune)
LoB - Line of Business
MRH – Multirisque Habitation
MRC – Multirisque Commerce
ODR - Objet de risque
PC - Prudence Créole
PE - Pertes d'exploitation
POT - Peaks Over Threshold ou méthode des excès au-delà d'un seuil
QP - Quote-Part
RC - Responsabilité civile
RI - Risque industriel
RUN - Réunion (code département)
SA - Sommes assurées
SL - Stop-Loss
S/P - Sinistres / Primes
TOC - Tempêtes, ouragans, cyclones
TVE - Théorie des valeurs extrêmes
XS - Excess of Loss (Excédent de Sinistres)

Note de synthèse

Contexte

Située dans une zone propice à la formation de cyclones tropicaux, La Réunion subit régulièrement des événements climatiques d'intensité variable, laissant des impacts durables sur l'île. Des cyclones emblématiques comme Jenny (1962), Hyacinthe (1980), Firinga (1989), Dina (2002), et Gamède (2007) ont causé d'importants dégâts matériels et humains. Le cyclone Dina, souvent considéré comme le plus dévastateur de ces trois dernières décennies, s'est distingué par des rafales de vent atteignant 300 km/h, accompagnées de plus de 28 heures consécutives de vents violents. Cette durée exceptionnelle a causé des destructions massives, notamment sur les infrastructures électriques et de communication, et des inondations importantes. Le cyclone Gamède de 2007, quant à lui, s'est caractérisé par des précipitations hors normes, atteignant près de 4 000 mm en seulement trois jours, entraînant des crues et des glissements de terrain ravageurs. Ces événements soulignent que la dangerosité des cyclones réside autant dans la force des vents que dans les précipitations intenses, renforçant la nécessité de prendre en compte ces deux aléas conjointement.

Bien que tous les cyclones n'aient pas la même intensité, Prudence Créole, filiale de Generali France à La Réunion, voit son activité directement impactée par ces événements. Lors des années cycloniques, son ratio combiné augmente significativement, nécessitant parfois un recours à la réassurance. Prudence Créole bénéficie de deux niveaux de couverture climatique via des réassureurs. La première, souscrite auprès de la Caisse Centrale de Réassurance (CCR), inclut deux traités successifs pour couvrir les catastrophes naturelles (CatNat) : une quote-part de 50 %, suivie d'une couverture en stop-loss consolidée par Generali France. Cette dernière s'active lorsque le ratio sinistres/primes (S/P) de la branche dommages aux biens dépasse 250 % après application des quotes-parts de Generali France consolidé. En complément, Generali France prend en charge les sinistres climatiques postérieurs aux cessions effectuées à la CCR. Cette prise en charge inclut les sinistres reconnus comme CatNat nets de réassurance CCR, ainsi que les sinistres climatiques exclus de cette couverture, grâce à un traité de type 120 xs 5, dont une partie est rétrocédée aux réassureurs de Generali France. Cette structure permet à Prudence Créole de limiter sa sinistralité à un montant maximal de 5 millions d'euros par cyclone (*modulo* les primes de REC). En revanche, cela crée de la sinistralité climatique supplémentaire pour Generali France. A titre d'exemple, en 2024, les acceptations des cessions du cyclone Belal de Prudence Créole vers Generali France sont le deuxième événement climatique le plus coûteux pour Generali France (à la vision octobre 2024).

Les relations avec les réassureurs de Generali France sont par ailleurs devenues complexes en raison d'une perception des territoires d'outre-mer comme des zones particulièrement risquées. Tous les départements et territoires ultramarins sont souvent perçus comme homogènes, alors qu'ils présentent des risques climatiques et économiques très différents. Cette perception généralisée, amplifiée depuis le cyclone Irma en 2017, a également été exacerbée par les récentes émeutes en Nouvelle-Calédonie et aux Antilles.

Si Generali France dispose de modèles performants pour les principaux périls en France hexagonale, tels que l'inondation ou la sécheresse, aucune modélisation spécifique aux cyclones à La Réunion n'a encore été développée.

Par ailleurs, les modèles de marché adaptés à l'île sont quasi inexistants, en raison de son isolement géographique, de la rareté des données historiques, et de la complexité topographique amplifiant les aléas. Face à ces défis et à l'intensification des cyclones liée au changement climatique, la conception d'un modèle cyclonique dédié apparaît désormais incontournable pour mieux anticiper et gérer ces risques spécifiques. Ce mémoire s'articule alors autour de deux questions centrales : Comment identifier rapidement si un cyclone nécessite une réassurance pour Prudence Créole ? Et comment estimer les coûts associés à ces événements pour Prudence Créole et Generali France ?

Ces deux objectifs visent à approfondir la compréhension des risques cycloniques tout en fournissant des indicateurs compréhensibles afin de faciliter leur utilisation dans un cadre opérationnel. En répondant à ces enjeux, ce travail contribue également à renforcer la confiance des réassureurs, un levier indispensable pour négocier des conditions de couverture optimales.

Pour construire un tel modèle il est indispensable de s'appuyer sur un travail approfondi concernant les données disponibles qui doivent être soigneusement consolidées et préparées pour garantir la robustesse et la pertinence des analyses.

Données de sinistres

Les données de sinistralité historique utilisées dans cette étude ont été fournies par Prudence Créole et couvrent une période allant de 2002 à 2024. Cette base regroupe les sinistres liés aux cyclones ayant impacté l'île et offre un aperçu détaillé des charges financières associées, des caractéristiques des sinistres et des contrats concernés.

Pour les besoins de la modélisation, seules les variables essentielles et suffisamment renseignées ont été retenues. Cela inclut notamment les informations contractuelles (numéros d'identification, code postal, produit d'assurance) et les données relatives aux sinistres (date de survenance, montant des charges, identification des sinistres relevant du régime CatNat). En revanche, certaines variables, comme le code des garanties ou le péril principal, ont été exclues en raison de leur faible qualité ou de leur incohérence.

Pour garantir leur pertinence et leur homogénéité dans le temps, ces données ont fait l'objet de travaux approfondis de mise "AS-IF". Cette démarche consiste à réajuster les charges des sinistres à une vision actuelle, en tenant compte de l'évolution des prix (via des coefficients d'inflation basés sur les indices FFB et RI) ainsi que des changements dans le portefeuille assuré. Cette étape permet de revaloriser les cyclones à aujourd'hui en neutralisant les effets des variations temporelles et assurant la comparabilité des données d'exposition actuelle. Finalement, les sinistres éligibles au régime CatNat, ont été majorés de 6,5 % pour une meilleure prise en compte de la réforme CatNat.

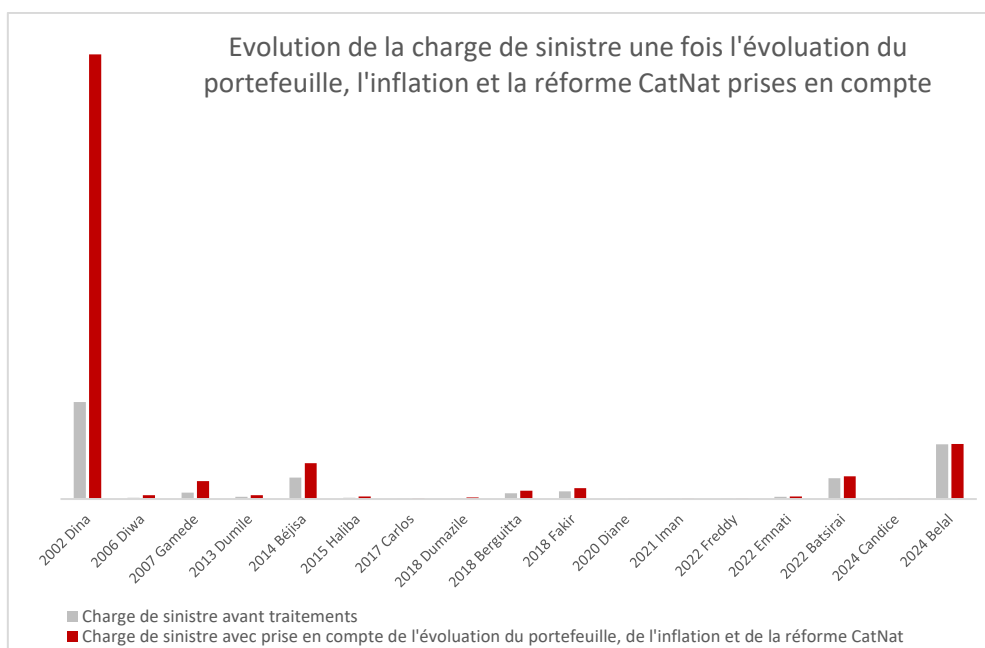


Figure 47 : Evolution de la charge de sinistre une fois l'évolution du portefeuille, l'inflation et la réforme CatNat prises en compte. Réalisé avec Excel.

En parallèle, les sinistres ont été rattachés à la maille communale en raison des limitations des anciennes bases de données, comme celles des cyclones Dina et Gamède, où les adresses précises n'étaient pas disponibles ou utilisables. Bien que les données à partir de 2018 permettent un

rattachement plus précis des sinistres aux données d'exposition, utiliser uniquement ces données aurait conduit à exclure les cyclones majeurs du passé, ce qui n'était pas envisageable. Le choix de la maille communale a ainsi permis de maintenir une cohérence globale dans l'analyse tout en contournant ces limitations.

Bien que la mise en AS-IF permette de revaloriser les données pour refléter les conditions actuelles, la dépendance aux données historiques peut restreindre la capacité à intégrer pleinement les évolutions futures, telles que l'intensification des cyclones liée au changement climatique ou les transformations socio-économiques, comme l'urbanisation ou l'amélioration des normes de construction. De plus, l'utilisation de la maille communale, bien qu'adaptée aux contraintes des données disponibles, ne permet pas de capturer les spécificités topographiques de La Réunion, comme les variations liées aux zones montagneuses ou aux vallées. Ces particularités influencent pourtant l'intensité des rafales et des précipitations. Intégrer un indice de topographie aurait nécessité des traitements longs et complexes, mais cela reste une piste intéressante pour affiner les estimations futures.

Données d'exposition

Les données d'exposition regroupent l'ensemble des informations relatives aux biens assurés, essentielles pour évaluer l'exposition financière de Prudence Créole en cas de sinistre. Issues du portefeuille dommages aux biens, elles incluent des détails tels que la branche d'assurance, le type de bâtiment, le mode d'occupation (propriétaire, locataire, etc.), les caractéristiques structurelles (construction en béton, etc.), les sommes assurées, ainsi que certains éléments contractuels comme les franchises et les taux de coassurance.

Pour préparer ces données à la modélisation, des traitements de nettoyage et de consolidation ont été nécessaires. Une attention particulière a été portée à la variable "type de construction", qui était la seule à présenter un niveau de complétude inférieur à 90 %. Cette lacune s'explique par une mauvaise transposition des données initiales de Prudence Créole. Alors que celles-ci contenaient 57 modalités très détaillées pour décrire les types de construction, cette granularité fine rendait les données difficiles à exploiter dans un cadre de collecte de données harmonisée avec les autres entités de Generali France.

Pour surmonter cette complexité, une nouvelle classification a été créée, regroupant les constructions en quatre grandes catégories, inspirées des méthodologies développées par RMS. Cette simplification, tout en réduisant la complexité des données, permet de mieux refléter les caractéristiques de vulnérabilité des bâtiments, notamment leur résistance aux vents violents. Les types de construction ainsi reclassifiés sont présentés dans le tableau ci-dessous, offrant une base robuste pour intégrer ces données dans la modélisation des impacts cycloniques.

Nom	Niveau de risque	Commentaire
Structure tout béton	Faible	Bâtiments dont la structure principale et le toit entièrement en béton offrent une grande résistance aux vents violents typiques des cyclones, mais aussi aux infiltrations et inondations.
Structure béton avec toiture différente	Intermédiaire	Ces bâtiments ont une structure en béton mais avec une toiture en matériaux plus légers, comme des tôles métalliques ou des matières plastiques. Les matériaux plus légers utilisés pour le toit peuvent être plus vulnérables aux infiltrations et aux détériorations lors des fortes rafales ou en cas de pluie intense. Les toitures en métal, notamment, sont plus sensibles aux fuites d'eau.
Structure acier et toiture autre que béton	Important	Les bâtiments avec une ossature en acier et une toiture en matériaux légers (comme les tôles ou les matières plastiques) sont particulièrement vulnérables aux cyclones. Ces toits peuvent être facilement endommagés par les vents violents, ce qui expose la structure en dessous à des infiltrations d'eau, aggravant ainsi les risques d'inondation à l'intérieur.
Inconnu	Important	Ils peuvent présenter des vulnérabilités inconnues face aux vents, inondations et infiltrations, ils seront donc considérés comme dangereux.

Tableau 22 : Tableau de catégories de type de construction.

En complément, des traitements ont été mis en œuvre pour rendre ces données exploitables à la maille communale. Les sommes assurées, initialement disponibles à des mailles très fines (comme les adresses précises), ont été agrégées par commune. Cette étape était essentielle pour relier les données d'exposition aux données de sinistralité elles-mêmes structurées à ce niveau géographique.

Données météorologiques

Les données climatiques exploitées dans ce mémoire proviennent des bases open source de Météo France et couvrent la période de 2002 à 2024. Ces informations, enregistrées par les stations météorologiques de La Réunion, incluent des mesures quotidiennes des vitesses de vent et des cumuls de précipitations.

