INSTITUT DES ACTUAIRES

ISUP

Mémoire présenté le :
pour l'obtention du diplôme
de Statisticien Mention Actuariat
et l'admission à l'Institut des Actuaires

Par : Arnaud Dalleau

Évaluation de l'impact du changement climatique sur le risque cyclonique dans les Antilles

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

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

Membres présents du jury de l'Institut des Actuaires signature

Entreprise :

Nom : Groupama Assurances Mutuelles

Signature :

Directeur de mémoire en entreprise :

Membres présents du jury de la filière

Nom : Marc Bagarry

Signature :

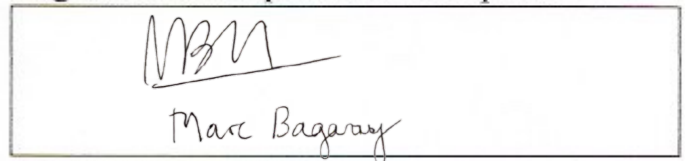
Invité :

Nom :

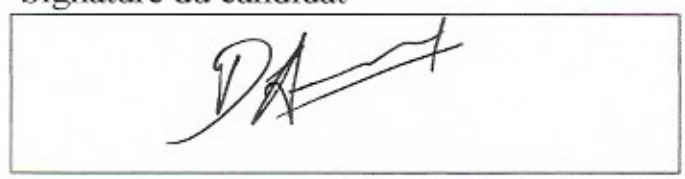
Signature :

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

Signature du responsable entreprise



Signature du candidat



Résumé

Comme l'indiquent les rapports du GIEC et de la CCR, le réchauffement climatique impactera très probablement l'activité cyclonique dans le bassin des Antilles d'ici la fin du siècle. Bien que la nature de cet impact ne soit pas encore clairement identifiée, les assureurs implantés dans des zones fortement exposées au risque cyclonique se doivent de suivre toute évolution de ce dernier afin de parer à toute éventualité.

Aussi bien pour les assureurs que pour l'ensemble des acteurs privés et publics exposés, apparaît alors la nécessité de se projeter et donc de simuler le climat futur. Face à l'ampleur de l'impact du réchauffement de notre climat sur notre société, la communauté scientifique a souhaité rendre accessibles certaines projections. Cette démarche s'inscrit dans un contexte plus global de données ouvertes (ou *open-data*) : des données auparavant réservées à des communautés expertes sont aujourd'hui accessibles librement *via* internet.

Un des objectifs de ce mémoire est de montrer qu'il est possible pour un assureur de tirer profit du contexte de données ouvertes actuel. Dans notre cas, il nous permet de fournir une estimation l'impact financier du réchauffement climatique sur le risque cyclonique pour Groupama, assureur exposé à ce risque dans le bassin des Antilles.

Dans un premier temps, nous nous attachons à extraire les informations relatives à l'activité cyclonique dans nos données, à savoir les ouragans et leurs caractéristiques principales. Cette extraction est réalisée grâce à un système de *tracking* : les ouragans s'apparentent à des objets à suivre sur un laps de temps indéterminé.

Une fois les informations extraites, nous pouvons étudier l'activité cyclonique observable sur les projections climatiques en s'intéressant au nombre annuel d'ouragans, à leur puissance ou encore à leur trajectoire. Cette étude prend la forme d'une comparaison entre l'activité observable sur les simulations du climat futur et celle observable sur les simulations aux conditions historiques. Afin d'identifier les biais des modèles climatiques générant ces projections, l'activité cyclonique générée sous conditions historiques est également comparée à l'activité observée historiquement.

Pour transformer l'activité cyclonique observée sur les projections en pertes financières, nous faisons appel à un modèle catastrophe existant proposé par un modélisateur extérieur à Groupama. Deux approches sont alors envisagées pour coupler ce modèle catastrophe aux informations extraites des projections.

La première, reposant grandement sur les projections climatiques, engendre des pertes difficilement exploitables, notamment à cause des biais des modèles climatiques. La seconde, n'exploitant que partiellement les projections, engendre des pertes stables au cours du XXI^{ème} siècle dans le cas d'émissions de gaz à effet de serre décroissantes dès 2050. Dans le scénario le plus pessimiste, les pertes constatées en fin de siècle sont bien plus élevées qu'historiquement et ceci est notamment lié à une proportion grandissante d'ouragans puissants.

Abstract

As the IPCC and Caisse Centrale de Réassurance reports indicate, global warming will most likely impact the cyclonic activity in the Caribbean Basin by the end of the century. Although the nature of this impact is not yet clearly identified, insurance companies doing business in areas highly exposed to cyclone risk must monitor any changes in this risk in order to be prepared for any eventuality.

Both for the insurers and for all the private and public stakeholders exposed, it is therefore necessary to make projections, in other words to simulate the future climate. Faced with the severity of the impact of global warming on our society, the scientific community wanted to make certain projections accessible. This approach is part of a more global context of open-data: data previously reserved for expert communities are now freely accessible *via* the Internet.

One of the objectives of this paper is to show that it is now possible for an insurer to take advantage of the current open-data context. In our case, it enables us to provide an estimate of the financial impact of global warming on the cyclonic risk for Groupama, an insurer exposed to this risk in the Caribbean basin.

First, we focus on extracting information on cyclonic activity from our data, that is to say hurricanes and their main characteristics. This extraction is carried out using a tracking system: hurricanes appear as objects to be tracked over an undetermined period of time.

Once the information has been extracted, we can study the cyclonic activity observable on the climate projections by looking at the annual number of hurricanes, their intensity or their trajectory. This study takes the form of a comparison between the activity observable on simulations of the future climate and that observable on simulations under historical conditions. In order to identify the biases of the climate models generating these projections, the simulations under historical conditions are also compared with the hurricanes observed historically.

To transform the observed cyclonic activity into financial losses, we use an existing catastrophe model proposed by an external modeller. Two approaches are then envisaged to couple this catastrophe model to the information extracted from the projections.

The first, which relies heavily on climate projections, generates losses that are difficult to exploit, particularly because of the biases of climate models. The second, only partially exploiting the projections, generates stable losses during the 21st century in the case of decreasing greenhouse gas emissions from 2050 onwards. In the most pessimistic scenario, the losses observed at the end of the century are much higher than historically and this is notably linked to a growing proportion of powerful hurricanes.

Synthèse

Contexte

Le cinquième rapport du GIEC, le Groupe d'experts Intergouvernemental sur l'Évolution du Climat, constitue la dernière évaluation détaillée de l'état des connaissances scientifiques, techniques, socio-économiques actuelles sur les changements climatiques, leurs causes et leurs répercussions.

Selon les experts du premier groupe de travail, l'augmentation globale des températures ainsi que du niveau des mers à l'horizon 2100 ne fait plus aucun doute, de même que le rôle majeur de l'Homme dans ces évolutions. En revanche, pour des phénomènes plus locaux tels que les ouragans, de nombreux doutes subsistent.

Face à ces doutes et aux conséquences que pourrait avoir le réchauffement climatique sur le risque cyclonique et donc sur la santé financière des assureurs implantés dans des régions exposées à ce risque, la CCR a évalué l'impact financier de ce réchauffement sur le risque cyclonique en Outre-mer à l'horizon 2050. Là encore, les incertitudes sont nombreuses et la CCR fait part de « résultats très divergents » selon le modèle climatique utilisé pour se projeter. La CCR est cependant confiante sur l'augmentation à venir des dégâts liés à la submersion marine et aux fortes précipitations induites par les cyclones.

Ainsi, alors même que l'attractivité des Antilles pour les assureurs est déjà affaiblie en raison, notamment, du risque cyclonique élevé, de possibles évolutions de ce dernier pourraient encore davantage l'affaiblir. Les assureurs exposés dans les Antilles comme Groupama doivent donc dès maintenant se projeter afin de parer à toute éventualité, en s'appuyant sur le régime CAT-NAT existant.

Dans le cadre de ce mémoire, nous saisissons l'opportunité créée par le contexte de données ouvertes actuel : des projections climatiques sont librement accessibles *via* internet et il est alors possible pour Groupama d'accéder à une activité cyclonique simulée jusqu'à la fin du siècle pour la zone Antilles.

Éléments de météorologie

Avant d'exploiter les projections climatiques pour la zone Antilles, c'est-à-dire d'extraire l'activité cyclonique observable, il est important de comprendre ce qu'est un cyclone.

Un cyclone tropical (ou ouragan tropical) est une perturbation particulièrement puissante se formant dans la zone tropicale, délimitée par les tropiques du Cancer et du Capricorne. Cette perturbation se forme sous l'action de la force de Coriolis et au-dessus d'une eau suffisamment chaude permettant le phénomène d'évaporation.

Dès lors que la vitesse des vents engendrés dépasse 33 m/s, la perturbation passe au stade d'ouragan.

Afin de caractériser sa puissance, différentes échelles ont été mis en place. La plus populaire de toutes est l'échelle de Saffir-Simpson :

	Vitesse de vent maximale enregistrée	
	m/s	km/h
Ouragan catégorie 5	≥ 70	≥ 252
Ouragan catégorie 4	59-70	209-251
Ouragan catégorie 3	50-58	178-208
Ouragan catégorie 2	43-49	154-177
Ouragan catégorie 1	33-42	119-153
Tempête tropicale	18-32	63-118
Dépression tropicale	≤ 17	≤ 62

Tableau 1 - Échelle de Saffir-Simpson (source : NOAA)

N'exploitant que la vitesse maximale des vents engendrés, cette échelle est particulièrement simple à utiliser. Étant donnée la faiblesse des vitesses de vent engendrées par les tempêtes tropicales et les dépressions, nous nous focalisons sur l'étude des ouragans de catégorie 1 ou plus.

Projections climatiques utilisées et extraction des ouragans

Les projections climatiques utilisées dans le cadre de ce mémoire sont celles issues du projet CORDEX (***CO****ordinated* ***R****egional* ***D****ownscaling* ***EX****periment*). L'objectif de ce projet est double : d'une part donner un cadre de référence pour la production de projections climatiques régionales, d'autre part fournir une interface permettant à toute personne d'avoir accès à ces simulations *via* internet.

Les projections utilisées sont réalisées sur le domaine *Central America*, domaine couvrant le bassin des Antilles. La variable utilisée est la vitesse des vents instantanée à 10 mètres de hauteur : cette dernière est projetée sur l'ensemble du domaine (avec une maille de 25 km) jusqu'en 2100 avec un pas de temps de 3 heures. Ces projections sont générées par 3 modèles de circulation générale (ou GCM) : HadGEM2-ES, fortement sensible au changement climatique, MPI-ESM, moyennement sensible au changement climatique et NorESM, faiblement sensible au changement climatique. Les projections étant générées sur une maille très large, un modèle climatique régional (ou RCM) est ensuite appliqué afin d'obtenir les projections sur la maille de 25 km.

Réaliser des projections climatiques jusqu'en 2100 implique de devoir prendre en compte différentes évolutions possibles de l'environnement socio-économique terrestre. Ces évolutions sont formalisées par des scénarii RCP, *Representative Concentration Pathway*, décrivant les émissions de gaz à effet de serre tout au long du siècle. Deux scénarii sont utilisés dans notre cas : le scénario RCP 2.6, traduisant une baisse progressive des émissions de gaz à effet de serre à partir de 2050 et le scénario RCP 8.5, scénario le plus pessimiste traduisant une inaction globale dans la réduction des émissions de gaz à effet de serre.

En plus des scénarii RCP utilisés pour projeter le climat, des simulations sont également réalisées sous conditions historiques :

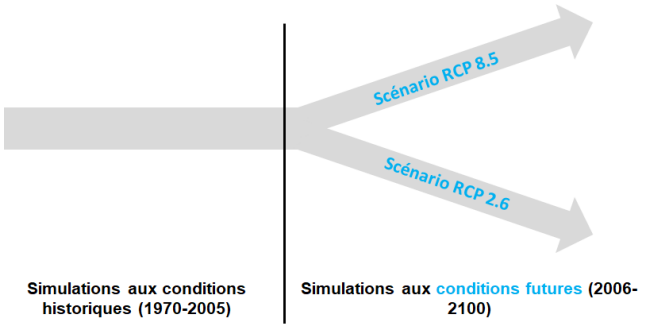


Figure 1 - Structure des scénarii utilisés pour simuler les données

Ces simulations, une fois comparées aux observations historiques, permettent d'évaluer les biais des modèles climatiques générant les projections. Aussi bien en scénario RCP qu'en conditions historiques, nos projections se présentent sous la forme d'une succession de matrices dont les coefficients donnent la vitesse des vents en un endroit du domaine CAM :

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,417} \\ \vdots & \ddots & \vdots \\ a_{225,1} & \cdots & a_{225,417} \end{pmatrix}$$

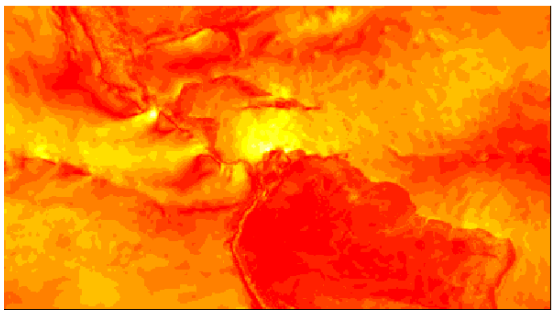


Figure 2 – Forme des matrices composant notre jeu de données (gauche) et visualisation d'une de ces matrices (droite), les coefficients jaunes correspondent aux vitesses élevées.