Pour garantir leur pertinence dans le cadre de la modélisation, des traitements spécifiques ont été réalisés. Tout d'abord, certaines stations situées à plus de 1 000 mètres d'altitude ont été exclues. Ces stations, bien qu'enregistrant souvent des rafales de vent particulièrement élevées, ne reflètent pas les conditions météorologiques des zones habitées où se concentrent les biens assurés. Leur prise en compte aurait artificiellement amplifié l'intensité des aléas dans des zones peu exposées économiquement.

En raison du nombre limité d'enregistrement des données de vent lors des différents événements, une interpolation spatiale s'est avérée nécessaire. Cette méthode permet d'estimer les vitesses de vent pour des points où aucune station n'est présente. La méthode des distances inverses pondérées (IDW) a été utilisée pour cette interpolation, et la moyenne des vitesses de vent par commune a été privilégiée. Ce choix offre une représentation plus fidèle de l'exposition réelle des assurés, évitant de surestimer les aléas en se basant sur des valeurs maximales enregistrées dans des zones peu habitées.

Enfin, pour les précipitations, les cumuls maximaux sur la durée de l'événement ont été agrégés à la maille communale. Ce choix méthodologique repose sur plusieurs considérations. Tout d'abord, les stations situées en altitude, susceptibles de surestimer les précipitations pour certaines zones, ont été exclues, ce qui rend le cumul maximal par commune plus représentatif des conditions extrêmes réellement subies par les assurés. Ensuite, bien que l'identification précise du péril à l'origine des sinistres ne soit pas systématiquement possible pour les événements historiques, des échanges avec Lionel LAFFIN, expert construction à La Réunion chez Saretec, ont mis en évidence que les infiltrations liées aux précipitations représentent une part significative des dommages survenus ces dernières années à La Réunion.

Ces traitements aboutissent à deux indicateurs clés par cyclone et par commune : une vitesse de rafale moyenne et un cumul maximal de précipitations sur la durée de l'événement. Ces résultats sont visualisés dans les cartes ci-dessous, qui illustrent respectivement les vitesses de rafales et les cumuls de précipitations pour l'ensemble du territoire réunionnais lors du cyclone Batsirai de 2022.

Données de vent interpolées (à gauche) et les cumuls de précipitations (à droite) lors du cyclone Batsirai de 2022

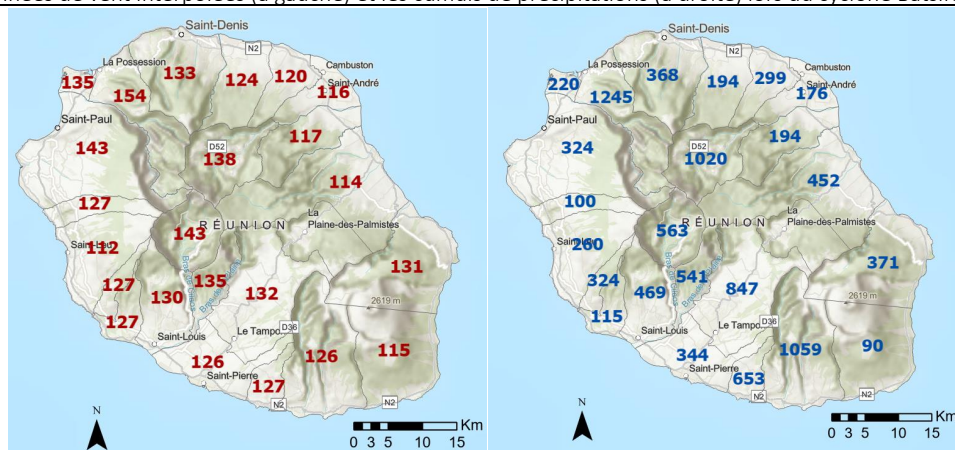


Figure 48 : Vitesse des rafales de vents à gauche et cumuls des précipitations à droite lors du cyclone Batsirai en 2022. Réalisé avec ArcGIS.

Pour les événements majeurs tels que Dina ou Gamède, les résultats des données interpolées ont été comparés aux publications officielles de Météo France et de la CCR afin de vérifier la cohérence des chiffres obtenus.

Présentation du modèle CAT simplifié

L'utilisation conjointe des données d'exposition, de sinistralité et météorologiques rendait particulièrement adaptée la conception d'un modèle CAT. Ce type de modèle permet de relier directement les caractéristiques des biens assurés et des aléas cycloniques à leurs impacts financiers, offrant ainsi une vision holistique des risques. Cependant, leur complexité et le temps nécessaire à leur développement sont incompatibles avec les contraintes de ce mémoire.

Face à ces limitations, une approche pragmatique a été adoptée, privilégiant la fiabilisation des données pour construire un modèle CAT simplifié. Adapté aux besoins opérationnels et conçu dans un contexte de faible granularité des données, ce modèle répond aux exigences de clarté et de vulgarisation imposées par les réassureurs et les décideurs. Il offre des premières estimations rapides à J+3, complétant la méthode des triangles utilisée à J+10 par Prudence Créole, apportant un cadre sécurisant pour les premières analyses. Le graphique ci-dessous offre une vue d'ensemble du processus de modélisation qui est détaillé dans le corps du mémoire.

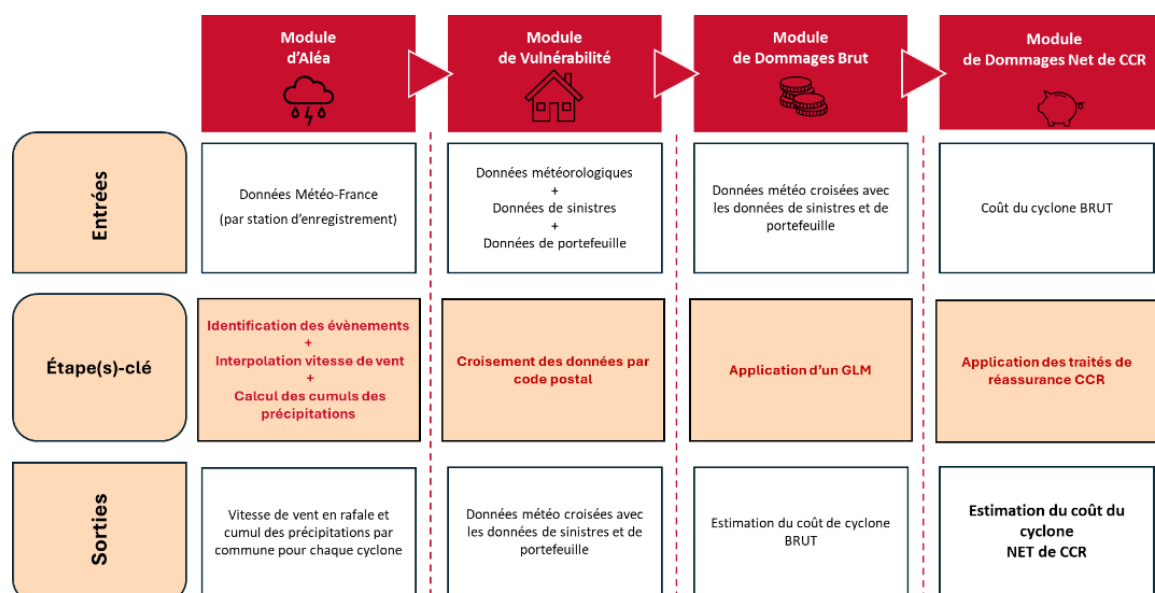


Figure 49 : Schéma du modèle CAT simplifié pour l'estimation du coût d'un cyclone. Réalisé avec PowerPoint.

Dans ce contexte de besoin de vulgarisation, en amont du modèle CAT simplifié, des indicateurs clairs et simples, tels que le pourcentage d'assurés potentiellement impactés ou les seuils météorologiques seront également élaborés. Cela permettra d'identifier rapidement si un événement sera couvert par la réassurance. Ces informations compléteront l'expérience terrain de Prudence Créole et serviront de base à une communication compréhensible avant de présenter le modèle CAT simplifié.

Identification des cyclones nécessitant un suivi de la réassurance

L'objectif de cette première étape de la modélisation est de déterminer si le coût d'un cyclone dépasse le seuil de déclaration auprès des réassureurs, en s'appuyant sur une démarche en deux étapes. La première consiste à évaluer si les données météorologiques indiquent une intensité élevée à travers deux seuils distincts : un seuil pour les précipitations intenses et un autre pour les rafales de vent violentes. La seconde étape analyse le pourcentage des sommes assurées impactées par des conditions dépassant ces seuils. Si ce pourcentage atteint ou dépasse un certain seuil critique, l'événement est considéré comme nécessitant un suivi de la réassurance.

L'analyse doit être rapide et précise, s'appuyant sur les données Météo-France disponibles dans les jours suivants le passage d'un cyclone. Cette exigence opérationnelle justifie l'adoption de seuils distincts, qui traduisent la complexité du risque cyclonique en indicateurs simples, pratiques et directement exploitables.

Pour identifier des seuils représentatifs de phénomènes cycloniques intenses, la méthode *Peaks-Over-Threshold* de la Théorie des Valeurs Extrêmes a été utilisée. Elle modélise les excès au-delà d'un certain seuil, capturant ainsi les caractéristiques des événements les plus marquants. Cette méthode, combinée à des analyses graphiques comme le *Mean Excess Plot* et le graphe de Hill, a permis d'identifier deux seuils pertinents : 810 mm pour les précipitations cumulées et 154 km/h pour les rafales de vent. Ces valeurs traduisent les niveaux extrêmes des aléas susceptibles d'impacter significativement le portefeuille assuré de Prudence Créole.

Enfin, l'approche a été complétée par l'évaluation du pourcentage des sommes assurées affectées par des rafales ou des précipitations intenses. Sur les données empiriques ci-dessous, un premier constat suggère que ce pourcentage se situe entre 11% et 18% de sommes assurées impactées. Puis, grâce à un arbre de décision, il a été déterminé qu'un seuil de 12% de sommes assurées impactées par des conditions intenses correspond à un événement nécessitant un suivi de réassurance.

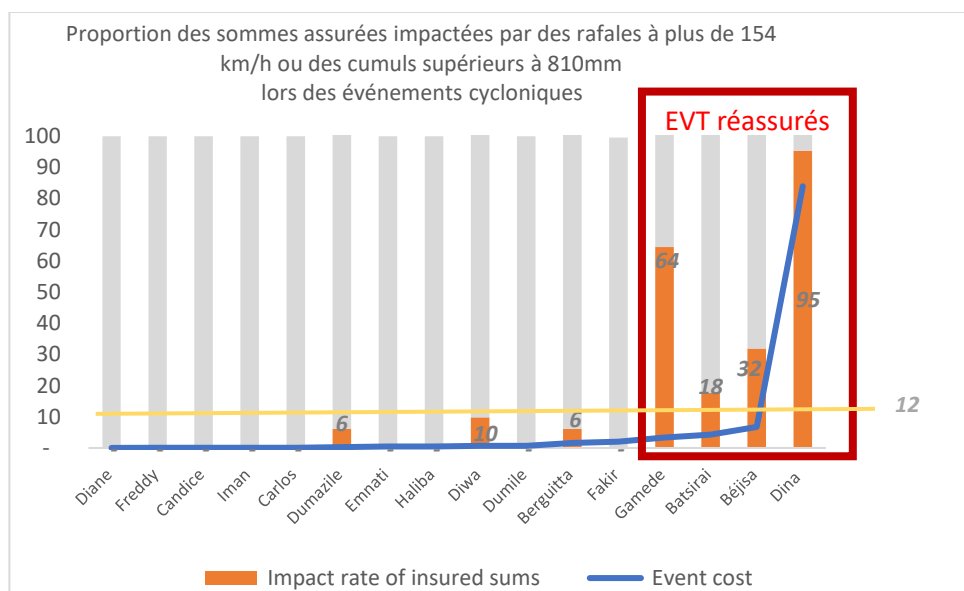


Figure 50 : Proportion des sommes assurées impactées par des rafales à plus de 154km/h ou des cumuls supérieurs à 810mm lors des événements cycloniques. Réalisé avec Excel.

Cette méthodologie a été validée par un back-test concluant réalisé sur le cyclone Belal, mais il est important de noter que les seuils déterminés dans cette partie, s'appuient sur un volume de données limité à 16 cyclones historiques. Si cet échantillon constitue une base précieuse pour cette étude, un enrichissement progressif des données permettrait de renforcer la robustesse des seuils et d'améliorer leur précision.

Estimation de la charge d'un cyclone réassuré

L'estimation de la charge associée aux cyclones réassurés a nécessité une préparation rigoureuse des données et la construction d'un modèle linéaire généralisé (GLM), structuré en deux étapes principales.

La première étape a consisté à traiter les données pour garantir leur pertinence dans la modélisation. Tout d'abord, un plafonnement a été utilisé pour limiter l'impact des sinistres exceptionnels dans certaines communes, tels que ceux liés au cyclone Dina, connu pour sa charge unique dans l'histoire récente de La Réunion. En effet, depuis 2000, sur 17 cyclones ayant impactés Prudence Créole, Dina est le seul à avoir été reconnu CatNat pour le vent, ce qui souligne son caractère

exceptionnel. Toutefois, Dina n'a pas été complètement exclu du modèle, car il constitue une référence incontournable pour représenter les cyclones majeurs ayant impacté l'île. Cette approche permet de conserver une certaine représentativité des événements rares dans la modélisation. Par ailleurs, pour éviter de sous-estimer les résultats, un ajustement des paramètres de la régression est effectué ultérieurement dans le mémoire, garantissant ainsi la cohérence et la robustesse du modèle.

Ensuite, les variables explicatives ont été regroupées en classes grâce à l'algorithme de *K-means*, garantissant une homogénéité des regroupements et facilitant l'établissement de relations linéaires nécessaires au GLM. Cette classification réduit également le risque de corrélations excessives, améliorant ainsi la stabilité et l'interprétabilité des coefficients estimés.

Enfin, un processus de corrélation a permis d'éliminer les variables trop fortement liées, ne conservant que les rafales de vent, les cumuls de précipitations, les sommes assurées totales, les années et les communes.

Pour identifier la loi de probabilité offrant la meilleure adéquation avec les données, plusieurs distributions reconnues pour leur pertinence dans la modélisation des sinistres élevés ont été testées : Gamma, Weibull, Pareto et Cauchy. Ces lois ont été sélectionnées pour leur capacité à capturer des caractéristiques clés des données cycloniques, comme les queues épaisses et l'asymétrie, fréquemment observées dans les charges de sinistres extrêmes. Par un score BIC, un QQ-plot et un test de Kolmogorov-Smirnov, la loi Gamma a été retenue, validant ainsi l'utilisation d'un GLM, car cette loi appartient à la famille exponentielle.

Un *back-test* a ensuite été réalisé sur le cyclone Belal de 2024, confirmant la robustesse du modèle : l'écart entre la charge observée et celle prédite, après calibration, n'est que de 0,7 %. Cela représente une amélioration significative par rapport à la méthode initiale utilisée par Prudence Créole, où les estimations présentaient un delta de l'ordre de 15 % à 20 %. Ce *back-test*, bien qu'il repose sur un unique cyclone, offre des résultats encourageants sur la capacité du modèle à estimer le coût d'un cyclone. Il sera également un bon complément au retour terrain de Prudence Créole.

Même si le modèle présente des performances suffisantes, certaines pistes d'amélioration peuvent être envisagées. Conçu pour capturer les sinistres significatifs, il montre toutefois des limites face à des événements rares et extrêmes, tels que le cyclone Dina, en raison du faible nombre d'observations disponibles. Pour ces types d'événements, une méthode spécifique est recommandée : appliquer le taux de destruction observé lors de Dina, avec une charge augmentée de 6,5 % pour intégrer les dispositions de la réforme CatNat, notamment les frais de relogement d'urgence. Ce mémoire a facilité cette démarche en récupérant et centralisant les données historiques détaillées liées à cet événement exceptionnel, offrant ainsi une base solide pour mieux traiter ce type de sinistre.

Le modèle suppose des relations stables entre les variables explicatives et les charges de sinistres dans le temps. Si cette hypothèse est adaptée au contexte actuel, des ajustements réguliers permettraient de tenir compte des évolutions futures, qu'elles soient liées au changement climatique ou à des modifications des normes de construction.

Les données d'exposition, en raison de leur qualité et de leur disponibilité à une maille très fine, offrent un fort potentiel d'amélioration pour le modèle. Il en va de même pour les données météorologiques qui peuvent être disponibles à l'adresse grâce au travail d'interpolation effectué. Lorsque les sinistres seront suffisamment nombreux à une échelle aussi précise que l'adresse, il sera possible de mieux capturer les impacts locaux et d'améliorer considérablement la granularité et la précision des analyses.

Ces perspectives soulignent le potentiel d'amélioration continue du modèle. Bien qu'il repose sur les données historiques disponibles, chaque nouvel événement cyclonique documenté offre une opportunité d'enrichir la base existante, permettant ainsi d'affiner les estimations et de mieux anticiper les impacts futurs, tout en espérant que ces événements restent rares pour le bien des habitants de La Réunion.

Executive summary

Context

Located in a zone prone to the formation of tropical cyclones, Réunion regularly experiences climatic events of varying intensity, leaving lasting impacts on the island. Emblematic cyclones such as Jenny (1962), Hyacinthe (1980), Firinga (1989), Dina (2002) and Gamède (2007) have caused significant material and human damage. Cyclone Dina, often considered the most devastating of the last three decades, was characterized by wind gusts of up to 300 km/h, accompanied by more than 28 consecutive hours of violent winds. This exceptional duration caused massive destruction, notably to power and communications infrastructures, and extensive flooding. Cyclone Gamède in 2007, for its part, was characterized by unprecedented rainfall, reaching almost 4,000 mm in just three days, leading to devastating floods and landslides. These events underline the fact that the danger posed by cyclones lies as much in the force of the winds as in the intense rainfall, reinforcing the need to take these two hazards into account together.

Although not all cyclones have the same intensity, Prudence Créole, Generali France's subsidiary in La Réunion, sees its business directly impacted by these events. In cyclone years, its combined ratio rises significantly, sometimes necessitating recourse to reinsurance. Prudence Créole benefits from two levels of climate cover through reinsurers. The first, underwritten by Caisse Centrale de Réassurance (CCR), includes two successive treaties to cover natural catastrophes (CatNat): a 50% quota share, followed by stop-loss coverage consolidated by Generali France. The latter is activated when the claims/premiums (S/P) ratio of the property damage branch exceeds 250%, after application of the consolidated Generali France quota shares. In addition, Generali France assumes responsibility for weather-related claims subsequent to cessions to CCR. This includes claims recognized as CatNat net of CCR reinsurance, as well as weather-related claims excluded from this coverage, thanks to a 120 x 5 treaty, part of which is retroceded to Generali France's reinsurers. This structure enables Prudence Créole to limit its losses to a maximum of 5 million euros per cyclone (modulo REC premiums). On the other hand, it creates additional climatic claims for Generali France. For example, in 2024, the acceptance of cessions of cyclone Belal from Prudence Créole to Generali France is the second most costly climatic event for Generali France (at the October 2024 vision).

Relations with Generali France's reinsurers have also become complex, due to the perception of overseas territories as particularly risky zones. All overseas departments and territories are often perceived as homogeneous, even though they present very different climatic and economic risks. This widespread perception, amplified since Cyclone Irma in 2017, has also been exacerbated by the recent riots in New Caledonia and the West Indies.

While Generali France has high-performance models for the main perils in mainland France, such as flooding or drought, no specific modeling for cyclones on La Réunion has yet been developed.

Moreover, market models adapted to the island are virtually non-existent, due to its geographical isolation, the scarcity of historical data, and the topographical complexity that amplifies hazards. Faced with these challenges and the intensification of cyclones linked to climate change, the design of a dedicated cyclone model now seems unavoidable to better anticipate and manage these specific risks. This thesis focuses on two central questions: How can we quickly identify whether a cyclone requires reinsurance for Prudence Créole? And how can we estimate the costs associated with these events for Prudence Créole and Generali France?

These two objectives aim to deepen our understanding of cyclonic risks, while at the same time providing understandable indicators to facilitate their use in an operational context. By meeting these challenges, this work also contributes to strengthening reinsurers' confidence, an essential lever for negotiating optimal coverage conditions.

To build such a model, it is essential to carry out in-depth work on the available data, which must be carefully consolidated and prepared to guarantee the robustness and relevance of the analyses.

Claims data

The historical claims data used in this study were supplied by Prudence Créole and cover the period from 2002 to 2024. This database includes claims related to cyclones that have impacted the island and provides a detailed overview of the associated financial expenses, the characteristics of the claims and the contracts concerned.

For the purposes of modeling, only essential variables with sufficient information were retained. This includes contractual information (identification numbers, zip code, insurance product) and claims data (date of occurrence, amount of expenses, identification of CatNat claims). On the other hand, certain variables, such as the cover code or principal peril, were excluded due to their poor quality or inconsistency.

To guarantee their relevance and consistency over time, these data have been the subject of in-depth “AS-IF” updating. This involves adjusting claims expenses to a current view, taking into account price trends (via inflation coefficients based on FFB and RI indices) as well as changes in the insured portfolio. This step makes it possible to revalue cyclones to the present day, neutralizing the effects of temporal variations and ensuring the comparability of current exposure data. Finally, claims eligible for the CatNat system have been increased by 6.5% to take better account of the CatNat reform.

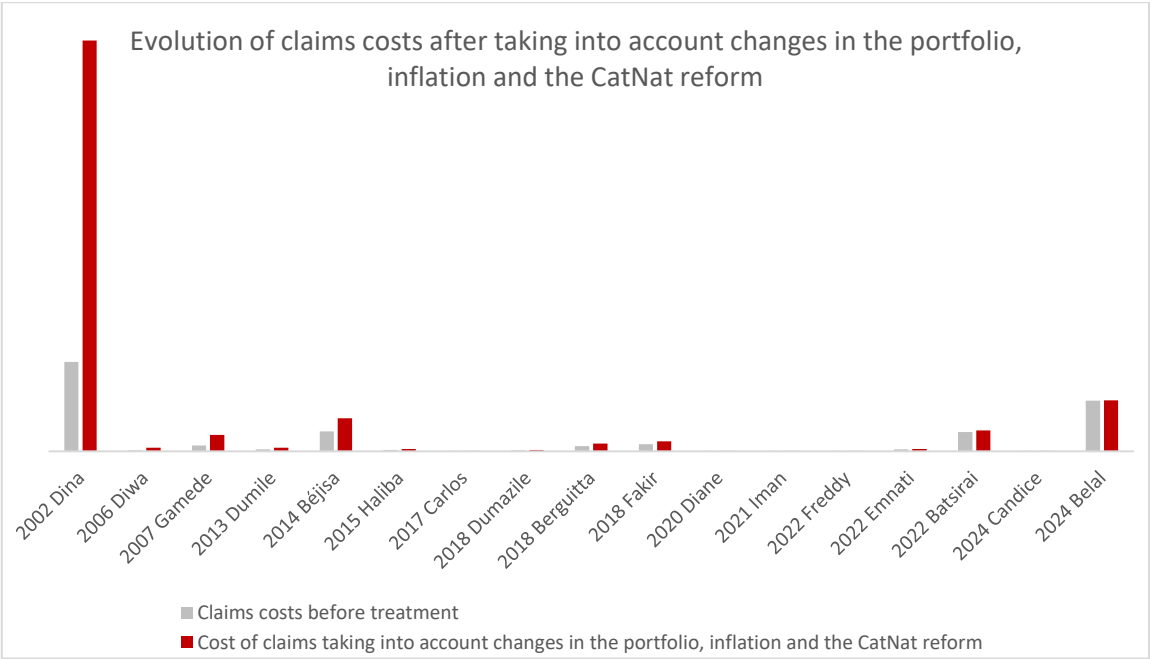


Figure 1: Claims cost evolution after taking into account portfolio evolution, inflation and the CatNat reform.
Created using Excel.

At the same time, claims were attached to the communal grid due to the limitations of older databases, such as those for cyclones Dina and Gamède, where precise addresses were not available or usable. Although the data apart from 2018 enable a more precise attachment of losses to exposure data, using these data alone would have led to the exclusion of major cyclones of the past, which was not feasible. The choice of the communal grid thus enabled us to maintain overall consistency in the analysis, while circumventing these limitations.

Although AS-IF allows data to be upgraded to reflect current conditions, dependence on historical data may restrict the ability to fully integrate future developments, such as the intensification of cyclones linked to climate change, or socio-economic transformations, such as urbanization or improved building standards. In addition, the use of the communal grid, although adapted to the

constraints of the available data, does not allow us to capture the topographical specificities of La Réunion, such as variations linked to mountainous areas or valleys. Yet these features influence the intensity of gusts and precipitation. Incorporating a topography index would have required lengthy and complex processing, but it remains an interesting avenue for refining future estimates.

Exposure data

Exposure data includes all information relating to insured property, essential for assessing Prudence Créole's financial exposure in the event of a claim. Taken from the property portfolio, it includes details such as class of insurance, type of building, type of occupancy (owner, tenant, etc.), structural characteristics (concrete construction, etc.), sums insured, as well as certain contractual elements such as deductibles and co-insurance rates.

To prepare this data for modeling, cleansing and consolidation processes were necessary. Particular attention was paid to the “type of construction” variable, which was the only one to present a level of completeness below 90%. This was due to a poor transposition of the initial Prudence Créole data. While the data contained 57 very detailed construction type descriptions, this fine granularity made it difficult to use the data in a harmonized data collection framework with other Generali France entities.

To overcome this complexity, a new classification was created, grouping constructions into four main categories, inspired by the methodologies developed by RMS. This simplification, while reducing the complexity of the data, better reflects the vulnerability characteristics of buildings, particularly their resistance to high winds. The reclassified building types are presented in the table below, providing a robust basis for integrating these data into cyclonic impact modeling.

Name	Risk Kevel	Commentary
All-concrete structure	Low	Buildings where the main structure and roof are entirely made of concrete offer significant resistance to the strong winds typical of cyclones, as well as to water infiltration and flooding.
Concrete structure with different roofing	Intermediate	These buildings have a concrete structure but a roof made of lighter materials, such as metal sheets or plastic materials. These lighter materials are more vulnerable to infiltration and deterioration from strong gusts during intense cyclones. Metal roofs, in particular, are also more prone to water leaks.
Steel structure and non-concrete roofing	High	Buildings with a steel framework and a roof made of lighter materials (such as metal sheets or plastic) are especially vulnerable to cyclones. These roofs are easily damaged by strong winds, exposing the structure and interior to water infiltration and increasing the risks of flooding inside.
Unknown	High	These buildings may present unknown vulnerabilities to wind, flooding, and infiltration. As a result, they are considered hazardous.

Table 1: Table of construction type categories.

In addition, processing was carried out to make the data usable at the commune level. The sums insured, initially available at very fine levels (such as precise addresses), were aggregated by commune. This step was essential to link exposure data to claims data, which were themselves structured at this geographical level.