Afin d'extraire l'activité cyclonique présente dans notre jeu de données, c'est-à-dire d'extraire les ouragans observables sur les simulations climatiques, un système de *tracking* est mis en place : les ouragans sont vus comme des objets à suivre sur un laps de temps indéterminé.

Dans un premier temps, filtrer les matrices au seuil de 33 m/s permet d'avoir un aperçu des pics d'intensité des ouragans présents dans notre jeu de données. L'algorithme Mean-Shift, une fois appliqué sur ces matrices, permet d'obtenir le nombre et la localisation des ouragans. Il s'agit ensuite de s'intéresser aux matrices situées entre ces pics d'activité cyclonique : après les avoir filtrées au seuil de 16 m/s, l'algorithme Mean-Shift permet une nouvelle fois de donner le nombre et la localisation des perturbations présentes. Suivre un ouragan sur deux matrices successives i et $i+1$ revient alors à chercher la perturbation en $i+1$ la plus proche de l'ouragan situé en i. Nous appliquons la règle suivante : si la perturbation la plus proche en $i+1$ est située à moins de 120 km de l'ouragan observé en i, le rattachement a lieu et nous considérons qu'il s'agit d'un seul et même ouragan.

Divers retraitements doivent ensuite être faits. Nous excluons dans un premier temps toutes les perturbations se formant dans le Pacifique. Dans 95% des cas, les ouragans durent plus de 4.5

jours et nous excluons donc de nos données toute perturbation dont la durée de vie est inférieure à ce seuil. Les coefficients des matrices, c'est-à-dire les vitesses de vent, doivent également être corrigés. En effet, aucun ouragan de catégorie 4 ou 5 n'est visible sur les simulations : une correction quantile-quantile est alors utilisée.

Étude des ouragans extraits

Dans un premier temps, nous comparons les ouragans générés sous conditions historiques aux ouragans observés sur la période 1970-2005. La base de données historiques utilisée est la base HurDat, base de référence créée par le National Hurricane Center situé à Miami et étudiant les ouragans du bassin des Caraïbes.

	HadGEM2-ES	**MPI-ESM**	**NorESM**
Fréquence annuelle et mensuelle	Adéquation	Inadéquation	Adéquation
Empreintes	Inadéquation	Adéquation	Adéquation
Vitesses de vent	Adéquation partielle	Adéquation	Inadéquation

Tableau 2 - Points d'adéquation/inadéquation entre ouragans observés historiquement et ouragans simulés selon le GCM utilisé

Ces points d'inadéquation peuvent surprendre au premier abord mais les modèles utilisés projettent de très nombreuses variables (une centaine) jusqu'en 2100. L'adéquation partielle avec les données historiques d'un phénomène aussi local et extrême que les ouragans, observé à l'aide d'une seule de ces variables, ne doit donc pas surprendre.

Dans le cadre du scénario RCP 2.6, nous pouvons constater que l'activité cyclonique observable sur les simulations est très similaire à celle observée sous conditions historiques et ce, peu importe le GCM utilisé. Le nombre annuel d'ouragans générés est constant tout au long du siècle, aucune déviation des trajectoires n'est à constater et les proportions de catégories de Saffir-Simpson détectées sont semblables à celles obtenues sous conditions historiques.

Dans le cadre du scénario RCP 8.5, nous pouvons constater que le changement climatique impacte le nombre annuel d'ouragans générés dans la deuxième partie du siècle ainsi que la puissance de ces derniers. Si les modèles sont unanimes sur la diminution des proportions d'ouragans faiblement puissants et l'augmentation des proportions d'ouragans les plus puissants générés, l'impact sur le nombre annuel d'ouragans est quant à lui moins tranché. Deux des trois modèles sont en accord avec les prévisions du GIEC, à savoir un nombre d'ouragans annuel stable ou en légère baisse, mais le modèle le plus sensible au changement climatique indique, lui, une augmentation légère de ce nombre passé 2050. Les modèles ne sont pas non plus unanimes quant à l'évolution de la fréquence à laquelle sont frappés les territoires de la caisse Antilles-Guyane.

Modèle catastrophe utilisé pour l'estimation des pertes financières

Afin de transformer l'activité cyclonique observée sur les simulations climatiques en pertes financières pour Groupama, nous faisons appel à un modèle catastrophe proposé par un modélisateur A, calibré sur la période 1950-2007 de la base de données historiques HurDat. Il est donc essentiel de l'évaluer.

Dans un premier temps, l'étude des ouragans utilisés par le modèle révèle que ces derniers passent tous à proximité des expositions de la caisse : l'activité cyclonique utilisée par le modèle ne correspond donc pas à une activité cyclonique traditionnelle. Notons également que les ouragans utilisés par le modèle sont légèrement moins puissants que ceux observés sur les données historiques.

Concernant les pertes engendrées, celles-ci sont disponibles île par île pour un même ouragan et correspondent aux pertes engendrées par le péril vent ainsi que par les fortes précipitations. Un effet d'amplification est également intégré pour les événements majeurs. Bien que l'historique de pertes disponible pour Groupama ne soit que très peu profond, nous pouvons constater que les périodes de retour générées pour les événements Irma, Maria et Dean sont trop courtes.

Le modèle utilisé est donc conservateur et il faudra en tenir en compte lors de l'estimation des pertes financières pour le Groupe.

Estimation des pertes financières

Deux méthodes de couplage entre le modèle catastrophe et les ouragans observés sur les simulations climatiques sont proposées : la première exploite au maximum les informations tirées des simulations, la seconde repose davantage sur le modèle catastrophe proposé par le modélisateur.

La première méthode de couplage, méthode de « recherche d'ouragans voisins », s'inspire de l'algorithme des k plus proches voisins. La méthode consiste à associer à un ouragan observé sur les simulations climatiques des ouragans similaires provenant de la table du modélisateur. Cette association se fait île par île : les caractéristiques physiques de l'ouragan à proximité d'une des îles sont observées et 3 ouragans aux caractéristiques similaires proposés par le modélisateur sont associés. La perte engendrée sur l'île considérée est égale à la moyenne des 3 pertes des ouragans associés.

Bien que l'association fonctionne, c'est-à-dire que les ouragans observés sur les simulations sont associés à des ouragans similaires proposés par le modélisateur, les biais des modèles climatiques utilisés, et notamment les trajectoires, ne permettent pas d'obtenir des pertes exploitables. De plus, pour un scénario RCP et un GCM donnés, le XXI^{ème} siècle n'est simulé qu'une seule fois : le calcul de périodes de retour robustes est alors impossible.

La seconde méthode de couplage, méthode de « resimulation », repose moins sur les projections climatiques et permet alors de contourner les différents biais des modèles. Cette méthode consiste à simuler 10 000 années d'activité cyclonique : chaque année, un nombre d'ouragans N est simulé (N suit une loi Binomiale Négative) et la catégorie de chaque ouragan est déterminée de façon aléatoire à partir des proportions de catégories de Saffir-Simpson observées. Pour simuler une sinistralité à un horizon donné, les proportions utilisées sont celles observées sur les simulations climatiques à cet horizon.

Si les pertes générées par cette dernière méthode en scénario RCP 2.6 sont très similaires à celles obtenues sous conditions historiques, les pertes générées à l'horizon 2100 dans le cadre du scénario le plus pessimiste à l'égard des émissions de gaz à effet de serre sont, elles, bien supérieures à celles obtenues sous conditions historiques.

Remerciements

Je tiens tout d'abord à remercier l'ensemble de la Direction Réassurance de Groupama et plus particulièrement l'équipe Actuariat, au sein de laquelle ces travaux ont été réalisés.

Je tiens à remercier Marc pour m'avoir introduit le sujet, aiguillé sur la façon de l'aborder et pour m'avoir permis d'exposer une partie de ces travaux au Colloque Actuariel de Paris 2020. Un grand merci à Vincent pour son soutien, ses conseils techniques et m'avoir permis d'utiliser son ordinateur personnel pendant la période du confinement afin de réaliser ce mémoire. Merci à Jocelyn pour son soutien, pour avoir pris le temps de répondre à toutes mes questions et pour sa bonne humeur.

Merci à Guillaume Levavasseur, climatologue à l'IPSL, pour ses précieuses explications concernant la modélisation en météorologie. Merci à Lars Bärring, climatologue au sein de l'Institut Suédois de Météorologie et d'Hydrologie et associé au projet CORDEX, pour m'avoir éclairé sur certains aspects techniques du projet.

Je tiens également à remercier mon tuteur académique, Olivier Lopez, pour m'avoir conseillé sur la structure à adopter dans ce mémoire.

Enfin, un grand merci à ma famille et à mes amis qui m'ont soutenu tout au long de ces 5 années d'études.

Confidentialité

Dans un souci de confidentialité, les montants de pertes présentés dans ce mémoire, aussi bien les montants de pertes historiques enregistrés par le Groupe que ceux simulés par le modélisateur, ont été modifiés.

Table des matières

Résumé 3
Abstract 4
Synthèse 6
Contexte 6
Éléments de météorologie 6
Projections climatiques utilisées et extraction des ouragans 7
Étude des ouragans extraits 9
Modèle catastrophe utilisé pour l'estimation des pertes financières 10
Estimation des pertes financières 10
Remerciements 12
Confidentialité 12
I. Contexte 15
1. Cinquième rapport du GIEC 15
2. Rapport de la Caisse Centrale de Réassurance 16
3. Présentation de GMA et de la caisse Antilles-Guyane 17
4. Régime d'indemnisation des catastrophes naturelles 19
II. Éléments de météorologie 20
1. Dépressions, tempêtes et ouragans en zone tropicale 20
2. Caractériser la puissance d'un ouragan 22
3. Réaliser des projections 24
III. Projections climatiques utilisées et extraction des ouragans 26
1. Projet CORDEX 26
2. Cadre associé au projet et données utilisées 27
3. Mise en place d'un système de *tracking* 33
4. Retraitements 42
IV. Étude des ouragans extraits 48
1. Présentation de la base de données historiques 48
2. Études des ouragans générés aux conditions historiques 50
3. Études des ouragans générés selon le scénario RCP 2.6 58
4. Études des ouragans générés selon le scénario RCP 8.5 64
V. Modèle catastrophe utilisé pour l'estimation des pertes financières 69

1. Modèles catastrophes 69
2. Étude de la table d'événements 71
3. Évaluation des pertes proposées par la table 77
VI. Estimation des pertes financières 78
1. Méthode de recherche d'ouragans voisins 79
2. Méthode de resimulation à partir de la table 87
Conclusion générale 96
Bibliographie 98
Annexes 100
Suivi des ouragans 100
Algorithme Mean-Shift 102
Algorithme des ***k*** plus proches voisins 103
Estimation par la méthode du maximum de vraisemblance 104

I. Contexte

1. Cinquième rapport du GIEC

Le Groupe d'experts Intergouvernemental sur l'Évolution du Climat est un organisme intergouvernemental ouvert à tous les pays membres de l'ONU. Depuis 30 ans, le GIEC fournit des évaluations détaillées de l'état des connaissances scientifiques, techniques, socio-économiques sur les changements climatiques, leurs causes et leurs répercussions.

La dernière évaluation proposée date de 2014 et constitue le cinquième rapport du GIEC (ou *Fifth Assessment Report*). Ce dernier résulte des travaux de 3 groupes de travail rassemblant des milliers d'experts et de scientifiques à travers le monde. Le premier groupe s'intéresse aux aspects scientifiques de l'évolution du climat, le second traite de la vulnérabilité des systèmes socio-économiques aux changements climatiques, le troisième évalue les solutions envisageables pour atténuer ces changements. Attardons-nous donc sur les travaux du premier groupe de travail.

Le premier groupe s'attache à mettre en évidence les mécanismes physiques à l'œuvre dans l'évolution du climat ainsi que les causes potentielles d'évolution (en particulier le rôle des activités humaines) : les températures, les concentrations en gaz à effet de serre, en aérosols, l'élévation du niveau de la mer ou encore l'évolution des événements extrêmes sont ainsi étudiées. Des projections sont faites et probabilisées à l'aide de termes définis précisément :

Terme*	Probabilité d'occurrence
Quasiment certain	99% - 100%
Très probable	90% - 100%
Probable	66% - 100%
Aussi probable que non probable	33% - 66%
Improbable	0% - 33%
Très improbable	0% - 10%
Exceptionnellement improbable	0% - 1%

*Notons que des termes additionnels (extrêmement probable : 95% - 100%, plus probable qu'improbable : >50% - 100%, extrêmement improbable : 0% - 5%) sont également utilisés dans le rapport.