Weather data

The climatic data used in this report come from Météo France's open-source databases and cover the period from 2002 to 2024. This information, recorded by meteorological stations on Reunion Island, includes daily measurements of wind speeds and rainfall totals.

To ensure their relevance for modeling purposes, specific processing was carried out. Firstly, certain stations located at altitudes of over 1,000 meters were excluded. Although these stations often record particularly high wind gusts, they do not reflect the weather conditions in the populated areas where

the insured properties are concentrated. Their inclusion would have artificially amplified the intensity of hazards in areas with little economic exposure.

Due to the limited number of wind data recorded during the various events, spatial interpolation proved necessary. This method enables wind speeds to be estimated for points where no stations are present. The weighted inverse distance (WID) method was used for this interpolation, and the average wind speed per commune was preferred. This choice offers a more accurate representation of the actual exposure of policyholders, avoiding overestimation of hazards by basing data on maximum values recorded in sparsely populated areas.

Finally, for precipitation, maximum accumulations over the duration of the event have been aggregated to the municipal level. This methodological choice is based on several considerations. Firstly, high-altitude stations, which are likely to overestimate precipitation for certain zones, were excluded, making the maximum accumulation per commune more representative of the extreme conditions actually experienced by policyholders. Secondly, although it is not always possible to precisely identify the peril causing the loss in the case of historical events, discussions with Lionel LAFFIN, Saretec's construction expert, have shown that rainfall-related seepage accounts for a significant proportion of the damage incurred in Réunion in recent years.

Two key indicators were derived for each cyclone and each commune: average gust speed and maximum cumulative rainfall over the duration of the event. These results are visualized in the maps below, which illustrate respectively gust speeds and rainfall totals for the whole of Reunion during cyclone Batsirai in 2022.

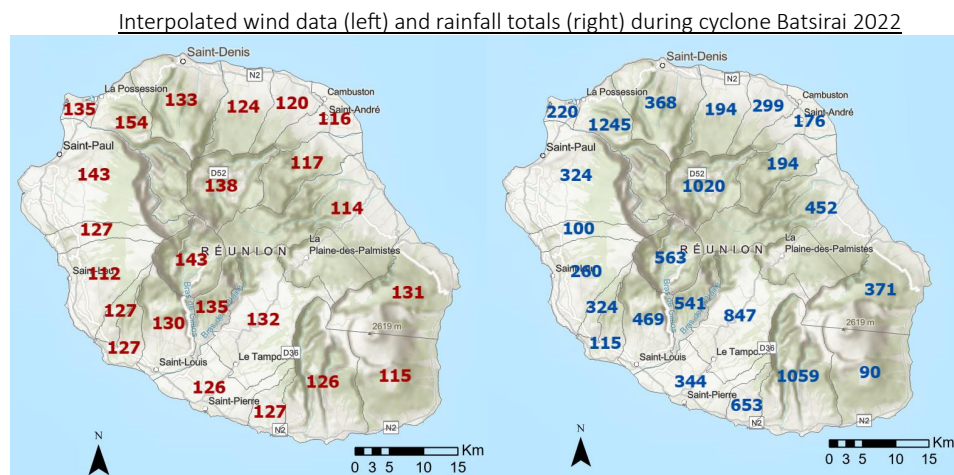


Figure 2: Wind gust speeds on the left and rainfall totals on the right during cyclone Batsirai in 2022. Realized with ArcGIS

For major events such as Dina or Gamède, the results of interpolated data were compared with official publications from Météo France and the CCR to check the consistency of the figures obtained.

Presentation of the simplified CAT model

The combined use of exposure, claims and meteorological data made the design of a CAT model particularly appropriate. This type of model makes it possible to directly link the characteristics of insured assets and cyclonic hazards to their financial impact, thus offering a holistic view of risks. However, their complexity and the time required to develop them are incompatible with the constraints of this thesis.

Faced with these limitations, a pragmatic approach was adopted, focusing on making data more reliable to build a simplified CAT model. Adapted to operational needs and designed in a context of low data granularity, this model meets the requirements of clarity and popularization imposed by reinsurers and decision-makers. It offers rapid initial estimates at D+3, complementing the triangle method used

at D+10 by Prudence Créole, and providing a reassuring framework for initial analyses. The diagram below provides an overview of the modeling process, which is detailed in the body of the report.

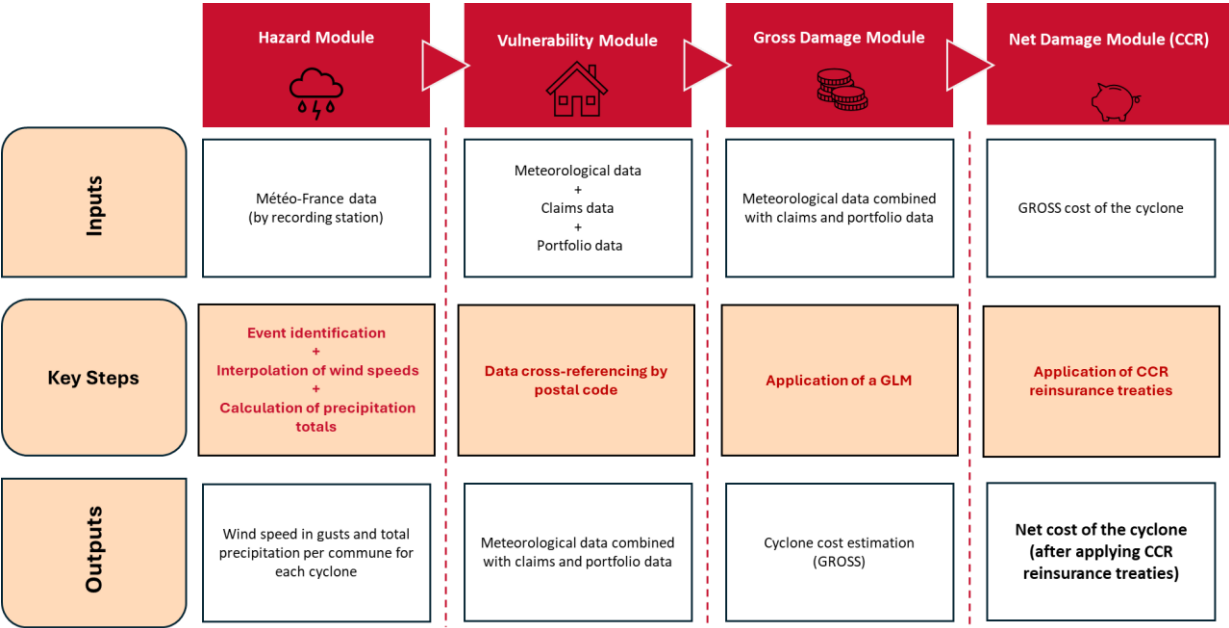


Figure 3: Diagram of the simplified CAT model for estimating the cost of a cyclone. Made with PowerPoint.

In this context of need for popularization, upstream of the simplified CAT model, clear and simple indicators, such as the percentage of policyholders potentially impacted or meteorological thresholds, will also be developed. This will enable rapid identification of whether an event will be covered by reinsurance. This information will complement Prudence Créole's experience in the field and serve as a basis for comprehensible communication before presenting the simplified CAT model.

Identifying cyclones requiring reinsurance monitoring

The aim of this first modelling stage is to determine whether the cost of a cyclone exceeds the threshold for reporting to reinsurers, using a two-stage approach. The first step consists in assessing whether the meteorological data indicate a high intensity through two distinct thresholds: one for intense precipitation and another for violent wind gusts. The second step analyses the percentage of insured sums impacted by conditions exceeding these thresholds. If this percentage reaches or exceeds a certain critical threshold, the event is considered to require reinsurance follow-up.

The analysis must be rapid and precise, based on Météo-France data available in the days following the cyclone's passage. This operational requirement justifies the adoption of distinct thresholds, which translate the complexity of cyclonic risk into simple, practical and directly exploitable indicators.

To identify thresholds representative of intense cyclonic phenomena, the Peaks-Over-Threshold method of Extreme Value Theory was used. It models excesses above a certain threshold, thus capturing the characteristics of the most significant events. This method, combined with graphical analyses such as the Mean Excess Plot and the Hill graph, enabled us to identify two relevant thresholds: 810 mm for cumulative precipitation and 154 km/h for wind gusts. These values reflect the extreme levels of hazards likely to have a significant impact on Prudence Créole's insured portfolio.

Finally, the approach was completed by assessing the percentage of insured sums affected by gusts of wind or intense precipitation. Based on the empirical data below, an initial assessment suggests that this percentage lies between 11% and 18% of insured sums impacted. Then, using a decision tree, it was determined that a threshold of 12% of insured sums impacted by intense conditions corresponds to an event requiring reinsurance follow-up.

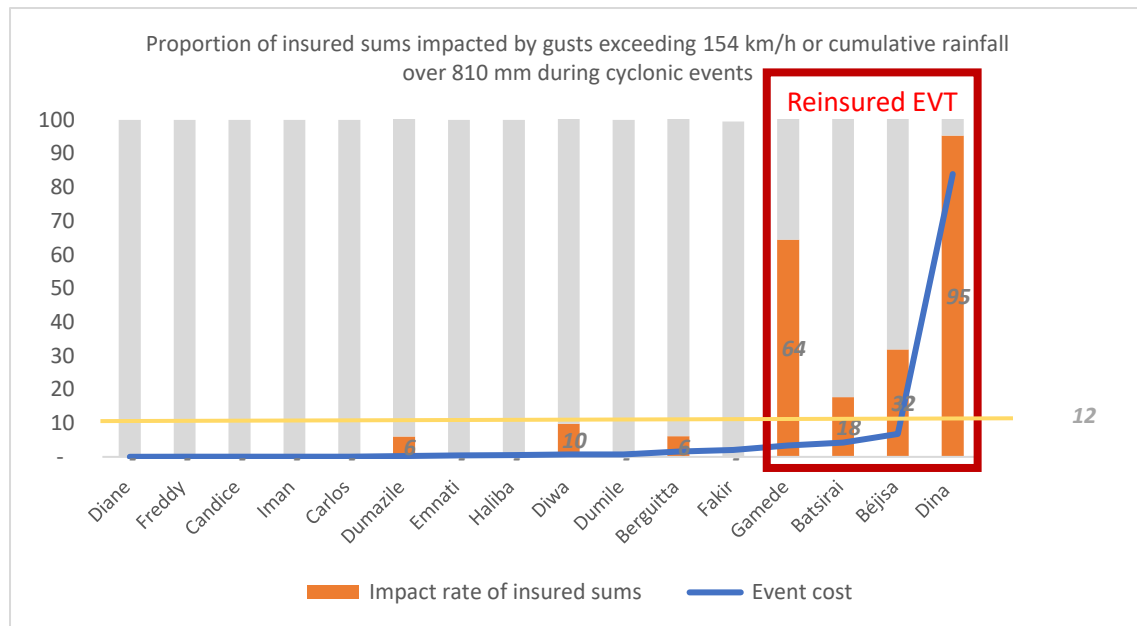


Figure 4 : Proportion of insured sums impacted by gusts in excess of 154km/h or cumulative rainfall in excess of 810mm during cyclonic events. Made with Excel.

This methodology was validated by a conclusive back-test carried out on cyclone Belal, but it is important to note that the thresholds determined in this section are based on a volume of data limited to 16 historical cyclones. While this sample provides a valuable basis for this study, a progressive enrichment of the data would reinforce the robustness of the thresholds and improve their accuracy.

Estimating the load of a reinsured cyclone

Estimating the load associated with reinsured cyclones required rigorous data preparation and the construction of a generalized linear model (GLM), structured in two main stages.

The first stage involved processing the data to ensure its relevance to the model. Firstly, capping was used to limit the impact of exceptional disasters in certain communes, such as those linked to cyclone Dina, known for its unique load in Réunion's recent history. Indeed, since 2000, of the 17 cyclones to have impacted Prudence Créole, Dina is the only one to have been recognized as a CatNat for wind, which underlines its exceptional nature. However, Dina has not been completely excluded from the model, as it is an essential reference for representing the major cyclones that have impacted the island. This approach ensures that the model retains a certain representativeness of rare events. Furthermore, to avoid underestimating the results, an adjustment of the regression parameters is carried out later in the thesis, thus guaranteeing the model's consistency and robustness.

Next, the explanatory variables were grouped into classes using the *K-means* algorithm, guaranteeing homogeneous groupings and facilitating the establishment of the linear relationships required for GLM. This classification also reduces the risk of excessive correlations, improving the stability and interpretability of the estimated coefficients.

Finally, a correlation process was used to eliminate variables that were too strongly related, retaining only wind gusts, rainfall totals, total insured sums, years and communes.

To identify the probability distribution offering the best fit with the data, several distributions recognized for their relevance in modeling high claims were tested: Gamma, Weibull, Pareto and Cauchy. These distributions were selected for their ability to capture key features of cyclonic data, such as thick tails and asymmetry, frequently observed in extreme claims loads. By means of a BIC score, a QQ-plot and a Kolmogorov-Smirnov test, the Gamma distribution was selected, validating the use of a GLM, as this distribution belongs to the exponential family.

A back-test was then carried out on the Belal cyclone of 2024, confirming the robustness of the model: the difference between the observed and predicted load, after calibration, was only 0.7%. This

represents a significant improvement on the initial method used by Prudence Créole, where deviations typically reached between 15% and 20%. This back-test, although based on a single cyclone, offers encouraging results on the model's ability to estimate the cost of a cyclone. It will also be a good complement to the field feedback from Prudence Créole.

Although the model's performance is adequate, there are several areas for improvement. While designed to capture significant losses, the model shows certain limitations when faced with rare and extreme events, such as cyclone Dina, due to limited available observations. For these types of events, a specific method is recommended: applying the destruction rate observed during Dina, with a charge increased by 6.5% to incorporate the provisions of the CatNat reform, notably emergency rehousing costs. This work has facilitated this process by recovering and centralizing detailed historical data relating to this exceptional event, thus providing a solid basis for better handling this type of loss.

The model assumes stable relationships between explanatory variables and claims costs over time. While this assumption is appropriate for the current context, regular adjustments could be made to take account of future developments, whether linked to climate change or changes in building standards.

Because of their quality and availability at a very fine scale, exposure data offer great potential for improving the model. The same applies to meteorological data, which may be available at the address thanks to the interpolation work carried out. When there are enough claims on a scale as precise as the address, it will be possible to better capture local impacts and considerably improve the granularity and accuracy of analyses.

These prospects underline the model's potential for continuous improvement. Although it is based on available historical data, each new documented cyclonic event offers an opportunity to enrich the existing database, enabling us to refine estimates and better anticipate future impacts, while hoping that such events remain rare for the sake of Réunion's inhabitants.

Annexes

<i>Annexe 1 : Réforme CatNat - Prise en charge des frais de relogement d'urgence</i>	125
<i>Annexe 2 : Liste des variables dans la base de données d'exposition reçue par Prudence Créole.</i>	126
<i>Annexe 3 : Base de données à l'issue du module vulnérabilité du modèle CAT simplifié</i>	130
<i>Annexe 4 : Tableaux de regroupement en classes (variable rafales de vents et variables non utilisées)</i>	131
<i>Annexe 5 : Définitions d'appartenir aux domaines d'attractions de Fréchet, Gumbel et Weibull (d'après le cours de Christian Y. Robert de février 2024 pour l'ENSAE)</i>	132
<i>Annexe 6 : Graphique des périodes de retour des cyclones par pays selon les vitesses de vent</i>	133
<i>Annexe 7 : Modélisation de la charge cyclonique sans plafonnement</i>	134
<i>Annexe 8 : Compléments sur les modèles linéaires généralisés</i>	137
<i>Annexe 9 : Fonctionnement de Ensemble Face Aux Risques Prudence Créole. Source : document interne.</i>	141

Annexe 1 : Réforme CatNat - Prise en charge des frais de relogement d'urgence

La loi du 28 décembre 2021 marque une avancée majeure en incluant les frais de relogement d'urgence dans les indemnisations CatNat. Désormais, les personnes sinistrées dont la résidence principale est devenue inhabitable pour des raisons de sécurité, salubrité ou hygiène peuvent bénéficier d'une prise en charge spécifique.

Initialement prévue pour le 1er janvier 2024, cette mesure a été anticipée au 1er novembre 2023 par décret, à la suite des graves inondations dans le Nord et le Pas-de-Calais. Les modalités, définies par le décret n° 2022-1737 et l'arrêté du 30 décembre 2022, précisent que cette indemnisation s'applique uniquement aux frais d'hébergement des occupants assurés, pour une durée maximale de six mois à compter du début du relogement.

L'indemnisation est calculée en fonction de la situation de l'assuré :

1. Propriétaires occupants : prise en charge à hauteur de la valeur locative du logement sinistré.
2. Locataires ou occupants à titre gratuit : remboursement des loyers payés, charges incluses, ou de la valeur locative du logement.
3. Locataires dont le bail a pris fin : indemnisation du surcoût de relogement par rapport aux loyers précédents, dans la limite de trois mois.

Enfin, cette indemnisation est exclusive : si l'assuré bénéficie de cette prise en charge par son assureur, il ne peut cumuler une aide financière de l'État pour les mêmes frais.

Ce dispositif renforce la protection des sinistrés tout en ajoutant une nouvelle dimension financière à la gestion des risques pour les assureurs.

Annexe 2 : Liste des variables dans la base de données d'exposition reçue par Prudence Créole.

Nom du champs	Libellé du champs	Descriptif
cie	CODE CIE	P = Prudence Créole
police	NUM POLICE	Numéro du contrat
numero_de_client	Numéro de client	
nom_de_l_assure	Nom de l'assuré	
id_site	ID SITE	Numéro du site
nb_sites	NOMBRE DE SITES	Nombre total de sites du contrat
typ_risque	TYPE DE RISQUE	C : Multirisque Commerciale R : Multirisque habitation L : Multirisque immeuble/ PNO G : Multirisque Garage H : Multirisque Fabricant/Artisan A : Multirisque Agricole M : Multirisque collectivités I : Multirisque Industrielle
csd	CSD	Catégorie statistique et détaillé, Code Produit si non dispo
typ_odr	Type d'Objet De Risque	Permet de qualifier le type d'ODR B = Bâtiment C = Contenu M = Machine E = Activité
date_extraction	Date d'extraction des données contrat	Date à laquelle l'extraction du système est effectuée au format JJ/MM/AAAA
annee_sousc	ANNEE SOUSCRIPTION	Année de souscription : Renseigner l'année N ou N-1 suivant la date d'extraction au format AAAA
date_effet	DATE EFFET	Date de première souscription du contrat au format JJ/MM/AAAA
date_expiration	DATE EXPIRATION	Date de fin d'effet du contrat (si différente de la date anniversaire) au format JJ/MM/AAAA
date_anniversaire	Date d'anniversaire du contrat	Date d'échéance principale du contrat au format JJ/MM
etat_police	ETAT_POLICE	Etat/Code état de la police (joindre une table de correspondance)

typ_adress	TYPE D'ADRESSE	R = Adresse du risque; F = Adresse de facturation.
adresse	ADRESSE DU RISQUE	
cpostal	CODE POSTAL	
cinsee	CODE INSEE	Code Insee de la commune du risque
ville	VILLE	Ville du risque
pays	PAYS	Pays du risque
typ_bat	TYPE DE BATIMENT	A : Appartement I : Immeuble M : Maison T : Atelier B: Bureau C : Commercial S : Usage Industriel Y : Usage Agricole E : Entrepot
typ_constr	TYPE CONSTRUCTION	Type de construction des batiments - joindre la table de correspondance
annee_constr	ANNEE CONSTRUCTION	Année de construction du bâtiment assuré
typ_occ	TYPE OCCUPATION	C : Co-proprétaire (occupant) CPN : Co-proprétaire non occupant L : Locataire LP : Locataire agissant pour le compte du propriétaire P : Propriétaire (occupant) PNO : Propriétaire Non Occupant SY : Syndic
mod_occ	MODE D'OCCUPATION	P= Résidence principale S= Résidence secondaire
nb_etages	NBRE ETAGES	Nombre d'étages du bâtiment où se situe le risque
num_etage	NUMERO ETAGE	Etage du bien assuré
pieces	NBRE PIECES	Nombre de pièces principales du bâtiment assuré . Concerne essentiellement le portefeuille Habitation.
surface	SURFACE	Surface du bâtiment assuré.
valeur_m2	VALEUR PAR M²	y compris frais annexes

Sommes assurées	SOMMES ASSUREES BATIMENT / INCENDIE	Sommes assurées pour les bâtiments au titre de la garantie Incendie au 100% du site
	SOMMES ASSUREES BATIMENT / TOC	Sommes assurées pour les bâtiments au titre de la garantie Tempêtes au 100% du site
	SOMMES ASSUREES CONTENU	Sommes assurées pour le contenu au 100% du site
	SOMMES ASSUREES FRAIS ET PERTES	Sommes assurées pour les Frais et Pertes Indirectes au 100% du site
	SOMMES ASSUREES BRIS DE MACHINE	Sommes assurées pour les bris de machine au 100% du site, au titre des contrats spécifiques.
	SOMMES ASSUREES DOMMAGES DIRECTS	Sommes assurées pour les dommages directs au 100% du site
	SOMMES ASSUREES PE	Sommes assurées pour les pertes d'exploitation au 100% du site
	SOMMES ASSUREES PE CONTRAT	Sommes assurées pour les pertes d'exploitation au 100% de la police
val_venale	VALEUR VENALE	Valeur vénale du fonds de commerce

Primes	PRIME INCENDIE	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME PE	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME RC	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME BDM	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME CAT NAT	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME TOC TGN	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME ATTENTAT TERRORISME	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME AUTRES DOMMAGES	Prime émise annuelle hors taxe brute de commission au 100% du site
	PRIME TOTAL HT	Prime émise annuelle hors taxe brute de commission au 100% du site
coass	COASSURANCE	Engagement de la compagnie sur la police . il s'agit de l'engagement à la part compagnie exprimée en %, ex. 100, 75, 50....
typ_coass	TYPE DE CO-ASS	A = Apériteur; C=co-assureur, Non renseigné si sans objet

Tableau 23 : Tableau de toutes les variables initialement fournies par Prudence Créole dans la base de données d'exposition. Source : données internes.

Annexe 3 : Base de données à l'issue du module vulnérabilité du modèle CAT simplifié

Le tableau, y compris les deux premières lignes pour donner des exemples, a été transposé pour des questions d'affichage. Les éléments grisés sont là pour respecter la confidentialité des données.

Commune	Saint Paul	Sainte Suzanne
Année	2022	2022
Nom événement	Batsirai	Batsirai
CHARGE inflatée AS IF (y compris réforme CatNat)		
Vent moyen	71	76
Vent en rafale	133	126
Précipitations	368	344
Engagement total de la commune		
Prime totale de la commune		
SA béton		
SA structure béton avec toiture diff		
SA acier		
SA Propriétaire		
SA Locataire		
SA Appartement		
SA usage Commercial		
SA Immeuble		
SA Maison		
SA usage Industriel		

Tableau 24 : Base de données à l'issue du module Vulnérabilité du modèle CAT simplifié.

Annexe 4 : Tableaux de regroupement en classes (variable rafales de vents et variables non utilisées)

Catégorie Saffir-Simpson des vitesses de vent en km/h	Effectif
Groupe 1 : < 29	1
Groupe 2 : 29-63 (dépression tropicale)	33
Groupe 3 : 64-118 (tempête tropicale)	158
Groupe 4 : 119-153 (Ouragan Cat 1)	85
Groupe 5 : 154-177 (Ouragan Cat 2)	27
Groupe 6 : 178-209 (Ouragan Cat 3)	4

Catégorie cyclone océan Indien des vitesses de vent en km/h	Effectif
Groupe 1 : < 29	1
Groupe 2 : 29-51 (perturbation tropicale)	14
Groupe 3 : 52 - 63 (dépression tropicale)	19
Groupe 4 : 64-88 (tempête tropicale modérée)	62
Groupe 5 : 89 - 118 (forte tempête tropicale)	96
Groupe 6 : 119 - 165 (Cyclone tropical)	104
Groupe 7 : 167 - 208 (cyclone tropicale intense)	12

Tableau 25 : Tableaux des tentatives de classification de la variable rafale de vent.

Variable	Nombre de classe optimale
SA en tout béton	3
SA en béton avec une toiture dans un autre matériel	5
SA en acier	4
SA appartement	4
SA commercial	4
SA immeuble	3
SA maison	4
SA indus	4

Figure 51 : Tableau du nombre de classe optimale des variables non gardées dans le mémoire.

Domaine d'attraction de la loi de Fréchet

Une fonction de répartition F appartient au domaine d'attraction de la loi de Fréchet Φ_α , avec $\alpha > 0$, si et seulement si il existe une fonction lentement variable L telle que :

$$\bar{F}(x) = x^{-\alpha}L(x)$$

Où

- $\bar{F}(x) = 1 - F(x)$ est la fonction de survie de F
- L est mesurable et positive sur $]0, \infty[$ et satisfait la condition suivante :

$$\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1, \quad \text{pour tout } x > 0.$$

Domaine d'attraction de la loi de Gumbel

Une fonction de répartition F appartient au domaine d'attraction de la loi de Gumbel Λ si et seulement s'il existe une fonction positive g telle que :

$$\lim_{x \rightarrow x_F} \frac{\bar{F}(x + tg(x))}{\bar{F}(x)} = \exp(-t), \quad \text{pour tout } t \in \mathbb{R}.$$

Une forme possible pour la fonction $g(x)$ est donnée par :

$$g(x) = \frac{\int_x^{x_F} \bar{F}(t) dt}{\bar{F}(x)} = E(X - x \mid X > x),$$

Où x_F correspond au bord supérieur du support de F , avec $x_F = \infty$ si le support est non borné.

Domaine d'attraction de la loi de Weibull

Une fonction de répartition F appartient au domaine d'attraction de la loi de Weibull Ψ_α , avec $\alpha > 0$, si et seulement si $x_F < \infty$ et $\bar{F}(x_F - x^{-1}) = x^{-\alpha}L(x)$.