Tableau 3 - Définition des termes utilisés dans le cinquième rapport du GIEC relatifs à l'occurrence d'un événement (source : GIEC, Technical Summary, 2013)

D'ici la fin du XXI^{ème} siècle, le premier groupe de travail fait notamment part de journées et de nuits froides moins froides et/ou moins nombreuses (jugé comme *quasiment certain*), de journées et nuits chaudes plus chaudes et/ou plus fréquentes (*quasiment certain*) et d'une amplitude accrue d'une très haute élévation du niveau des mers (*très probable*). Notons que la contribution de l'Homme à ces 3 changements est jugée *probable* par les experts du premier groupe de travail.

Bien que le réchauffement de l'air ainsi que l'élévation du niveau de la mer d'ici la fin du XXI^{ème} siècle soient considérés comme au moins *très probable*, les conclusions portant sur les événements extrêmes, et notamment les cyclones, sont moins tranchées. En effet, s'il est *probable* qu'au cours du XXI^{ème} siècle la fréquence des cyclones tropicaux va diminuer ou

rester stable pour l'essentiel au niveau mondial et que la vitesse maximale des vents ainsi que l'intensité des précipitations imputées aux cyclones va augmenter, il est également *probable* que l'influence des changements climatiques sur les cyclones tropicaux varie selon les régions. Cependant, un *faible degré de confiance* est accordé aux projections par région.

Les projections concernant les cyclones de l'Atlantique Nord, ceux passant dans le bassin des Antilles, font apparaître une légère diminution de la fréquence de ces derniers (de l'ordre de -10%) accompagnée d'une forte augmentation de la fréquence des cyclones les plus puissants (de l'ordre +50%). Il est très important de noter que les intervalles de confiance donnés sont très larges : $[-60\%, +60\%]$ pour l'augmentation de fréquence générale, $[-100\%, +200\%]$ pour l'augmentation de fréquence des cyclones les plus puissants. Les projections font aussi apparaître une légère augmentation des vitesses maximales de vent par événement et des précipitations relatives aux cyclones : l'estimation est de +5% pour l'évolution de la vitesse de vent maximale par cyclone (avec un intervalle de confiance de $[-5\%, +10\%]$) et de +15% pour l'évolution de l'intensité des précipitations qui leur sont associées (avec un intervalle de confiance de $[+5\%, +20\%]$). En ce qui concerne l'évolution de la trajectoire des cyclones, il est *improbable*, selon les experts du premier groupe de travail, qu'un simple déplacement vers le pôle Nord des trajectoires se produise au cours du XXI^{ème} siècle. Par ailleurs, deux articles plus récents publiés dans la revue scientifique *Nature* (Baldini *et al.*, Kossin *et al.*) font part d'un déplacement progressif vers le Nord-Est du maximum d'intensité des ouragans au cours des dernières décennies et laissent penser que ce déplacement se poursuivra.

Les experts du premier groupe de travail concluent que seul un *faible degré de confiance* doit être accordé aux projections à long terme sur l'activité des cyclones tropicaux et précisent que ceci est lié à la faible profondeur de l'historique disponible. Un *faible degré de confiance* est également accordé à l'influence des activités humaines sur l'évolution de l'activité cyclonique et ceci provient de la méconnaissance actuelle des corrélations entre phénomènes physico-chimiques relatifs à l'activité humaine et les cyclones tropicaux.

Ainsi, bien que la première étude sur le réchauffement actuel du climat date de la fin du XXI^{ème} siècle, les conséquences de ce réchauffement sur le climat futur demeurent encore, en de nombreux points, incertaines. Dans la lignée de ce cinquième rapport, la Caisse Centrale de Réassurance a quant à elle étudié les conséquences assurantielles de l'évolution du climat.

2. Rapport de la Caisse Centrale de Réassurance

De par son rôle majeur dans l'indemnisation de la sinistralité liée aux catastrophes naturelles, la Caisse Centrale de Réassurance (CCR) s'intéresse de près au changement climatique et à son impact. Après avoir publié en 2018 une première étude à ce sujet portant sur la France métropolitaine, la CCR a rendu publique en février 2020 une seconde étude portant, elle, sur l'Outre-mer.

Intitulée *Évolution du risque cyclonique en Outre-mer à horizon 2050*, cette étude propose une estimation de l'évolution de la sinistralité liée à l'activité cyclonique au niveau marché. Pour ce faire, la CCR alimente ses modèles d'estimation de pertes avec des projections climatiques

provenant du modèle ARPEGE-Climat de Météo-France. Les périls modélisés sont les principaux périls consécutifs aux cyclones : le vent, la submersion marine et l'inondation. Précisons que les données issues du projet CORDEX, données que nous allons utiliser, sont en partie utilisées dans l'étude pour mettre en évidence une tendance à la hausse ou à la baisse de la fréquence des ouragans dans le bassin des Caraïbes.

L'étude livre dans un premier temps ses conclusions quant à l'évolution de la fréquence des événements frappant les côtes des territoires français d'Outre-mer. Selon l'étude, les données mettent en évidence une augmentation cette fréquence comprise entre +17% et +68% pour la Guadeloupe contre -35% et +51% pour la Martinique. La CCR explique la largeur de ces intervalles par des « résultats très divergents » selon le modèle climatique utilisé dans les données issues du projet CORDEX. De plus, la Guadeloupe et la Martinique étant deux territoires géographiquement proches, des intervalles aussi différents interpellent : les deux îles sont séparées par 170 km alors que la taille d'un ouragan varie entre 500 et 1 000 km. La CCR souligne donc « l'extrême prudence » avec laquelle ces résultats doivent être considérés. Sur l'ensemble de l'Outre-mer, la sinistralité augmenterait de 20% à l'horizon 2050 du fait de l'augmentation de la fréquence moyenne des cyclones et de la hausse du niveau de la mer.

Si le péril vent ne provoquerait pas davantage de dégâts à l'avenir pour les ouragans les plus intenses, l'étude pointe le rôle croissant que va jouer le péril submersion dans la sinistralité globale : augmentation de 14% des pertes dans le cas d'une submersion si le niveau de la mer augmente de 10 cm, de 34% si le niveau de la mer augmente de 20 cm. L'étude rapporte enfin que les inondations représentent un coût « bien inférieur » à celui lié au péril vent mais que pour les ouragans les plus intenses, le coût peut atteindre plusieurs centaines de millions d'euros. La CCR rappelle en effet que, selon le cinquième rapport du GIEC, les précipitations associées aux cyclones vont augmenter du fait du changement climatique.

Étant particulièrement présent sur le marché de l'assurance dommages dans les îles de Guadeloupe et de Martinique, Groupama se trouve exposé aux cyclones et sera donc impacté par le réchauffement climatique si l'on en croit l'étude de la CCR. Voyons comment est exposé le Groupe dans le bassin des Caraïbes.

3. Présentation de GMA et de la caisse Antilles-Guyane

Né au XIXème siècle de la volonté d'agriculteurs souhaitant protéger leurs terres contre différents aléas, Groupama Assurances Mutuelles (GMA) est aujourd'hui un des leaders de l'assurance en France. Fortement marqué par ses origines, le Groupe est assureur de la moitié des collectivités locales en France ainsi que leader sur le marché de l'assurance agricole et de la santé individuelle.

Comme la plupart des assureurs de premier plan du marché, GMA agit en tant que réassureur de ses entités. En France, GMA réassure donc 9 caisses régionales, 2 caisses d'Outre-mer ainsi que 2 caisses spécialisées fonctionnant de façon indépendante :

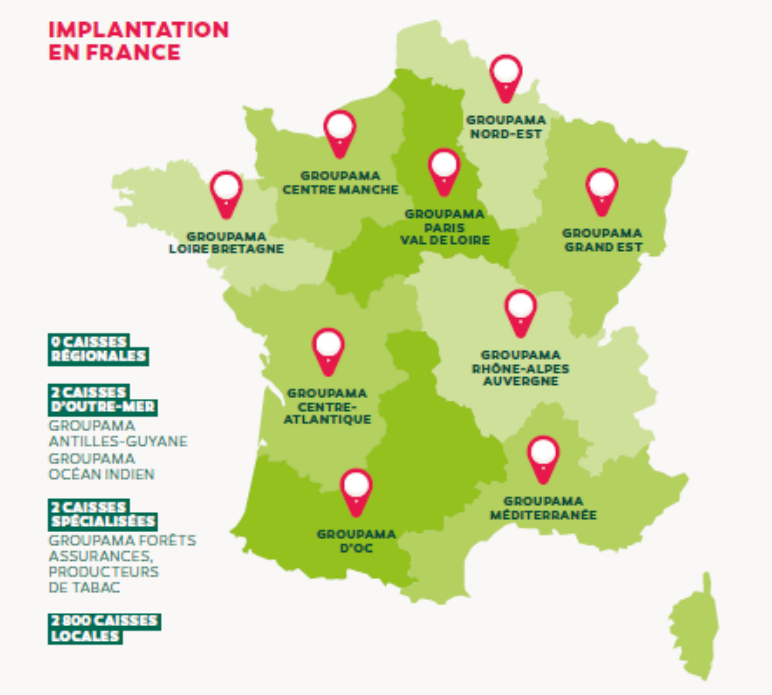


Figure 3 - Entités françaises de GMA (source : GMA)

La caisse régionale Antilles-Guyane existe depuis 30 ans et fait partie des 2 caisses régionales d'Outre-mer du Groupe. Elle couvre la Guyane ainsi que quatre îles situées dans la mer des Antilles : la Guadeloupe, la Martinique, Saint-Martin et Saint-Barthélemy.

Figure 4 – vue de l'Amérique Centrale avec indication des cinq territoires couverts par la caisse Antilles-Guyane (indiqués par un marqueur gris)

La caisse compte environ 66 000 sociétaires : des particuliers, professionnels, entreprises, commerces et collectivités. Les contrats couvrant les dommages aux biens climatiques sont répartis en différents segments : résidentiel, commerces et collectivités principalement. L'engagement total de la caisse au 31 décembre 2019 portant sur ces contrats (coûts de reconstruction, de remplacement des contenus, d'indemnisation de la perte d'exploitation pour les commerces et entreprises, …) est de 20 milliards d'euros. Dans la suite de ce mémoire, les « expositions » de la caisse Antilles-Guyane feront référence aux biens couverts par les contrats de dommages aux biens climatiques.

Groupama ainsi que deux autres assureurs français, Generali France et Allianz France, se partagent la majorité des contrats dans la zone Antilles. Le risque cyclonique y est particulièrement élevé et engendre une sous-représentation des assureurs. Cependant, créé en 1982, le régime d'indemnisation CAT-NAT permet d'accroître l'attractivité des Antilles en proposant aux assureurs, en cas de catastrophe naturelle, une couverture de réassurance illimitée, garantie par l'État.

4. Régime d'indemnisation des catastrophes naturelles

Le régime d'indemnisation des catastrophes naturelles, ou régime CAT-NAT, a été mis en place par l'État français dans le but de pallier une carence de couverture des risques naturels qui n'étaient que très peu assurés jusqu'alors.

Ce régime repose sur un partenariat public-privé : la Caisse Centrale de Réassurance, dont l'État est actionnaire à 100%, propose aux assureurs de les réassurer moyennant une surprime sur les contrats d'assurance dommages. Le taux de cette surprime est fixé par l'État.

Précisons que l'intervention de la CCR est conditionnée à la reconnaissance de l'état de catastrophe naturelle par l'État. Le *Code des Assurances* définit les catastrophes naturelles comme « des 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 pratique, la reconnaissance de l'état de catastrophe naturelle est effectuée commune par commune : le maire émet une demande de reconnaissance et doit prouver le caractère « anormal » de l'intensité du péril ayant touché sa commune. Dans le cas des ouragans, la reconnaissance est systématique si les vents engendrés dépassent 145 km/h en moyenne sur 10 minutes ou 215 km/h en rafales.

La couverture mise en place par la CCR est composée d'un premier traité Quote-Part à 50% et d'un second traité en excédent de pertes annuelles (ou *stop-loss*), sur la rétention du premier, à portée illimitée. Les territoires ultra-marins étant particulièrement exposés au risque cyclonique, le niveau de couverture élevé proposé par le régime CAT-NAT permet aux assureurs de limiter leur exposition et d'offrir des solutions d'assurance dans ces territoires.

Bien que ce régime s'avère très important pour les assureurs, le taux de pénétration de l'activité d'assurance dans les Antilles reste faible : il est de l'ordre de 50%. Il résulte du faible niveau de vie des populations dans la zone et des primes d'assurance plus élevées qu'en métropole : beaucoup de bâtiments sont en effet peu robustes et construits en zone inondable.

Le régime CAT-NAT permet donc d'accroitre l'attractivité des Antilles auprès d'assureurs souhaitant s'y implanter en limitant leur prise de risque. L'activité cyclonique étant particulièrement forte dans cette zone, les assureurs doivent par ailleurs se doter de leur propre vision du risque cyclonique et suivre son évolution sur le long terme. Grâce au rapport de la CCR, nous avons pu constater que cette évolution allait très probablement être impactée par le réchauffement climatique. Avant de voir comment le contexte de données ouvertes actuel permet d'étudier l'impact de ce réchauffement sur l'activité (et donc le risque) cyclonique,

essayons de comprendre comment se forment les cyclones et définissons les termes météorologiques d'usage.

II. Éléments de météorologie

1. Dépressions, tempêtes et ouragans en zone tropicale

Les trois phénomènes que nous allons définir maintenant sont des stades de développement plus ou moins avancés d'une perturbation atmosphérique. La notion de perturbation atmosphérique est très générale : Météo-France la définit comme toute interruption d'un état d'équilibre d'une région de l'atmosphère, de quelque nature que soit cette interruption. Les définitions de dépression, tempête ou ouragan sont tout aussi générales mais la création d'échelles (comme l'échelle de Saffir-Simpson que nous allons définir par la suite) a permis de de caractériser plus précisément ces phénomènes.

Notons que la zone tropicale fait référence à la zone située entre les tropiques du Cancer et du Capricorne :

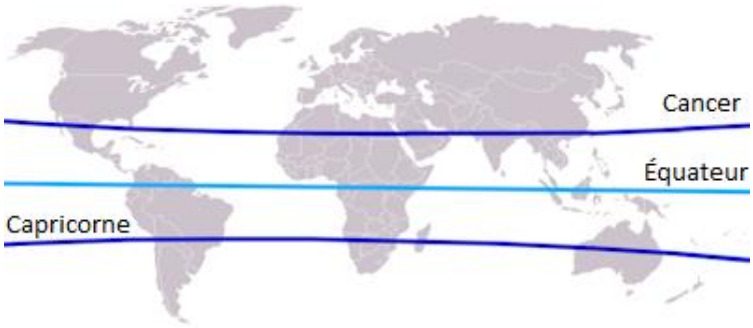


Figure 5 - Tropiques du Cancer, du Capricorne et équateur

Un phénomène est dit tropical s'il se forme dans la zone tropicale.

Une dépression tropicale est une perturbation atmosphérique tropicale caractérisée par :

- une zone de faible pression : à une altitude donnée, la pression est plus faible que dans les zones avoisinantes (d'où le terme de dépression)
- des vents tournant dans le sens anti-horaire dans l'hémisphère Nord, horaire dans l'hémisphère Sud
- un maximum des vitesses de vent moyennes sur une minute à 10 mètres du sol inférieur à 17 m/s

Une tempête tropicale est une dépression tropicale suffisamment puissante pour engendrer des vents moyens sur une minute dépassant les 17 m/s. A partir de 33 m/s, la tempête tropicale passe au stade d'ouragan.

Les ouragans sont des perturbations tropicales particulièrement violentes. Ils s'étendent sur 500 à 1000 km et sont caractérisés par une pression au centre très basse. Ce centre, appelé « œil », s'étend sur 30 à 60 km et est une zone d'accalmie : très peu de vent et pas de pluie.

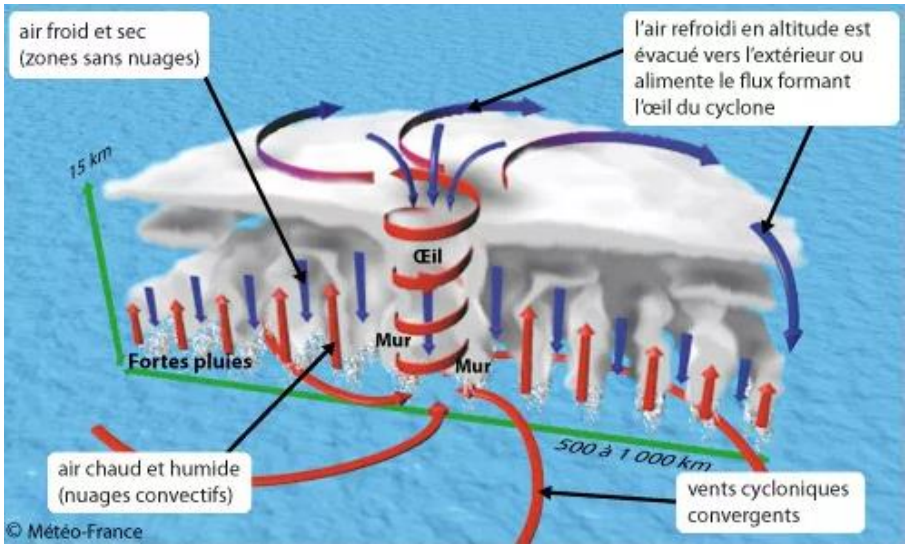


Figure 6 - Coupe verticale d'un ouragan (source : Météo-France)

Nous pouvons remarquer sur la coupe ci-dessus la présence d'un « mur » autour de l'œil : le mur engendre les vents les plus dévastateurs. Ceci aura son importance par la suite pour l'exploitation des données. A la base de l'ouragan, on peut noter la présence d'un courant ascendant d'air chaud et humide : le phénomène d'évaporation est en effet la principale source d'énergie d'un ouragan. Trois conséquences sont alors à relever :

- Les ouragans ne se forment qu'au-dessus d'une eau suffisamment chaude : au moins 28°C en surface, le processus d'évaporation est alors suffisamment intense. Ces eaux chaudes se trouvant principalement dans la zone tropicale, tous les ouragans (sauf cas exceptionnels) se forment dans cette zone.
- La mer dans la zone tropicale étant particulièrement chaude entre début juin et fin novembre, la majorité des ouragans se forment durant cette période.
- Un ouragan perd petit à petit en puissance si l'évaporation n'a plus lieu, comme par exemple dans le cas où il pénètre dans les terres.

Voici les principaux bassins d'ouragans :

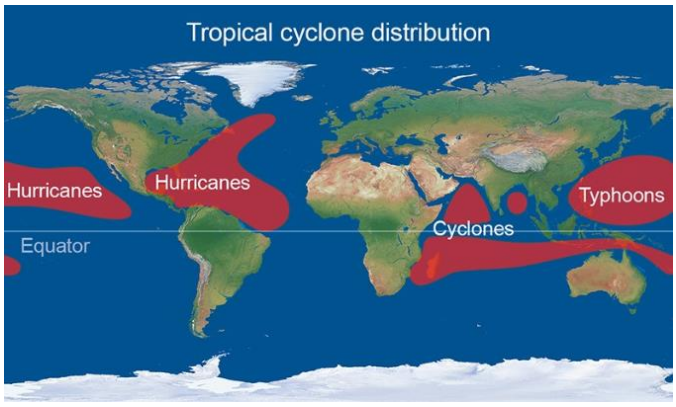


Figure 7 - Les différents bassins d'ouragans (source : Met Office)

Comme le souligne la carte ci-dessus, les termes ouragan, typhon ou cyclone sont propres à chaque bassin de formation mais désignent un seul et même phénomène. Nous pouvons remarquer que ces bassins sont bien répartis de part et d'autre de l'équateur, dans la zone tropicale.

La carte ci-dessus expose un autre élément ayant son importance dans la formation des ouragans : la force de Coriolis. Cette force amorce la rotation des vents autour de l'œil et contribue au maintien de l'ouragan. La force de Coriolis étant quasi nulle au voisinage de l'équateur, aucun ouragan ne se forme dans cette zone. Pour la même raison, un ouragan ne peut traverser l'équateur au cours de son cycle de vie.

Un ouragan est donc une perturbation tropicale puissante se formant grâce à une perturbation initiale, à l'évaporation et à la force de Coriolis. Voyons maintenant comment caractériser la puissance d'un tel phénomène.

2. Caractériser la puissance d'un ouragan

Nous l'avons vu, il existe plusieurs bassins d'ouragans tout autour du globe. Chacun de ces bassins est rattaché à un Centre Météorologique Régional Spécialisé (CMRS) qui a pour but, entre autres, d'étudier les ouragans se formant dans la zone. Le National Hurricane Center (NHC) est le centre de référence pour le bassin des Caraïbes et c'est ce centre qui a mis en place l'échelle de Saffir-Simpson en 1969.

Cette échelle permet la classification d'un système dépressionnaire en fonction de la vitesse des vents qu'il engendre :

	Vitesse de vent maximale enregistrée		Description des dommages
	m/s	km/h	
Ouragan catégorie 5	≥ 70	≥ 252	Évacuation en urgence des zones résidentielles en amont de l'événement, bâtiments de tous types sévèrement endommagés, manque d'eau et d'électricité pouvant durer plusieurs mois
Ouragan catégorie 4	59-70	209-251	Évacuation en d'urgence de la côte nécessaire ainsi que des maisons de plain-pied sur une distance de 400m à partir de la côte en amont de l'événement, dégâts irréparables sur les petites résidences, érosion importante sur les plages, coupures importantes d'eau et d'électricité possibles
Ouragan catégorie 3	50-58	178-208	Dégâts structurels sur les petites résidences, habitations aux fondations non solides détruites et toits arrachés, les inondations peuvent détruire des constructions légères sur les côtes, risque de glissement de terrain, coupures totales d'électricité et manque d'eau possibles
Ouragan catégorie 2	43-49	154-177	Dégâts structurels possibles, dommages importants infligés à la végétation, déracinement d'arbres, coupures de courant et manque d'eau potable pendant plusieurs jours possibles
Ouragan catégorie 1	33-42	119-153	Aucun dégât structurel significatif causé aux bâtiments mais des tuiles peuvent s'envoler et des coupures de courant peuvent avoir lieu
Tempête tropicale	18-32	63-118	Beaucoup de pluie, dégâts négligeables
Dépression tropicale	≤ 17	≤ 62	Beaucoup de pluie

Tableau 4 - Échelle de Saffir-Simpson (source : NOAA)

Notons que par *vitesse enregistrée,* on entend *vitesse moyenne sur une minute à 10 mètres du sol sur un terrain sans obstacle*. La vitesse maximale enregistrée est utilisée pour classifier le système dépressionnaire. Ainsi, un système dépressionnaire engendrant des vents atteignant au maximum 45 m/s est classifié comme Ouragan de catégorie 2.

Cette échelle de mesure d'intensité d'un système dépressionnaire n'est pas unique : chaque CMRS a, au fil du temps, développé sa propre échelle. Néanmoins, l'échelle de Saffir-Simpson est la plus populaire de toutes. Notons par ailleurs qu'il existe des méthodes de caractérisation de l'intensité d'un ouragan plus complexes, plus récentes et utilisant davantage de données. Par exemple, l'indice HSI (pour *Hurricane Severity Index*), mis en place en 2005 par Hebert *et al.*, exploite les vitesses de vent générées par l'ouragan et leur répartition spatiale. L'ouragan est alors décrit plus précisément par cet indice que par une échelle n'exploitant que les vitesses maximales, comme l'échelle de Saffir-Simpson. Nous verrons cependant qu'un tel indice ne pourra être calculé sur nos données.

Nous pouvons voir ici que les dépressions tropicales ainsi que les tempêtes tropicales ne causent que des dégâts mineurs. Pour cette raison, nous nous focaliserons uniquement sur les ouragans dans la suite de cette étude.

Ayant maintenant assez d'information sur le phénomène d'ouragan en lui-même et sur une façon de caractériser sa puissance, intéressons-nous aux projections de l'activité cyclonique et aux scénarii dans lesquels elles sont réalisées.

3. Réaliser des projections

Afin de réaliser des projections jusqu'en 2100, il est nécessaire de prendre en compte différentes évolutions possibles de l'environnement socio-économique terrestre. Depuis le 5ème rapport du GIEC publié en 2014, ces différentes évolutions sont formalisées par quatre scénarii dits *RCP*, *Representative Concentration Pathway*. Les simulations du climat futur seront donc à chaque fois faites dans le cadre d'un de ces scénarii.

Chacun des quatre scénarii est un profil d'évolution des concentrations en gaz à effet de serre jusqu'en 2100. C'est par l'intermédiaire de la notion de forçage radiatif que sont mesurées de manière quantitative ces évolutions.

Le forçage radiatif est une notion physique : c'est la différence entre l'énergie reçue et produite par un même système. Si ce système est la planète Terre, il existe un équilibre entre le rayonnement solaire entrant et les rayonnements sortant de l'atmosphère : cet équilibre contrôle la température terrestre. Ainsi, le forçage radiatif évalue l'éloignement par rapport à cette situation d'équilibre : plus il est élevé, plus il y a déséquilibre et donc réchauffement de la Terre et de son atmosphère. Précisons que le forçage radiatif s'exprime en W/m².

Les quatre scénarii sont les suivants :

- RCP 2.6
- RCP 4.5
- RCP 6
- RCP 8.5

Les noms de ces scénarii viennent du forçage radiatif attendu en 2100 : par exemple, en se plaçant dans le scénario RCP 4.5, on s'attend à ce que le forçage radiatif atteigne environ 4.5 W/m² en 2100.