Où

- x_F est le point terminal du support de F ,
- $\bar{F}(x) = 1 - F(x)$.
- L est mesurable et positive sur $]0, \infty[$ et satisfait la condition suivante :

$$\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1, \quad \text{pour tout } x > 0.$$

Annexe 6 : Graphique des périodes de retour des cyclones par pays selon les vitesses de vent

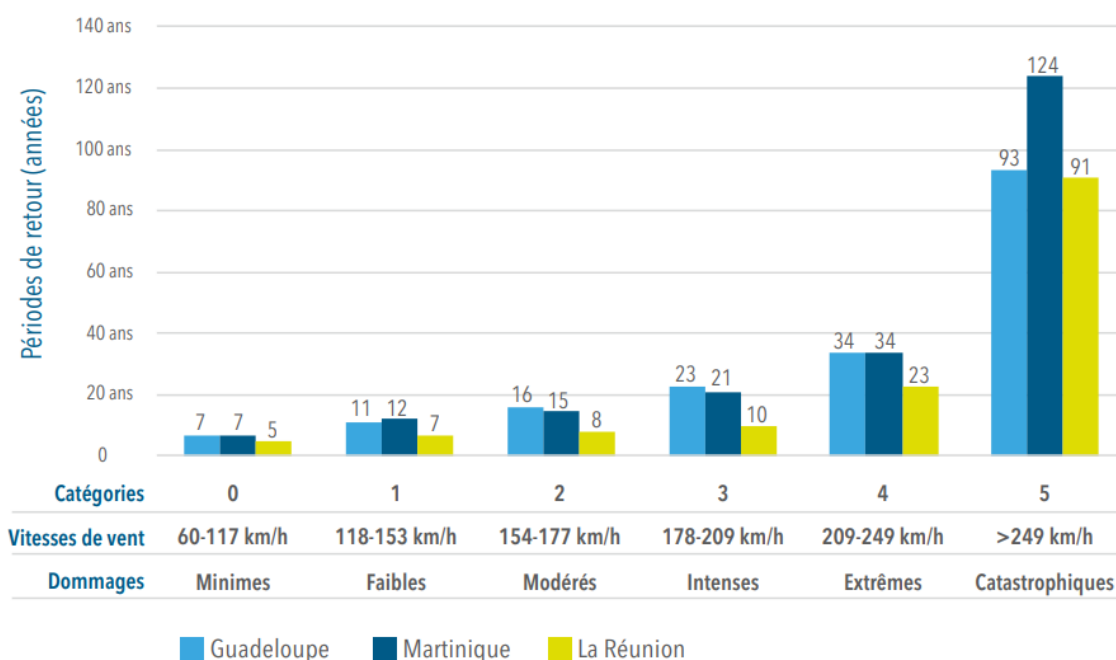


Figure 52 : période de retour des cyclones selon l'approche historique avant la survenance d'Irma et Maria.
Source : CCR.

La figure met en évidence des différences marquées entre La Réunion et les Antilles en termes de périodes de retour des rafales de vent. À La Réunion, les vents extrêmes (209-249 km/h) et catastrophiques (>249 km/h) sont moins fréquents, avec des périodes de retour de 23 ans et 91 ans respectivement, comparées aux Antilles. Ces observations confirment que les rafales de vent à La Réunion, bien qu'intenses, sont moins violentes et fréquentes, en accord avec l'analyse basée sur la théorie des valeurs extrêmes qui indique des queues bornées pour les vitesses de vent dans cette région.

Annexe 7 : Modélisation de la charge cyclonique sans plafonnement

Cette annexe a pour objectif de démontrer l'importance de plafonner la sinistralité de certaines communes dans la modélisation et de mettre en évidence les performances inférieures d'un modèle construit sans plafonnement.

Pour rappel, le plafonnement vise à limiter l'influence des valeurs extrêmes lors de la phase de modélisation, afin de garantir une meilleure stabilité statistique et d'éviter une surestimation artificielle des charges. Cette méthode est particulièrement adaptée au contexte des cyclones à La Réunion, où des événements exceptionnellement intenses, comme le cyclone Dina en 2002, ont été observés. Étant donné que le modèle n'est pas conçu pour représenter un cyclone d'une telle envergure, les taux de destruction spécifiques à Dina sont directement appliqués dans ce cas précis.

Comparaison des modèles

Un modèle GLM log-gamma a été construit sans appliquer de plafonnement préalable aux données mais en appliquant tous les autres traitements détaillés dans le corps du mémoire. Le tableau ci-dessous compare les indicateurs BIC d'un modèle avec et sans plafonnement. A ce stade, bien que le critère BIC du modèle avec plafonnement soit légèrement meilleur, il n'est pas encore possible de conclure que ce modèle est plus adéquat.

	Sans plafonnement	Avec plafonnement
BIC	7 180	7 115

Tableau 26 : Score BIC du modèle avec et sans plafonnement. Réalisé avec Excel.

Coefficients du GLM

Les coefficients estimés par le GLM avec et sans plafonnement est présenté ci-dessous :

	Classe	Coefficients du GLM sans plafonnement	Coefficients du GLM sans plafonnement
SA	Classe SA 1	-1,31	-1,24
	Classe SA 2	-0,5	-0,45
	Classe SA 4	0,89	0,83
Année	2013	-1	-0,95
	2021	-1,35	-1,41
	2006 - 2015 - 2024 - 2007	-0,54	-0,51
	2023 - 2020 - 2017	-1,58	-1,61
	2014 - 2002 - 2018	0,52	0,42
Commune	Saint Leu, Saint André, Saint Pierre, Sainte Marie, Le Port, La Possession, Saint Benoît	-0,80	-0,62
	Salazie, La Plaine des Palmistes, L'Étang Salé, Saint Philippe, Les Trois Bassins, Saint Joseph, Les Avirons, Entre Deux, Petite Île, Sainte Rose, Le Tampon, Sainte Suzanne, Bras Panon, Saint Louis, Cilaos, Saint Paul	-0,92	-0,66
Vent	[26 - 93,6]	-0,5	-0,54
	[110,98 - 128,75]	0,60	0,61
	[128,75 - 148,79]	1,53	1,54

	[149,79 - 193,67]	3,0	2,24
Précipitation	[58,6 - 194,5]	0,4	0,37
	[194,5 - 297,49]	0,5	0,51
	[297,49 - 437,77]	1	0,88
	[437,77 - 634,53]	1,22	1,1
	[634,53 - 3730]	1,31	1,17
Intercept		10,7	10,5

Tableau 27 : Coefficients du GLM avec et sans plafonnement. Réalisé avec Excel.

Le plafonnement a un impact significatif sur les coefficients, en particulier pour les valeurs extrêmes des classes de vent et de précipitation :

- Pour le vent, le coefficient des rafales les plus fortes ([149,79 - 193,67]) passe de 3,0 sans plafonnement à 2,24 avec plafonnement, illustrant une diminution de l'impact exagéré des vents violents sur les charges.
- De manière similaire, pour les précipitations les plus élevées ([634,53 - 3730 mm]), le coefficient diminue de 1,31 à 1,17, reflétant une réduction de l'influence des cumuls excessifs.

En revanche, pour des valeurs modérées de vent et de précipitation, les coefficients restent quasi identiques entre les deux modèles, démontrant que le plafonnement agit principalement sur les conditions météorologiques les plus violentes. Ce mécanisme contribue à limiter l'effet des valeurs extrêmes tout en conservant la stabilité et la cohérence globale du modèle.

Analyse des résidus

En ce qui concerne l'analyse des résidus, la différence entre les deux modèles est particulièrement visible sur l'histogramme des résidus et la courbe de Lorentz.

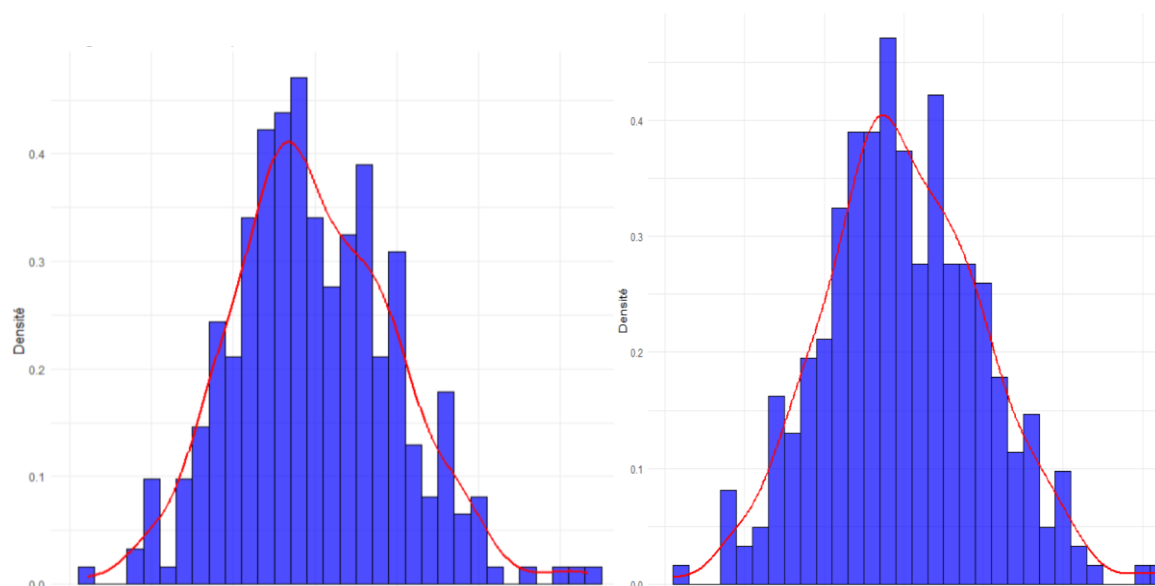


Figure 53: histogramme des résidus sans plafonnement à gauche et avec à droite. Réalisé sous R.

L'histogramme des résidus pondérés illustre l'impact du plafonnement sur la qualité du modèle. À gauche, sans plafonnement, la distribution des résidus est asymétrique, avec des queues plus longues traduisant la présence de valeurs extrêmes mal capturées par le modèle. Cette asymétrie indique un modèle moins stable et des prédictions biaisées, particulièrement pour les observations extrêmes.

En revanche, à droite, avec plafonnement, les résidus sont répartis de manière plus symétrique autour de zéro, et la courbe de densité s’aligne mieux sur les barres de l’histogramme, suggérant une meilleure adéquation à une distribution normale. Ce résultat reflète un modèle plus robuste, où l’influence des valeurs atypiques a été maîtrisée, améliorant ainsi la fiabilité et la stabilité des prédictions.

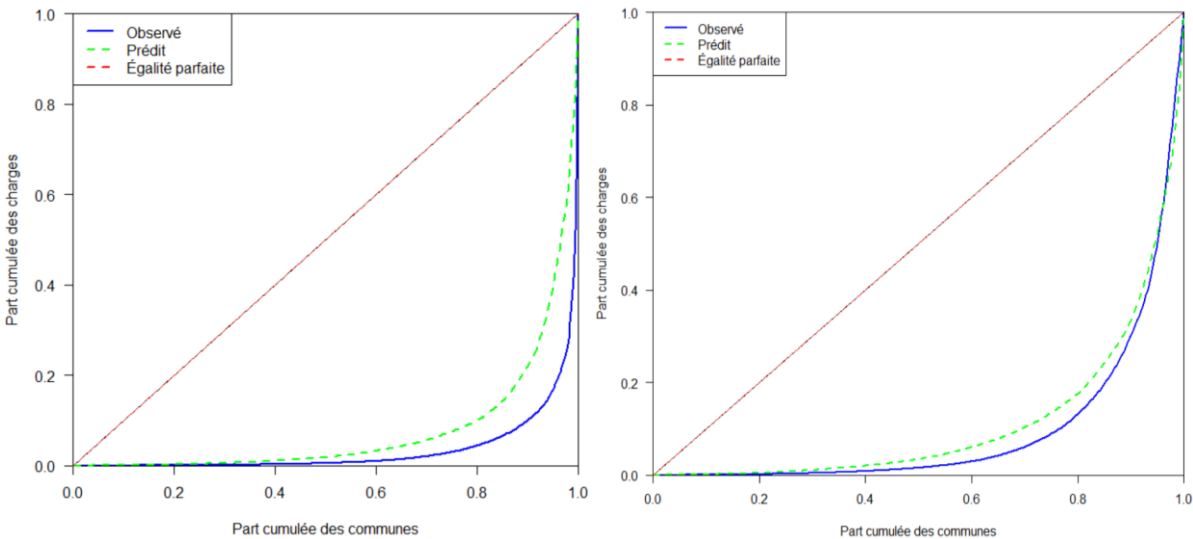


Figure 54 : courbe de Lorentz avec plafonnement à gauche et sans à droite. Réalisé sous R.

À gauche, le modèle sans plafonnement montre une courbe moins proche de la courbe observée, indiquant une sous-estimation des inégalités dans la répartition des sinistres. À droite, le modèle avec plafonnement suit de plus près les données observées, confirmant une meilleure adaptation. Les courbes de Lorenz mettent en évidence une concentration plus réaliste des charges dans le modèle avec plafonnement.

Back-test

Lors du back-test réalisé sur le cyclone Belal (2024), les performances du modèle sans plafonnement se sont avérées inférieures :

	Modèle sans plafonnement	Modèle avec plafonnement
Ecart par rapport à la charge observée	70%	0,7%
RMSE	783 036	235 556

Tableau 28 : Résultat du back-test avec et sans plafonnement. Réalisé sur Excel.

Le modèle avec plafonnement montre des performances nettement supérieures, avec une meilleure adéquation entre les charges observées et prédites.

Ainsi, le plafonnement apparaît comme une approche essentielle pour améliorer la robustesse et la fiabilité du modèle GLM dans le cadre de la modélisation des risques cycloniques de Prudence Créole.

L'estimation des paramètres d'un modèle linéaire généralisé se fait par maximum de vraisemblance, l'objectif est de résoudre :

$$\frac{\partial l(S_{j,e}, \theta_{j,e})}{\partial \beta} = 0.$$

Pour des raisons de simplification des notations, β désigne $(\beta_0, \beta) \in R^{p+1}$.