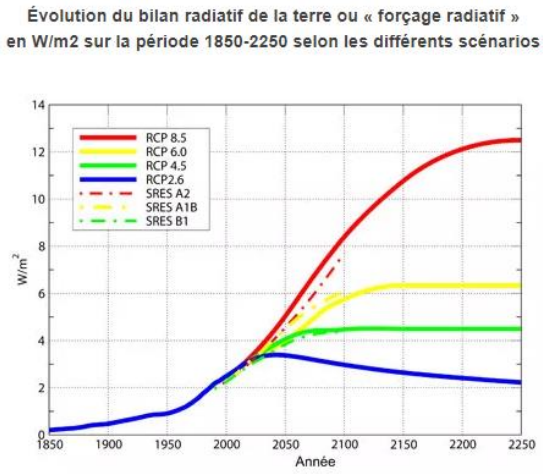


Figure 8 - Évolution du forçage radiatif selon les 4 scénarii RCP (source : Météo-France)

Comme nous pouvons le voir sur le graphique ci-dessus, le scénario RCP 2.6 correspond à un scénario optimiste où les politiques de réduction des émissions de gaz à effet de serre permettraient une baisse du forçage radiatif avant 2050 et donc de limiter fortement le

réchauffement planétaire. Le scénario RCP 8.5 est quant à lui le scénario le plus pessimiste : il correspond au scénario où les émissions de gaz à effet de serre restent dans leur état actuel. C'est dans le cadre de ce scénario que les phénomènes engendrés par le réchauffement climatique comme la montée du niveau des mers ou l'augmentation des températures à la surface du Globe sont les plus intenses.

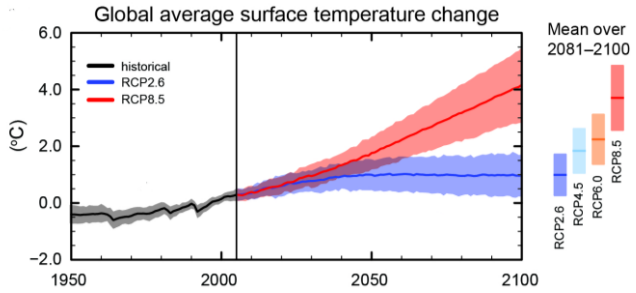


Figure 9 – Évolution de la moyenne globale des températures à la surface de la Terre selon les différents scénarii RCP par rapport à la température moyenne sur la période 1986-2005 (source : GIEC, Climate change 2013 : The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the intergovernmental panel on climate change, 2013)

Nous pouvons voir sur la figure ci-dessus que l'augmentation de la température de surface au niveau mondial à l'horizon 2100 augmente avec le forçage radiatif attendue à cette période : 1°C d'augmentation pour le scénario RCP 2.6, jusqu'à près de 4°C d'augmentation pour le scénario RCP 8.5. A titre de de comparaison, selon une récente étude publiée dans la revue *Nature* (Tierney *et al.*), la température moyenne de la Terre au plus fort de la dernière glaciation était 6,1°C plus basse que la température moyenne actuelle : une augmentation de l'ordre de 4°C à l'horizon 2100 serait alors lourde de conséquences.

L'augmentation du niveau des mers varie également en selon le forçage radiatif :

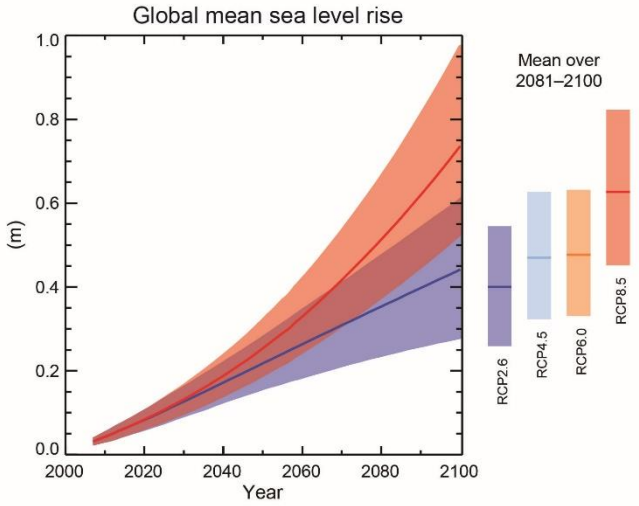


Figure 10 - Évolution du niveau moyen des mers selon les différents scénarii RCP par rapport au niveau observé en 2005 (source : GIEC, Climate change 2013 : The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the intergovernmental panel on climate change, 2013)

Si l'augmentation du niveau global des mers serait de 40 cm environ à l'horizon 2100 dans le cadre du scénario RCP 2.6, elle serait de plus de 60 cm dans le cadre du scénario le plus pessimiste.

Nous savons maintenant comment se forment les ouragans et comment il est possible de caractériser leur puissance. Une des caractéristiques importantes des projections climatiques actuelles liées à l'étude du réchauffement climatique, les scénarii RCP, ont été présentés et vont nous permettre de comprendre les données utilisées dans le cadre de ce mémoire.

III. Projections climatiques utilisées et extraction des ouragans

Nous allons voir dans cette partie quelles sont les projections climatiques que nous avons utilisées et comment les ouragans ont été extraits de ces projections puis étudiés.

1. Projet CORDEX

Le projet international CORDEX a été lancé en 2009 par le Programme Mondial de Recherche sur le Climat (WCRP). L'objectif de ce projet est double : d'une part donner un cadre de référence pour la production de projections climatiques régionales, d'autre part fournir une interface permettant à toute personne d'avoir accès à ces simulations *via* internet. Pour comprendre tout l'intérêt du projet CORDEX, attardons-nous sur ces deux objectifs.

Afin d'étudier l'évolution du climat, la communauté météorologique réalise des projections climatiques. Ces projections sont réalisées à l'aide de modèles de circulation générale (*Global Circulation Model* ou *GCM* en anglais). Ces derniers sont des modèles mathématiques simulant le comportement de l'atmosphère et des océans à l'échelle du globe : ils reposent essentiellement sur des équations issues de la mécanique des fluides et de la thermodynamique. De par le nombre et la complexité des interactions qu'ils modélisent, les GCMs sont très coûteux à utiliser en termes de ressources informatiques.

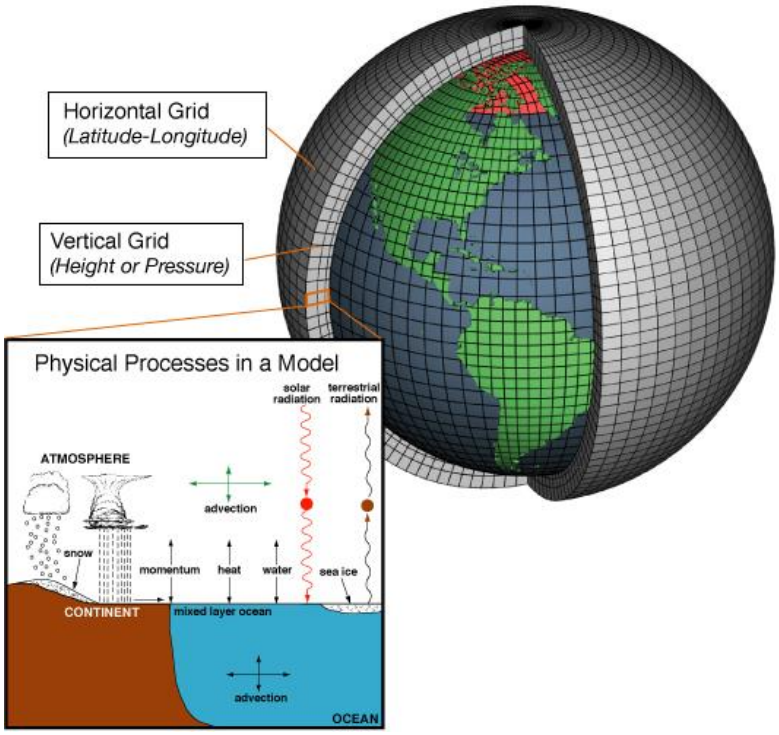


Figure 11 - Exemples de processus physico-chimiques modélisés par un GCM (source : NOAA)

Comme nous pouvons le voir sur le schéma ci-dessus, un GCM génère des résultats sur une grille tridimensionnelle : deux dimensions permettent de se repérer sur le globe terrestre (latitude et longitude), la troisième dimension est, elle, dédiée à l'altitude. Un détail a son importance : le maillage de cette grille est très large. Bien qu'il soit de plus en plus fin grâce à des ordinateurs toujours plus puissants, la maille de la grille bidimensionnelle recouvrant le Globe, c'est-à-dire l'espacement entre deux points successifs de cette grille, varie généralement entre 100 et 500 km. Un tel maillage ne permet alors pas de fournir de résultats locaux, valables pour un lieu précis comme une ville par exemple.

Bien conscient de l'intérêt que pourraient avoir des projections climatiques plus ciblées pour des décideurs locaux ou des entreprises privées, le WCRP a mis en place le projet CORDEX : ***CO****ordinated* ***R****egional* ***D****ownscaling* ***EX****periment*. Ce projet coordonne donc la production mondiale de projections climatiques locales en fixant un cadre. Ce dernier permet au projet de rendre accessible des données homogènes, facilement comparables entre elles. Pour cette raison, le GIEC utilisera ces projections climatiques régionales dans son 6ème rapport, publié en 2022.

Voyons donc quelles sont les caractéristiques du cadre mis en place.

2. Cadre associé au projet et données utilisées

Les projections climatiques accessibles en ligne grâce au projet CORDEX sont produites par différents centres de recherche en météorologie : citons l'Institut Pierre-Simon Laplace (IPSL) à Paris par exemple. Ces projections doivent être faites dans un cadre bien précis et c'est ce cadre que nous allons détailler ici.

➢ Les domaines

Les projections climatiques sont disponibles pour 14 zones du globe : ce sont les domaines. On retrouve parmi eux l'Asie du Sud, l'Amérique du Nord, l'Afrique ou encore l'Europe.

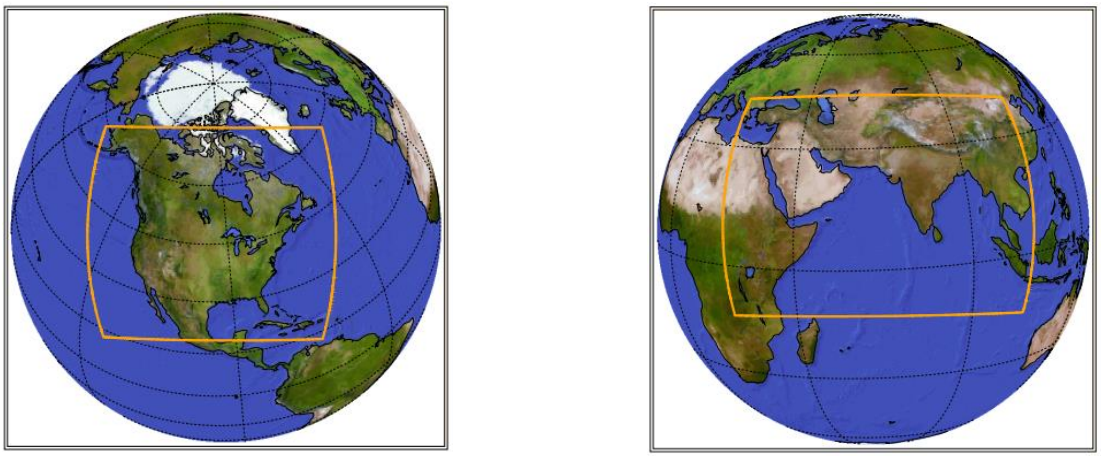


Figure 12 - Délimitation des domaines Amérique du Nord (gauche) et Asie du Sud (droite) (source : CORDEX)

Le domaine utilisé pour ce mémoire est le domaine couvrant le bassin des Caraïbes, c'est-à-dire *Central America* (ou *CAM*).

Figure 13 - Délimitation du domaine CAM (source : projet CORDEX)

➢ Les variables

Ces projections climatiques sont disponibles sous la forme de variables projetées jusqu'en 2100 : on retrouve parmi elles la pression à une altitude donnée, la vitesse des vents à la surface du sol ou encore la pluviométrie.

La variable utilisée dans ce mémoire est la vitesse de vent au niveau de la surface du sol sur une zone dégagée. Notons que cette variable est une vitesse instantanée : ce n'est pas une vitesse maximale ou une vitesse moyenne sur une période donnée. D'autres variables caractérisant la puissance d'un ouragan auraient pu être utilisées également comme la pression à une altitude donnée. Cependant, en plus d'augmenter considérablement le volume de données, de telles variables n'apportent pas davantage d'information.

Bien que le vent soit le péril causant la majorité des dégâts pour un ouragan, nous avons vu précédemment que le niveau de la mer ainsi que les précipitations jouent un rôle important. Des variables telles que la pluviométrie auraient alors pu être prises en compte dans l'étude. Cependant, aussi bien dans la base de données historique que dans les bases d'ouragans simulés par des modélisateurs, de telles variables ne sont généralement pas disponibles : seules la pression et la vitesse des vents le sont systématiquement.

- Le pas de temps

Les variables sont projetées jusqu'en 2100 suivant un pas de temps fixe. Les différents pas de temps disponibles sont 1 heure, 3 heures, 6 heures, un jour, une semaine, un mois, une année. Dans notre cas, le pas de temps choisi est le plus court disponible pour la zone CAM, à savoir 3 heures. Ce pas de temps est satisfaisant et suffisamment fin pour identifier et suivre des ouragans dans les données.

- La maille

Comme dit précédemment, la raison d'être du projet CORDEX est de fournir des projections climatiques sur une grille de maille fine : cette maille varie entre 25 km et 12.5 km et recouvre le domaine choisi. Afin d'alléger le poids des données, les domaines les plus grands ont tendance à être associés avec la maille la plus large (25 km) : c'est le cas du domaine CAM et c'est donc cette maille de 25 km que nous utiliserons.

Cette maille est peu satisfaisante pour deux raisons. La première concerne la taille des territoires couverts par la caisse Antilles-Guyane : les quatre îles de de la mer des Antilles sont de taille très réduite et seul une très faible quantité de données décrira alors la vitesse des vents sur ces territoires. Seulement 24 km séparent la côte Est de la côte Ouest en Martinique par exemple, 45 km en Guadeloupe, alors que ces deux îles sont les plus grandes des quatre. Dans le cadre de cette étude, nous verrons cependant que ce problème n'entravera pas notre avancée.

La seconde raison concerne les ouragans présents dans les données issues du projet CORDEX. Le cinquième rapport du GIEC ainsi que de nombreuses publications scientifiques (Fuentes-Franco *et al.* ou Vishnu *et al.* par exemple) rappellent que les mailles des modèles actuels ne permettent pas d'étudier précisément des phénomènes aussi locaux que les ouragans : les mailles les plus fines sont encore trop larges. Pour cette raison, des méthodes précises de caractérisation de la puissance d'un ouragan, comme l'indice HSI évoqué plus haut, sont inapplicables sur nos données. Cependant, bien que d'un point de vue purement météorologique la maille ne soit pas optimale, nous verrons qu'il est tout à fait possible, dans le cadre de cette étude, de tirer de l'information de nos données.

- Le GCM et le RCM

Le GCM est le modèle climatique générant des projections sur une maille très large. Ce modèle comprend une part d'aléatoire : deux *runs* aux conditions initiales identiques donnent des projections différentes.

Afin d'obtenir les projections climatiques régionales du projet CORDEX, un modèle climatique régional (ou RCM) est appliqué sur les projections du GCM : ceci permet l'obtention de projections sur une maille plus fine. C'est donc le RCM qui réalise l'étape de descente d'échelle et permet d'aboutir aux données issues du projet CORDEX.

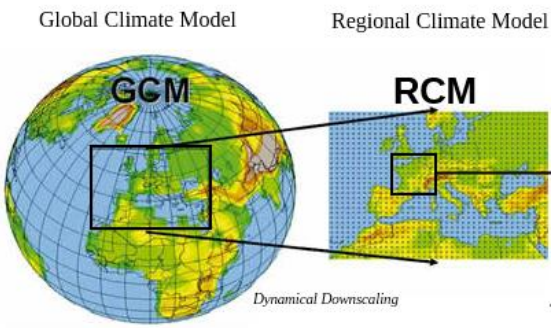


Figure 14 - Passage d'une modélisation climatique globale à régionale (source : projet DRIAS)

Notons qu'un RCM ne comporte pas d'aléatoire : appliquer un même RCM deux fois sur un même *run* de GCM donne les mêmes projections.

CORDEX rassemble les projections climatiques issues de 3 GCMs de référence :

- NorESM, Norwegian Earth System Model (faiblement sensible au changement climatique).
- MPI-ESM, Max Planck Institut Earth System Model (moyennement sensible au changement climatique).
- HadGEM2-ES, Hadley Centre Global Environment Model version 2 (fortement sensible au changement climatique).

Il est important de noter qu'ils couvrent l'ensemble du spectre des sensibilités au changement climatique. Bien que ces 3 GCMs produisent la majorité des données, des projections résultant d'autres GCMs sont disponibles pour certains domaines.

Le choix du RCM est quant à lui plus limité puisque CORDEX rassemble les projections issues de 2 RCMs :

- REMO, développé par le Climate Service Center Germany (GERICS) à Hambourg.
- RegCM, développé par l'International Centre for Theoretical Physics (ICTP) à Trieste.

Seules les projections produites par les 3 GCMs de référence sont disponibles dans notre cas : MPI-ESM, NorESM et HadGEM2-ES. Ce choix limité est suffisant dans le cadre cette étude puisque, comme vu précédemment, ces 3 GCMs couvrent l'ensemble du spectre des sensibilités au changement climatique. Quant au RCM, le seul disponible est REMO. Il aurait pu être intéressant de comparer les projections issues de ce RCM avec celles issues de RegCM afin de mettre en évidence le poids du RCM dans les projections créées.

➢ Scénarii RCP

Réaliser des projections climatiques jusqu'en 2100 implique de devoir faire des hypothèses sur l'évolution de l'environnement socio-économique dans lequel sont faites ces projections. Ainsi, différents scénarii sont proposés : ce sont les scénarii RCP, détaillés précédemment. Les projections sont toujours faites dans le cadre de ces scénarii. Dans la pratique, ces derniers agissent sur les conditions initiales du GCM.

Afin de s'assurer de la qualité des projections climatiques aux conditions futures, des jeux de données dits « historiques » sont proposés : les conditions initiales du GCM sont celles du passé. Vérifier l'adéquation de ces simulations historiques avec le climat passé est alors une façon de tester la fiabilité des projections proposées.

Ainsi, en fixant un domaine, une variable, une maille, un pas de temps et un couple GCM-RCM, 5 jeux de données sont en théorie disponibles :

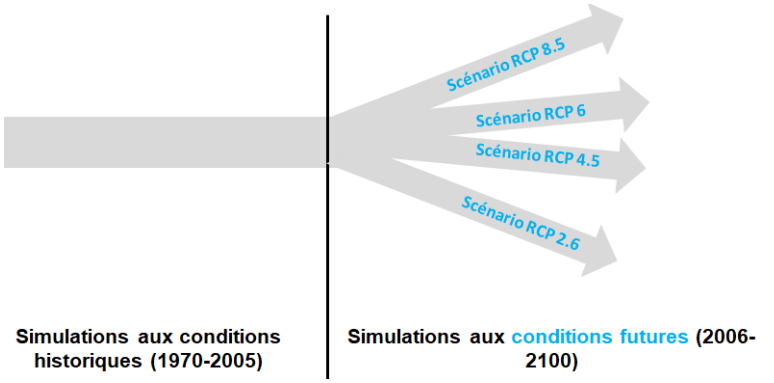


Figure 15 - Les cinq jeux de données disponibles pour les données issues du projet CORDEX

Dans notre cas, seules les projections faites dans le cadre des scénarii RCP 2.6 et RCP 8.5 sont disponibles. Rappelons que ce sont les deux scénarii extrêmes : RCP 2.6 est le plus optimiste quant aux émissions de gaz à effet de serre, RCP 8.5 le plus pessimiste.

La structure de nos données est alors la suivante :

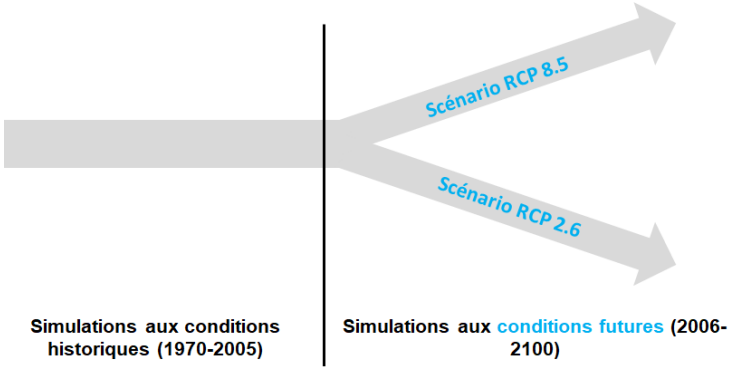


Figure 16 - Structure des scénarii utilisés pour simuler les données

Comme expliqué précédemment, seul le GCM comporte une part d'aléatoire et il faudrait alors plusieurs *runs* d'un même GCM pour étudier de façon optimale un phénomène précis.

Or, le projet CORDEX ne donne accès qu'à un seul et unique *run* : la période 2006-2100 n'est simulée qu'une seule fois pour chacun des scénarii. Interrogé sur le sujet, Guillaume Levavasseur, météorologue à l'IPSL, fait part de modèles RCMs « extrêmement coûteux » en termes de ressources informatiques. Le choix alors fait est de ne rendre qu'un seul *run* accessible : le *run* le plus en adéquation avec les observations historiques. Guillaume Levavasseur précise cependant qu'idéalement, seul une dizaine de *runs* est utilisée pour étudier un phénomène climatique précis. La part d'aléatoire n'est en effet pas aussi importante que lors du tirage d'un échantillon généré selon une loi de probabilité donnée par exemple.

Un internaute souhaitant accéder à ces projections doit faire son choix parmi les différentes caractéristiques du cadre que nous venons de présenter. Nous l'avons constaté, pour un domaine et une variable fixés, toutes les modalités des caractéristiques évoquées ci-dessus ne sont pas nécessairement disponibles. Ainsi, les pas de temps les plus courts ne sont pas toujours disponibles, de même pour la maille la plus fine ou certains scénarii RCP. Aussi, les données dites « disponibles » sont celles qui étaient disponibles au début de la réalisation de cette étude en décembre 2019. Davantage de données seront disponibles à l'avenir pour le domaine CAM.

Étant donné le pas de temps choisi, notre jeu de données est composé d'une matrice de ce type générée toutes les 3 heures :

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,417} \\ \vdots & \ddots & \vdots \\ a_{225,1} & \cdots & a_{225,417} \end{pmatrix}$$

Cette matrice possède 225 lignes et 417 colonnes. Chacun des coefficients de la matrice donne la vitesse instantanée des vents en m/s à 10 mètres de hauteur en un endroit du domaine CAM : par exemple, le coefficient $a_{1,1}$ donne la vitesse du vent à l'extrême Nord-Ouest du domaine. Deux coefficients successifs donnent la vitesse des vents en deux endroits différents et, étant donné le choix de la maille, sont espacés d'environ 25 km.

Ces vitesses sont instantanées et non moyennes sur une minute, comme les vitesses considérées dans l'échelle de Saffir-Simpson ou celles de la base de données historiques dont nous parlerons par la suite. Il est important d'en tenir compte mais notons que dans la littérature scientifique, il est commun de comparer directement ces deux types de vitesses, sans correction préalable. Citons par exemple les travaux de Fuentes-Franco *et al*. qui seront présentés dans la suite de ce mémoire.

Visualisons une de ces matrices composées de coefficients décrivant des vitesses de vent :

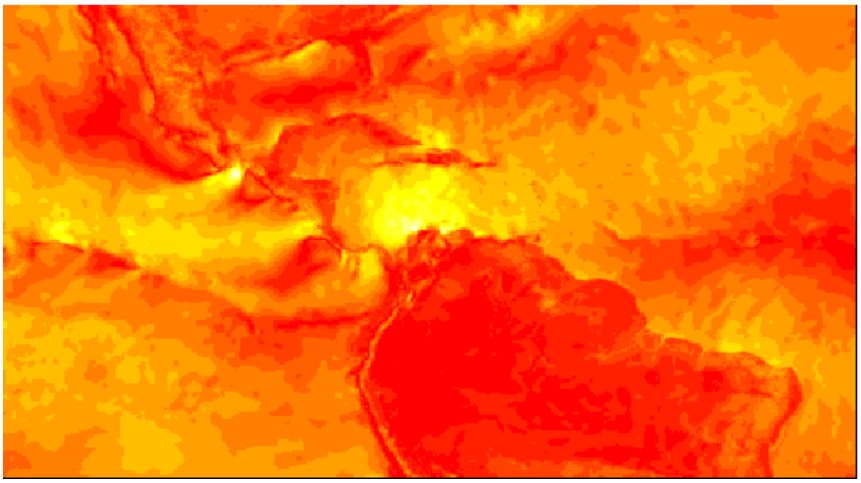


Figure 17 – Visualisation d'une matrice de vitesses de vent

Les zones de vents forts apparaissent en jaune, les zones de vents faibles en rouge. On distingue clairement le continent américain en rouge puisque les vents soufflent moins fort dans les terres.

A raison d'une matrice de ce type toutes les 3 heures, une année est représentée par 2920 matrices de vent. Nos données sont alors composées de :

- 105 120 matrices pour les 36 années de simulations aux conditions historiques
- 277 400 matrices pour les 95 années de simulations selon le scénario RCP 2.6
- 277 400 matrices pour les 95 années de simulations selon le scénario RCP 8.5

Au total, notre jeu de données est composé de 659 920 matrices de vent, ce qui représente un poids d'environ 800Go. Voyons maintenant comment mettre en évidence l'activité cyclonique présente dans ce jeu de données, c'est-à-dire extraire les ouragans apparaissant sur les matrices de vent, afin d'en déduire des pertes financières pour le Groupe.