Les observations étant supposées suivre une loi appartenant à la famille exponentielle, la log-vraisemblance du modèle est donnée par :

$$l(\beta) = \sum_{j=1}^n \frac{S_{j,e} \theta_{j,e} - b(\theta_{j,e})}{a(\phi)} + c(S_{j,e}, \phi)$$

De plus,

$$\forall i \in \{1, \dots, n\}, \quad \theta_{j,e} = (b')^{-1} \circ g^{-1}(x_{j,e} \beta) = h(x_{j,e} \beta) \quad \text{où} \quad h = (b')^{-1} \circ g^{-1}.$$

D'où :

$$l(\beta) = \sum_{j=1}^n \frac{S_{j,e} h(x_{j,e} \beta) - b(h(x_{j,e} \beta))}{a(\phi)} + c(S_{j,e}, \phi) = \sum_{j=1}^n l_j(\beta)$$

Il faut résoudre les équations de vraisemblance, i.e. $(\forall i \in \{1, \dots, p\})$, il faut résoudre :

$$\begin{aligned} \frac{\partial l(\beta)}{\partial \beta_i} &= \sum_{j=1}^n \frac{\partial l_j(\beta)}{\partial \beta_i} = 0. \\ \Leftrightarrow \frac{\partial l(\beta)}{\partial \beta_i} &= \sum_{j=1}^n \frac{S_{j,e} x_{j,e}^i h'(x_{j,e} \beta) - x_{j,e}^i (b \circ h)'(x_{j,e} \beta)}{a(\phi)} \\ \Leftrightarrow \frac{\partial l(\beta)}{\partial \beta_i} &= \sum_{j=1}^n \frac{S_{j,e} x_{j,e}^i h'(x_{j,e} \beta) - x_{j,e}^i h'(x_{j,e} \beta) b' \circ h(x_{j,e} \beta)}{a(\phi)} \\ \Leftrightarrow \frac{\partial l(\beta)}{\partial \beta_i} &= \sum_{j=1}^n \frac{S_{j,e} x_{j,e}^i h'(x_{j,e} \beta) - x_{j,e}^i h'(x_{j,e} \beta) g^{-1}(x_{j,e} \beta)}{a(\phi)} \\ \Leftrightarrow \frac{\partial l(\beta)}{\partial \beta_i} &= \sum_{j=1}^n \frac{x_{j,e}^i h'(x_{j,e} \beta) (S_{j,e} - g^{-1}(x_{j,e} \beta))}{a(\phi)}. \end{aligned}$$

De plus,

$$\begin{aligned} h'(x_{j,e} \beta) &= ((b')^{-1} \circ g^{-1})'(x_{j,e} \beta) \\ \Leftrightarrow h'(x_{j,e} \beta) &= (g^{-1})' \times (((b')^{-1})' \circ g^{-1}(x_{j,e} \beta)) \\ \Leftrightarrow h'(x_{j,e} \beta) &= (g^{-1})'(x_{j,e} \beta) \times \frac{1}{b'' \circ (b')^{-1} \circ g^{-1}(x_{j,e} \beta)} \end{aligned}$$

$$\begin{aligned}
\Leftrightarrow h'(x_{j,e}\beta) &= (g^{-1})'(x_{j,e}\beta) \times \frac{1}{b'' \circ h(x_{j,e}\beta)} \\
\Leftrightarrow h'(x_{j,e}\beta) &= (g^{-1})'(x_{j,e}\beta) \times \frac{1}{b''(\theta_{j,e})} \\
\Leftrightarrow h'(x_{j,e}\beta) &= (g^{-1})'(x_{j,e}\beta) \times \frac{a(\phi)}{V(S_{j,e})}.
\end{aligned}$$

ainsi,

$$\frac{h'(x_{j,e}\beta)}{a(\phi)} = \frac{(g^{-1})'(x_{j,e}\beta)}{V(S_{j,e})},$$

Les équations de vraisemblance obtenues sont données par la formule suivante :

$$\sum_{j=1}^n \frac{x_{j,e}^i h'(x_{j,e}\beta) (S_{j,e} - g^{-1}(x_{j,e}\beta))}{a(\phi)} = \sum_{i=1}^n \frac{(g^{-1})'(x_{j,e}\beta) x_{j,e}^i (S_{j,e} - g^{-1}(x_{j,e}\beta))}{V(S_{j,e})} = 0 \quad \forall i \in \{1, \dots, p\}$$

Les équations de vraisemblance n'ont pas de solution explicite en général, sauf lorsque $(b' = \text{id})$. Des méthodes itératives, telles que l'algorithme de Newton-Raphson, peuvent être utilisées pour approcher la solution.

Loi GAMMA

Si une variable aléatoire Y suit une loi Gamma de paramètres $r > 0$ et $\alpha > 0$, sa densité est donnée par :

$$f_{r,\alpha}(y) = \frac{\alpha^r}{\Gamma(r)} y^{r-1} \exp(-\alpha y), \quad y \geq 0$$

Où $\Gamma(r)$ est la fonction Gamma définie par : $\Gamma(r) = \int_0^\infty u^{r-1} e^{-u} du$.

Les moments de la loi Gamma sont :

$$E[Y] = \mu = \frac{r}{\alpha} \text{ et } \text{Var}(Y) = \frac{r}{\alpha^2}.$$

Demonstrations:

$$E[Y] = \int_R y \frac{\alpha^r}{\Gamma(r)} y^{r-1} \exp(-\alpha y) dy$$

$$\Leftrightarrow E[Y] = \frac{\alpha^r}{\Gamma(r)} \int_0^\infty y^r \exp(-\alpha y) dy$$

$$\text{Posons } u = \alpha y, \quad y = \frac{u}{\alpha}, \quad dy = \frac{1}{\alpha} du$$

$$E[Y] = \frac{\alpha^r}{\Gamma(r)} \int_0^\infty \left(\frac{u}{\alpha}\right)^r \exp(-u) \frac{1}{\alpha} du$$

$$\Leftrightarrow E[Y] = \frac{\alpha^r}{\Gamma(r)} \frac{1}{\alpha^{r+1}} \int_0^\infty u^r \exp(-u) du$$

$$\begin{aligned}\Leftrightarrow E[Y] &= \frac{\alpha^r}{\Gamma(r)} \frac{1}{\alpha^{r+1}} \Gamma(r+1) \\ \Leftrightarrow E[Y] &= \frac{\Gamma(r+1)}{\alpha \Gamma(r)} \\ \Leftrightarrow E[Y] &= \frac{r}{\alpha}\end{aligned}$$

$$\begin{aligned}\text{Puis, } E[Y^2] &= \int_0^{\infty} y^2 \frac{\alpha^r}{\Gamma(r)} y^{r-1} \exp(-\alpha y) dy \\ \Leftrightarrow E[Y^2] &= \frac{\alpha^r}{\Gamma(r)} \int_0^{\infty} y^{r+1} \exp(-\alpha y) dy\end{aligned}$$

$$\begin{aligned}\text{Posons } u &= \alpha y, \quad y = \frac{u}{\alpha}, \quad dy = \frac{1}{\alpha} du \\ \Leftrightarrow E[Y^2] &= \frac{\alpha^r}{\Gamma(r)} \int_0^{\infty} \left(\frac{u}{\alpha}\right)^{r+1} \exp(-u) \frac{1}{\alpha} du \\ \Leftrightarrow E[Y^2] &= \frac{\alpha^r}{\Gamma(r)} \frac{1}{\alpha^{r+2}} \int_0^{\infty} u^{r+1} \exp(-u) du \\ \Leftrightarrow E[Y^2] &= \frac{\alpha^r}{\Gamma(r)} \frac{1}{\alpha^{r+2}} \Gamma(r+2) \\ \Leftrightarrow E[Y^2] &= \frac{\Gamma(r+2)}{\alpha^2 \Gamma(r)} \\ \Leftrightarrow E[Y^2] &= \frac{r(r+1)}{\alpha^2} \\ \text{Donc } \text{Var}(Y) &= \frac{r}{\alpha^2}\end{aligned}$$

En posant $\mu = \frac{r}{\alpha}$, la densité peut également s'écrire sous la forme suivante :

$$f_{r,\alpha}(y) = \frac{\left(\frac{r}{\mu}\right)^r}{\Gamma(r)} y^{r-1} e^{-\frac{ry}{\mu}}$$

Puis en simplifiant,

$$\begin{aligned}f_{r,\alpha}(y) &= \exp\left(-\frac{ry}{\mu} - r \ln(\mu) + r \ln(r) - \ln(\Gamma(r)) + \ln(y) (r-1)\right). \\ \Leftrightarrow f_{r,\alpha}(y) &= \exp\left(-\alpha y - r \ln(r) + r \ln(\alpha) + r \ln(r) - \ln(\Gamma(\alpha)) + \ln(y) (r-1)\right). \\ \Leftrightarrow f_{r,\alpha}(y) &= \exp\left(-\alpha y + r \ln(\alpha) - \ln(\Gamma(\alpha)) + \ln(y) (r-1)\right).\end{aligned}$$

Par identification, il est possible de retrouver les paramètres de la famille exponentielle,

$$\begin{aligned}\theta &= -\alpha = -\frac{r}{\mu} \\ b(\theta) &= r \ln(-\theta) \\ a(\phi) &= 1 \\ c(y, \phi) &= -\ln(\Gamma(\alpha)) + \ln(y) (r-1)\end{aligned}$$

La fonction de vraisemblance pour un ensemble des observations y_i (avec $i = 1, \dots, n$) est donnée par :

$$\mathcal{L}(y; \beta) = \prod_{i=1}^n \frac{1}{\Gamma(r)} \left(\frac{ry_i}{\mu_i} \right)^r \exp\left(-\frac{ry_i}{\mu_i}\right) \frac{1}{y_i}.$$

Afin de simplifier les calculs et d'optimiser les paramètres β , on prend le logarithme naturel de la fonction de vraisemblance, ce qui permet de transformer le produit en une somme :

$$\ln(\mathcal{L}(y; \beta)) = \sum_{i=1}^n \ln\left(\frac{1}{\Gamma(r)} \left(\frac{ry_i}{\mu_i} \right)^r \exp\left(-\frac{ry_i}{\mu_i}\right) \frac{1}{y_i}\right)$$

Puis on décompose chaque terme comme suit :

$$\ln(\mathcal{L}(y; \beta)) = \sum_{i=1}^n \left[\ln\left(\frac{1}{\Gamma(r)}\right) + r \ln\left(\frac{ry_i}{\mu_i}\right) - \frac{ry_i}{\mu_i} + \ln\left(\frac{1}{y_i}\right) \right]$$

Ensuite, chaque composant est simplifié :

- $\ln\left(\frac{1}{\Gamma(r)}\right) = -\ln(\Gamma(r))$, un terme constant qui ne dépend pas de β .
- $r \ln\left(\frac{ry_i}{\mu_i}\right) = r \ln(r) + r \ln(y_i) - r \ln(\mu_i)$.
- $\ln\left(\frac{1}{y_i}\right) = -\ln(y_i)$, un autre terme constant.
- $-\frac{ry_i}{\mu_i}$ reste inchangé.

En regroupant les termes dépendant de μ_i et en éliminant les constantes $(-\ln(\Gamma(r)), -\ln(y_i), r \ln(r))$ qui n'interviennent pas dans la maximisation, on obtient :

$$\ln(\mathcal{L}(y; \beta)) = \sum_{i=1}^n \left[-\frac{ry_i}{\mu_i} - r \ln(\mu_i) \right].$$

La log-vraisemblance finale, exprimée en fonction des paramètres β , s'écrit donc :

$$l(y; \beta) = \sum_{i=1}^n r \left(-\frac{y_i}{\mu_i} - \ln(\mu_i) \right).$$

Pour estimer $\mu_i = \exp(\beta_0 + X_i' \beta)$, il suffit de résoudre le système d'équations suivant :

$$\frac{\partial}{\partial \beta_j} \sum_{i=1}^n r \left(-\frac{y_i}{\exp(\beta_0 + X_i' \beta)} - (\beta_0 + X_i' \beta) \right) = 0, \quad \text{pour } j = 1, \dots, p$$

1) Renseignement de l'adresse



PRUDENCE CRÉOLE

**DIAGNOSTIC D'EXPOSITION
AUX RISQUES CLIMATIQUES,
NATURELS ET
TECHNOLOGIQUES**

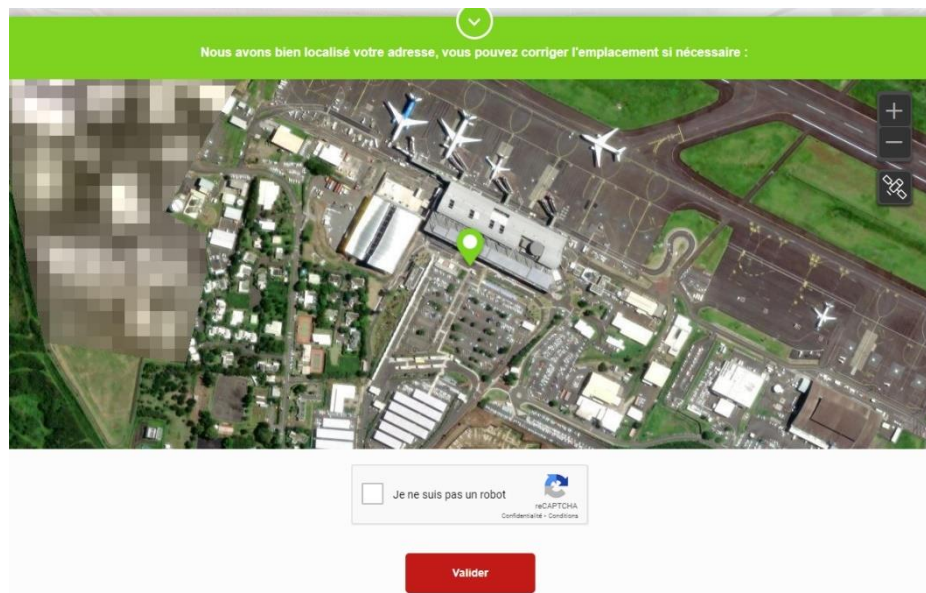
Réalisez le diagnostic d'exposition aux risques à l'adresse de votre choix à la Réunion ou à Mayotte et obtenez des conseils de prévention.