3. Mise en place d'un système de *tracking*

Le but est d'extraire les ouragans présents dans notre jeu de données. Cette extraction prend la forme d'un problème de *tracking* : nous cherchons à suivre automatiquement un objet sur plusieurs images, correspondant ici aux visualisations des matrices de vitesses de vent. Différents algorithmes de *tracking* déjà construits ont dans un premier temps été testés mais les résultats ne furent pas satisfaisants. Les détails de ces essais sont donnés en annexe. Pour cette raison, un système de *tracking* a dû être créé dans le cadre de cette étude.

L'extraction prend la forme d'une construction d'empreinte pour chaque ouragan : l'empreinte est le tracé suivi par l'ouragan au cours de sa vie. Notons que la méthode d'extraction des données est la même pour tout le jeu de données : pas de différences entre l'exploitation des simulations selon les scénarii RCP et les simulations aux conditions historiques.

a) Troncature des matrices

Avant d'entamer l'extraction, réduisons la taille de toutes les matrices de notre jeu de données. Comme nous pouvons le voir en figure 17, un certain nombre de coefficients sur la gauche de l'image concernent l'océan Pacifique, or nous nous intéressons aux ouragans du bassin des Caraïbes. Enlever des colonnes sur la gauche de la matrice permet alors de réduire le poids de nos données sans perdre d'information.

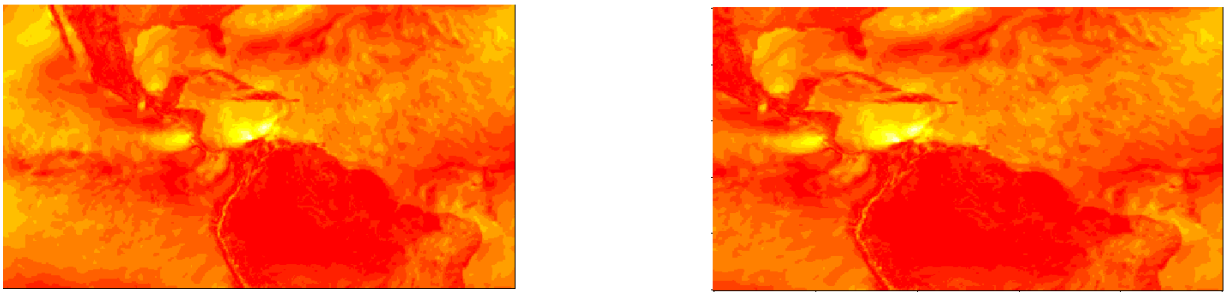


Figure 18 – Visualisation d'une matrice de vitesses de vent (gauche) et de la même matrice tronquée (droite)

b) Filtration des vents forts

Un ouragan tropical étant une perturbation tropicale dont les vents atteignent ou dépassent les 33 m/s, isoler les vitesses de vent dépassant ce seuil permet d'avoir un premier aperçu des ouragans présents dans nos données.

Filtrons alors les vitesses de vent supérieures à 33 m/s. Étant donnée une matrice initiale M de coefficients $m_{i,j}$, filtrer cette matrice au seuil s consiste à lui associer la matrice M' de coefficients $m'_{i,j}$ telle que :

$$\begin{cases} m'_{i,j} = m_{i,j} \text{ si } m_{i,j} \geq s \\ m'_{i,j} = 0 \text{ si } m_{i,j} < s \end{cases}$$

Ainsi, une matrice dont certaines vitesses de vent sont supérieures à ce seuil est transformée de cette manière :

Figure 19 – Visualisation d'une matrice de vitesses de vent (gauche) et de la même matrice filtrée au seuil de 33 m/s (droite)

La différence entre les vitesses de vent les plus fortes et les plus faibles étant trop grande, le

continent américain n'est plus visible sur l'image de gauche. La matrice filtrée apparaît, quant à elle, comme majoritairement rouge puisque la quasi-totalité des coefficients sont égaux à 0. Seuls quelques coefficients situés dans la partie haute de l'image apparaissent jaune car non nuls : ils représentent des vitesses de vent supérieures ou égales à 33 m/s, caractéristiques d'un ouragan. Nous pouvons même distinguer l'œil de ce dernier : les vitesses de vent au niveau de l'œil étant moins élevées, l'œil apparaît comme un point rouge au milieu du cercle jaune. Remarquons que les vitesses les plus élevées, celles supérieures à 33 m/s sont situés proche de l'œil : c'est le mur dont nous parlions en partie II.

Ainsi, il nous est possible d'identifier sur chaque année simulée les pics d'activité cyclonique, c'est-à-dire les matrices où des ouragans atteignent leur pic d'intensité. Appelons-les *matrices de vents forts*. Elles apparaissent en vert sur le schéma suivant :

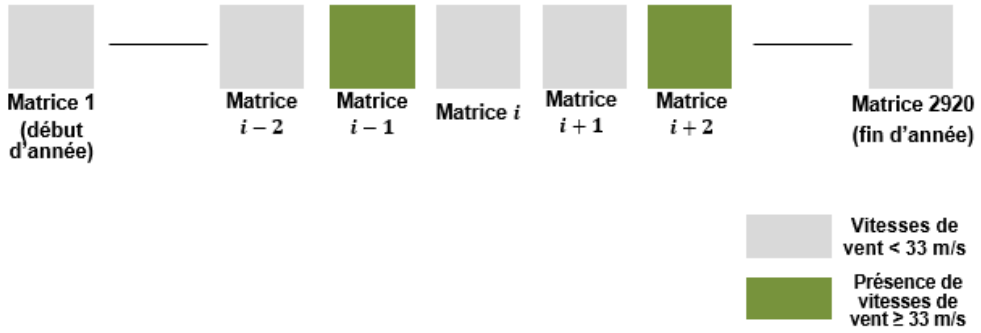


Figure 20 - Une année de vitesses de vent simulées

Précisons qu'un ouragan peut avoir plusieurs pics d'intensité, il n'en a pas nécessairement qu'un seul.

c) Étude des matrices de vents forts

L'étape de filtration des vents a permis d'identifier les matrices de vents forts. Nous n'avons cependant aucune information sur le nombre d'ouragans présents sur chacune de ces matrices et sur leur localisation.

Considérons une matrice de vents forts où se trouvent par exemple deux ouragans :

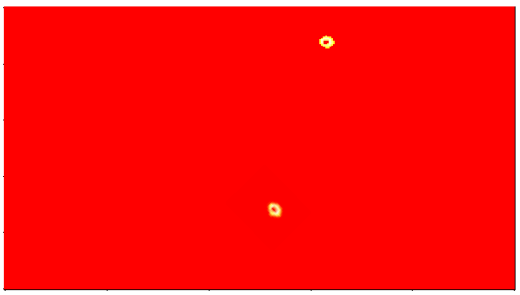


Figure 21 - Visualisation d'une matrice de vents forts filtrée à deux ouragans

Rappelons que sur cette matrice, tous les coefficients sont nuls sauf ceux étant supérieurs ou égaux à 33 m/s. Ces derniers sont affichés en jaune. Supposons qu'ils sont au nombre de N.

Nous pouvons alors assimiler cette matrice à un plan dans lequel se trouve N points : chaque point a pour coordonnées son indice de colonne et son indice de ligne dans la matrice initiale.

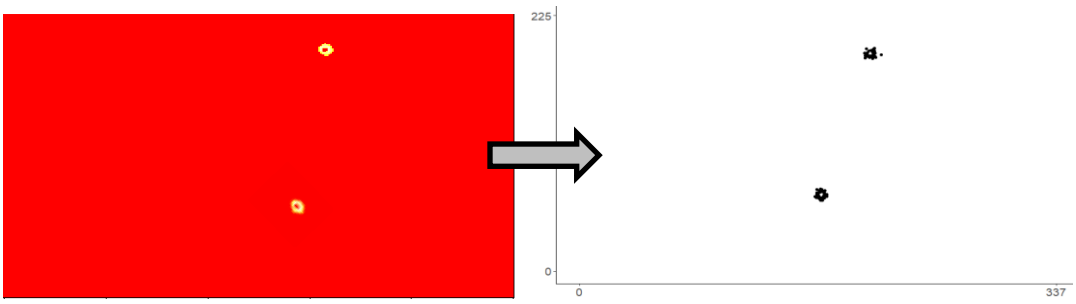


Figure 22 - Passage d'une matrice de vents forts filtrée dans le plan

Ne connaissant pas le nombre d'ouragans présents dans la matrice de vents forts, une étape de partitionnement de nos N points dans le plan s'avère nécessaire. A l'issue de ce partitionnement, nous saurons que les N points décrivent k ouragans (ci-dessus, $k = 2$).

Pour chacun de ces k ouragans, cherchons ensuite la position de l'œil. Nous aurons alors à notre disposition le nombre ainsi que la localisation des ouragans sur les matrices de vents forts.

Au lieu de procéder séparément pour ces deux étapes, nous allons utiliser une méthode de clustering combinant partitionnement et recherche des centres des classes : l'algorithme *Mean-Shift*.

d) Algorithme Mean-Shift

L'algorithme Mean-Shift est une méthode itérative permettant d'une part de partitionner un jeu de données composé d'individus décrits par des variables quantitatives mais également de trouver le centre des groupes formés. Précisons que le nombre de groupes à former est inconnu : il s'agit d'une méthode de classification non-supervisée. L'algorithme estime dans un premier temps une densité sur les données puis associe chaque observation au maximum local de l'estimateur de densité le plus proche.

- Estimation de densité

  Soient $x_1, x_2, \dots, x_n$ des réalisations de variables aléatoires $X_1, X_2, \dots, X_n$ à valeur dans $\mathbb{R}^d$ indépendantes et identiquement distribuées. La fonction de densité commune est inconnue et notée f. Dans notre cas, $d = 2$ et $n = N$, le nombre de coefficients non nuls de la matrice de vents forts filtrée.

  Une méthode classique d'estimation non-paramétrique de f est l'estimation par noyau. On définit l'estimateur par noyau de f, noté $\hat{f}_H$, par :

$$\hat{f}_H(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i)$$

Avec :

- x un point quelconque de $\mathbb{R}^d$
- H la matrice de lissage, symétrique de taille dxd et définie positive. Ici, H est choisie diagonale de coefficients diagonaux tous égaux.
- $x \mapsto K_H(x)$ un noyau symétrique, c'est-à-dire une fonction positive, d'intégrale égale à 1 et symétrique.

Le choix le plus classique pour le noyau est le noyau gaussien et K est alors la fonction de densité d'une loi normale multidimensionnelle centrée de matrice de variance H :

$$K_H(x) = (2\pi)^{-d/2}|H|^{-1/2} \exp\left(-\frac{1}{2}x^T H^{-1} x\right)$$

avec $|H|$ le déterminant de la matrice H. C'est ce noyau que nous utiliserons.

En un point x de l'espace, la valeur de l'estimateur est d'autant plus élevée qu'il y a de points de l'échantillon à proximité de x. En pratique, un poids est attribué à chaque point de l'échantillon en fonction de sa distance à x et la moyenne de ces poids donne la valeur de l'estimateur en ce point. Dans notre cas, les poids décroissent en suivant la densité d'une loi gaussienne. Le choix de la matrice de variance H de cette loi est alors crucial.

Une matrice H avec des coefficients trop faibles conduit à une situation de sur-ajustement : l'estimateur de la densité f s'est trop adapté aux données. A l'inverse, des coefficients trop élevés conduisent à un lissage abusif.

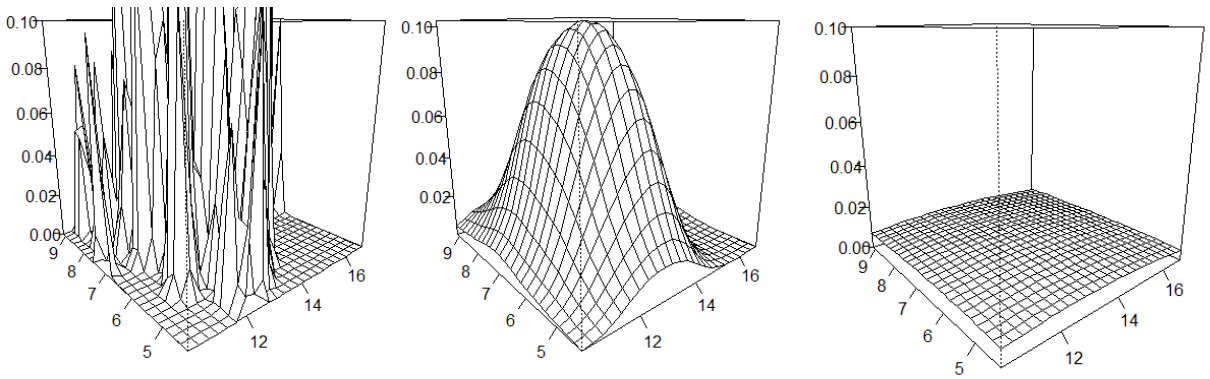


Figure 23 - Estimateurs par noyau d'une densité gaussienne avec 3 matrices de lissage différentes (sur-ajustement à gauche, lissage optimal au centre, lissage abusif à droite)

Dans notre cas, les points $x_1, x_2, \ldots, x_n$ correspondent aux coefficients non-nuls de la matrice de vents forts filtrée. Le but étant que l'estimateur $\hat{f}_H$ admette un seul maximum local par ouragan, prenons $H = \begin{pmatrix} 20^2 & 0 \\ 0 & 20^2 \end{pmatrix}$. Notons que les ouragans sont particulièrement espacés les uns des autres : le risque de lissage abusif (c'est-à-dire le risque

que $\hat{f}_H$ admette un seul maximum pour deux ouragans distincts) est donc moindre et plusieurs choix de matrice de lissage sont alors possibles.