Adresse

La réalisation du diagnostic, personnalisé et gratuit, et son envoi nécessitent l'utilisation des données de votre adresse postale, conformément à votre demande.

Le traitement de ces données, par Generali IARD, a pour base juridique le consentement et a pour finalité la réalisation ainsi que l'envoi du diagnostic.

2) Validation de la localisation avec la possibilité de la modifier en cas d'erreur.

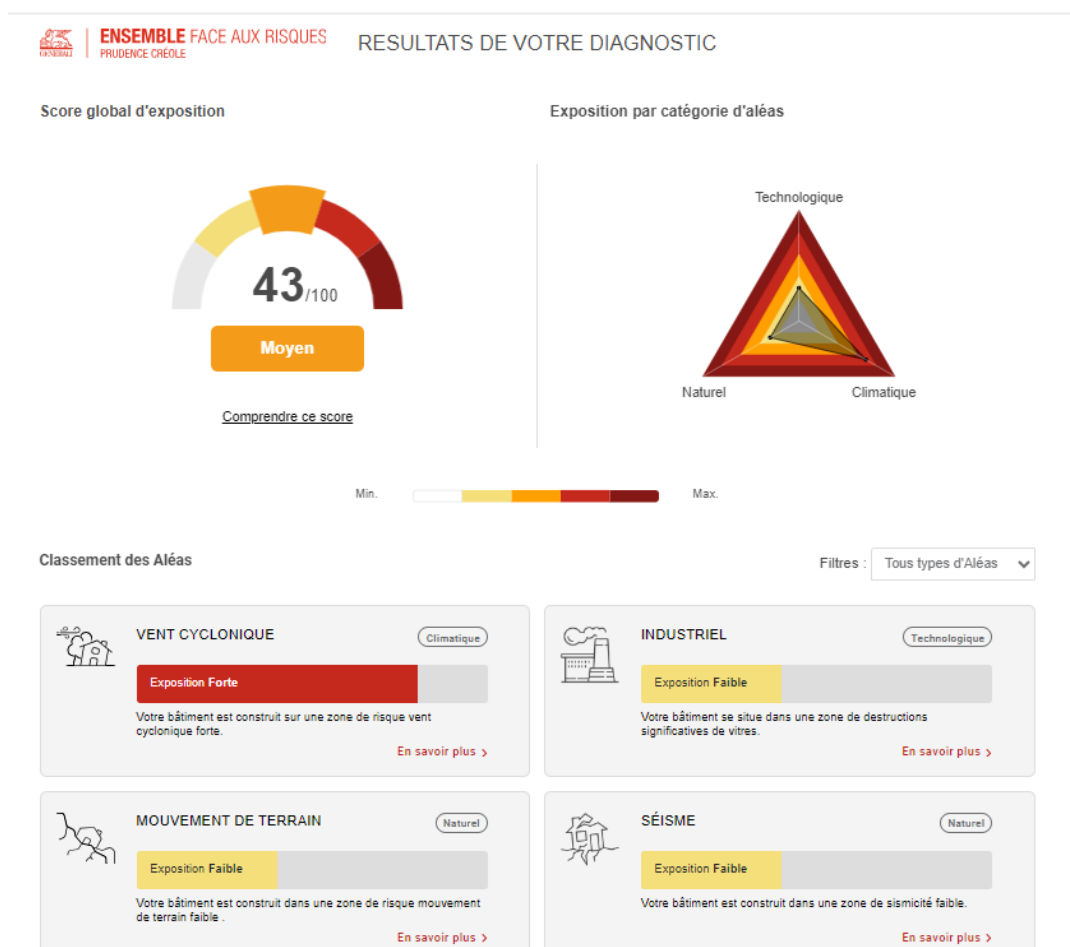


Nous avons bien localisé votre adresse, vous pouvez corriger l'emplacement si nécessaire :

☐ Je ne suis pas un robot

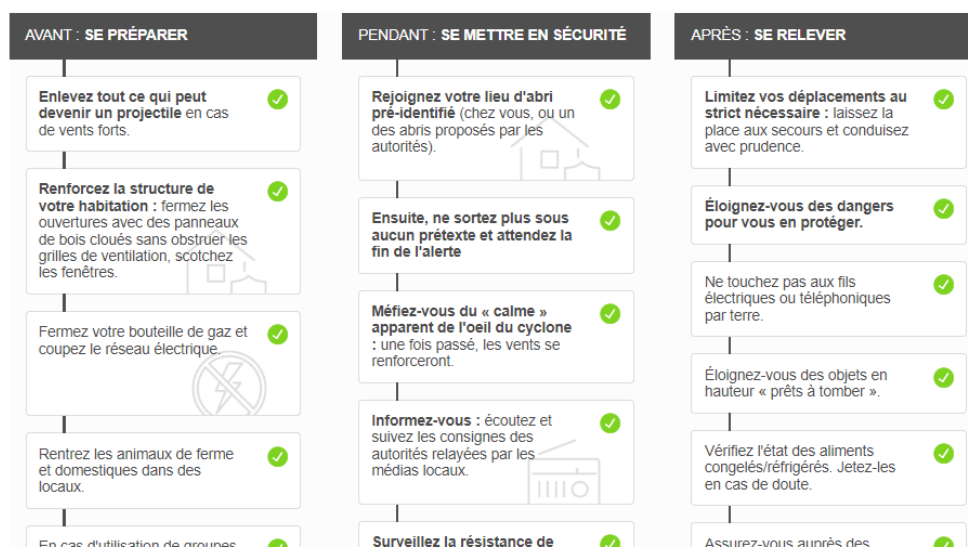
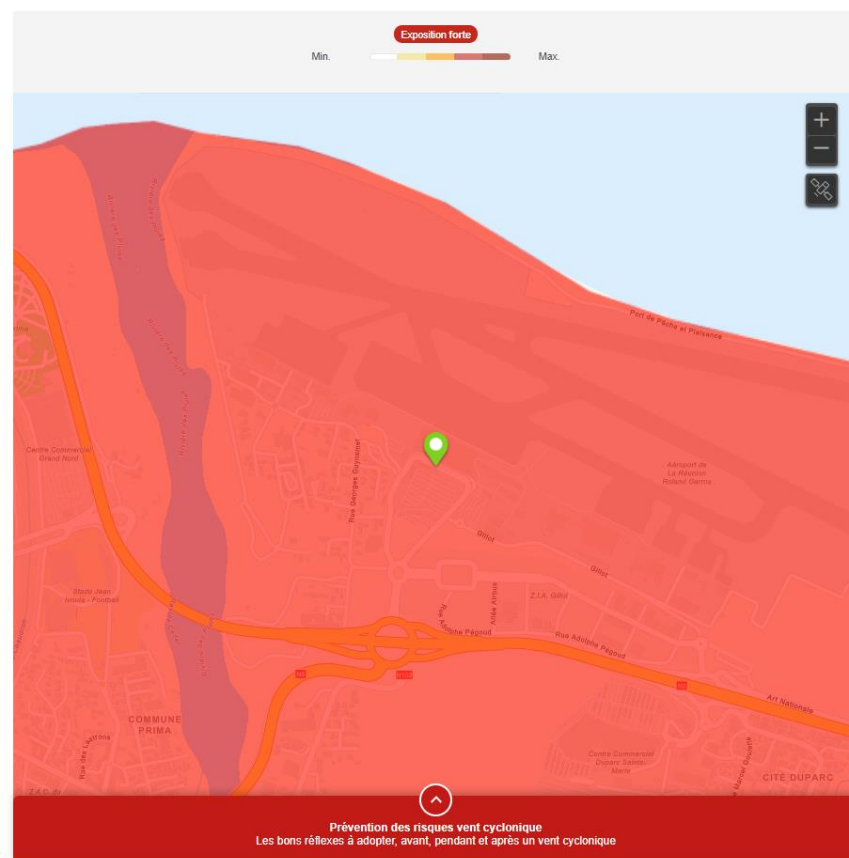
VALIDER

3) Accès au diagnostic avec la possibilité de zoomer sur chacun des périls



4) Exemple : Zoom sur le vent cyclonique

Après avoir cliqué sur un péril, la carte de l'aléa s'affiche avec les conseils de prévention avant, pendant et après l'occurrence du péril.



A chaque diagnostic de péril, une fiche de prévention est également associée.

Toujours dans le cas du vent cyclonique :

**PRUDENCE CRÉOLE PRÉVENTION
VENT CYCLONIQUE**



**PRUDENCE
CRÉOLE**

VENT CYCLONIQUE : QUELLES PRÉCAUTIONS PRENDRE ?



Nos conseils pratiques de prévention climatique



Un cyclone est un phénomène météorologique caractérisé par des pluies et des vents violents supérieurs à 117 km/h.

La Réunion et Mayotte se situent dans le bassin cyclonique du Sud-Ouest de l'Océan Indien, où la saison cyclonique se déroule de novembre à avril, avec un maximum de risque sur les trois mois d'été austral, entre janvier et mars. Cependant, il a déjà été observé dans cette région du monde, des cyclones matures dès octobre et jusqu'en mai.

QUE FAIRE POUR SE PRÉPARER EN DÉBUT DE CHAQUE SAISON ?

INSPECTEZ VOTRE HABITAT

- Vérifiez et renforcez si nécessaire la structure de votre habitation.
- Nettoyez les abords de votre logement.
- Enlevez, à l'extérieur, tout objet pouvant être emporté et devenir un projectile dangereux en cas de vents forts.

CHOISISSEZ L'ENDROIT OÙ VOUS VOUS METTREZ À L'ABRI EN CAS DE CYCLONE

- Repérez les lieux les plus solides du bâtiment et éloignés des baies vitrées.
- Si votre habitation n'est pas sûre, informez-vous sur les lieux de protection et d'accueil proposés par les autorités.
- Indiquez à vos proches où se trouve votre lieu d'abri et mutualisez les coordonnées de votre famille.

PRÉVOYEZ D'ORES ET DÉJÀ LA CRISE ET L'APRÈS-CRISE

- Vérifiez et préparez les équipements nécessaires aux situations d'urgence.
- Constituez des stocks pour l'après-crise.
- Placez vos réserves dans des sacs étanches.
- Stockez-les dans un lieu accessible et abrité.

QUE FAIRE À L'ANNONCE D'UN CYCLONE ?

PROTÉGEZ VOTRE HABITAT ET VOS BIENS

- Enlevez tout ce qui peut devenir un projectile en cas de vents forts.
- Renforcez la structure de votre habitation :
 - fermez les ouvertures avec des panneaux de bois cloués ;
 - n'obstruez pas les grilles de ventilation ;
 - scotchez les fenêtres.
- Fermez votre bouteille de gaz et coupez le réseau électrique pour notamment éviter les incendies.

- En cas d'utilisation de groupes électrogènes, veillez à les placer à l'extérieur du bâtiment et à respecter les consignes d'utilisation.
- Amarrez les canots le plus loin possible du rivage.
- Rentrez les animaux de ferme et domestiques dans des locaux.

SOYEZ VIGILANT

Ne vous approchez pas du rivage en cas de forte houle et ne franchissez jamais un radier submergé.

QUE FAIRE LORSQUE LE CYCLONE SURVIENT ?

METTEZ-VOUS EN SÉCURITÉ

- Rejoignez votre lieu d'abri pré-identifié (chez vous, ou un des abris proposés par les autorités).
- Ensuite, ne sortez plus sous aucun prétexte et attendez la fin de l'alerte.
- Méfiez-vous du « calme » apparent de l'œil du cyclone : une fois passé, les vents se renforceront.

SOYEZ ATTENTIF

- Informez-vous : écoutez et suivez les consignes des autorités relayées par les médias locaux.
- Surveillez la résistance de l'abri et le risque d'inondation.

TENEZ-VOUS INFORMÉ DE LA SITUATION

- Ne téléphonez qu'en cas d'urgence : laissez les secours disposer au mieux des réseaux.
- Redoublez de vigilance la nuit.

QUE FAIRE APRÈS LE CYCLONE ?

LIMITEZ VOS DÉPLACEMENTS AU STRICT NÉCESSAIRE

- Laissez la place aux secours.
- Conduisez avec prudence.

ÉLOIGNEZ-VOUS DES DANGERS POUR VOUS EN PROTÉGER

- Ne touchez pas aux fils électriques ou téléphoniques par terre.
- Éloignez-vous des objets en hauteur « prêts à tomber ».
- Vérifiez l'état des aliments congelés/réfrigérés. Jetez-les en cas de doute.

- Assurez-vous auprès des autorités locales que l'eau du robinet est potable. Dans tous les cas, faites couler l'eau avant de la consommer.
- En cas d'utilisation de groupes électrogènes, veillez à les placer à l'extérieur du bâtiment et à respecter les consignes d'utilisation.
- Réparez et confortez votre habitation, et faites votre déclaration à votre assureur.