En prenant par exemple la matrice de vents forts filtrée de la figure 21 nous obtenons alors l'estimateur $\hat{f}_H$ suivant :

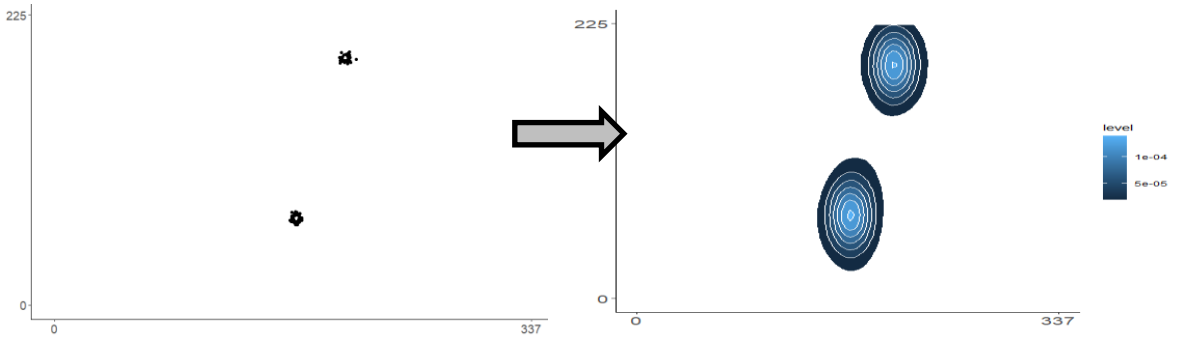


Figure 24 - Matrice de vents forts filtrée passée dans le plan (gauche) et visualisation des lignes de niveau de la densité estimée sur les points visibles à gauche par noyau gaussien (droite)

➢ Recherche des maxima

Après avoir obtenu l'estimateur $\hat{f}_H$ de la densité f, l'algorithme Mean-Shift rattache chaque point x_i de l'échantillon au maximum de $\hat{f}_H$ le plus proche. Les points se trouvant rattachés au même maximum appartiennent au même groupe de points : c'est ainsi que l'algorithme Mean-Shift partitionne. Dans notre cas, les groupes de points formés sont les ouragans et les positions des maxima de $\hat{f}_H$ sont les positions des centres de ces ouragans.

Pour chaque point x_i de l'échantillon est créée une suite de points $(x_i^t)_{t\geq 1}$ convergeant vers la position du maximum de $\hat{f}_H$ le plus proche.

En pseudo-code :

Pour chaque point x_i de l'échantillon :

Calcul des deux premiers termes de la suite : $x_i^1 \leftarrow x_i$ et $x_i^2 \leftarrow x_i^1 + m(x_i^1)$

$t \leftarrow 1$

Tant que $\|x_i^{t+1} - x_i^t\| > \varepsilon$:

$x_i^{t+2} \leftarrow x_i^{t+1} + m(x_i^{t+1})$

$t \leftarrow t + 1$

La fonction notée m est définie de la façon suivante :

$$m(x) = \frac{\sum_i x_i \exp(-\frac{1}{2} x^T H^{-1} x)}{\sum_i \exp(-\frac{1}{2} x^T H^{-1} x)} - x$$

Définir la fonction m de cette façon permet la convergence de la suite vers le maximum de $\hat{f}_H$ le plus proche. Davantage de détails sont donnés en annexes.

En choisissant un paramètre ε suffisamment petit, la suite $(x_i^t)_{t\geq 1}$ créée pour un point x_i de l'échantillon converge donc vers le maximum de $\hat{f}_H$ le plus proche. Comme dit précédemment, deux points x_i et x_j dont les suites $(x_i^t)_{t\geq 1}$ et $(x_j^t)_{t\geq 1}$ convergent vers le même maximum sont regroupés au sein du même groupe et donc décrivent un même ouragan. Le maximum local en question est considéré comme le centre du groupe et donc le centre de l'ouragan.

Ainsi, sur les matrices de vents forts, l'algorithme Mean-Shift nous permet d'obtenir le nombre et la localisation des ouragans. Notons que cette localisation correspond à un indice de ligne et un indice de colonne dans la matrice initiale, pas directement à une latitude et une longitude.

e) Suivi des ouragans

Étudier les matrices de vents forts nous a permis d'obtenir des informations sur les ouragans présents dans nos données durant leur période de forte intensité. Cependant, la plus grande partie de l'empreinte de ces ouragans nous est encore inconnue et cela pose de nombreux problèmes, pour compter les ouragans et ainsi obtenir des fréquences annuelles par exemple.

Considérons une succession de matrices de vent comportant une matrice de vents forts positionnée en i et supposons qu'elle ne contient qu'un seul ouragan :

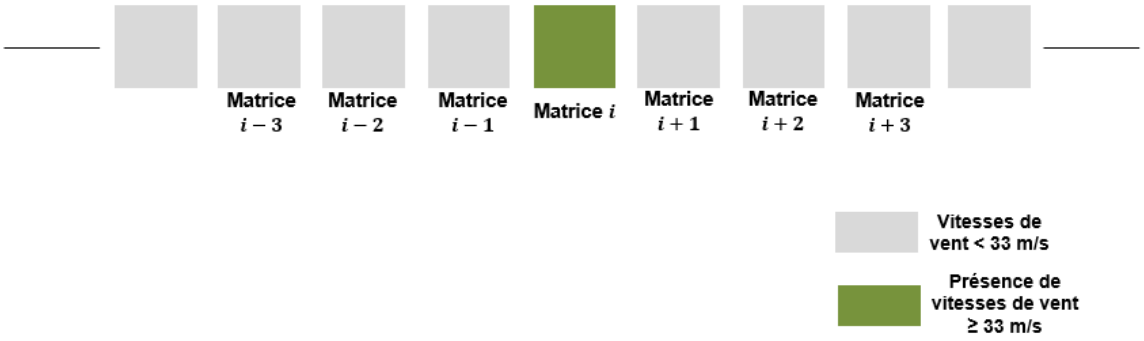


Figure 25 - Série de matrices de vitesses de vent comprenant une matrice de vents forts

Notre travail va porter sur les matrices grises du schéma et, dans un premier temps, sur celles à proximité de la matrice de vents forts, à savoir les matrices $i-1$ et $i+1$. Nous allons tenter de rattacher les ouragans observés sur la matrice de vents forts aux perturbations observées sur ces 2 matrices. Si ce rattachement est possible, nous nous intéresserons aux matrices $i-2$ et $i+2$. Nous continuerons ainsi jusqu'à ce que le rattachement soit, des deux côtés, impossible. Nous aurons ainsi reconstruit toute l'empreinte de l'ouragan.

Voici le pseudo-code de la démarche :

$k \leftarrow 0$

Tant que début ou fin de l'empreinte pas trouvé :
    Si rattachement $i + k / i + k + 1$ possible et fin pas trouvée:
        Compléter l'empreinte avec la position de
        l'ouragan en $i + k + 1$
    Sinon :
        Fin trouvée

    Si rattachement $i - k / i - (k + 1)$ possible et début pas trouvé :
        Compléter l'empreinte avec la position de
        l'ouragan en $i - (k + 1)$
    Sinon :
        Début trouvé

    $k \leftarrow k + 1$

Voyons maintenant plus en détail comment fonctionne le rattachement d'un ouragan observé sur une matrice i avec une perturbation observée sur une matrice $i + 1$.

Dans le cas où la matrice $i + 1$ est une matrice de vents forts, nous disposons déjà de la localisation du ou des ouragans présents sur cette matrice. Il suffit alors de comparer la position de l'ouragan en i avec la position du ou des ouragans en $i + 1$ pour dire si un de ces ouragans correspond à celui observé en i.

Notons C le centre de l'ouragan observé en i et $C_1, C_2, \dots, C_n$ les centres des n ouragans observés en $i + 1$. Notons $\|.\|$ la norme euclidienne sur $\mathbb{R}^2$. Supposons que C_k $(1 \leq k \leq n)$ soit le centre le plus proche de C, c'est-à-dire :

$$\|C_k - C\| = \min_{1 \leq i \leq n} \|C_i - C\|$$

Alors, en définissant un seuil $K \geq 0$, si $\|C_k - C\| \leq K$, nous rattachons les centres C et C_k : l'ouragan présent sur la matrice i est présent sur la matrice $i + 1$ et son centre en $i + 1$ est C_k . En revanche, si $\|C_k - C\| > K$, les centres présents en $i + 1$ sont considérés comme trop éloignés du centre de l'ouragan en i et le rattachement n'a pas lieu.

Le choix du seuil K est alors crucial. Ce dernier dépend de la distance que peut parcourir un ouragan entre les 2 instants où sont capturés les vitesses de vent qui figurent sur les matrices i et $i + 1$. Sachant que 3 heures se sont écoulées entre ces deux instants, avoir une idée de la vitesse à laquelle se déplace un ouragan permettrait d'évaluer la distance parcourue pendant ce laps de temps.

Tranche de latitude	Vitesse		Nombre de cas
	km/h	kt	
0°-5°N	25,5	13,7	100
5°-10°N	21,8	11,8	3282
10°-15°N	19,6	10,6	5808
15°-20°N	17,7	9,5	6086
20°-25°N	17,7	9,5	6817
25°-30°N	20,1	10,9	5321
30°-35°N	27,6	14,9	2835
35°-40°N	38,5	20,8	1026
40°-45°N	46,7	25,2	263
45°-50°N	51,4	27,7	36
50°-55°N	51,4	27,8	15
55°-60°N	55,8	30,1	1

Figure 26- Vitesses de déplacement des ouragans de l'Atlantique moyennées par tranches de 5 degrés de latitude (source : Météo-France)

Comme nous pouvons le voir dans ce tableau, la vitesse de déplacement d'un ouragan est étroitement liée à la latitude à laquelle il se trouve. Étant donné que le domaine CAM ne va pas au-delà de 40°N, nous pouvons raisonnablement considérer que les ouragans présents dans nos données ont une vitesse de déplacement maximale de 40 km/h environ. Ainsi, les ouragans que nous étudions ne peuvent parcourir plus de 120 km entre les instants i et $i + 1$. Prenons donc $K = 120$ km.

Dans le cas où la matrice $i + 1$ est une matrice de vents faibles, c'est-à-dire contenant des vitesses de vent inférieures à 33 m/s, la démarche est la même à une différence près : nous ne disposons pas des centres des perturbations sur la matrice $i + 1$ et nous allons donc devoir les déterminer dans un premier temps.

Pour déterminer les centres des perturbations en $i + 1$, commençons par filtrer la matrice $i + 1$ au seuil de 16 m/s. Ce seuil, volontairement très bas, correspond au seuil de vitesse de vent d'une dépression tropicale selon l'échelle de Saffir-Simpson. Des vents de vitesse inférieure à 16 m/s ne causeraient aucun dégât.

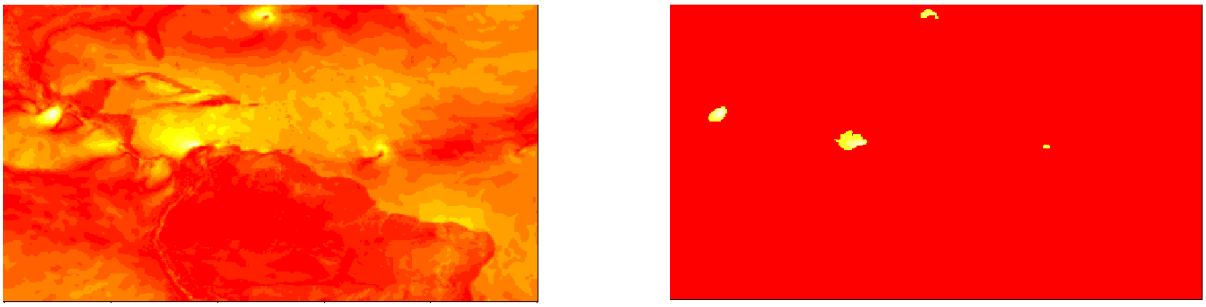


Figure 27 - Visualisation d'une matrice de vitesses de vent (gauche) et de la même matrice filtrée au seuil de 16 m/s (droite)

Une fois la matrice $i + 1$ filtrée, nous allons déterminer le centre de chaque perturbation grâce à l'algorithme Mean-Shift exécuté sur les coefficients non-nuls de la matrice (comme dans le cas d'une matrice de vents forts).

Ensuite, nous comparerons les distances entre le centre C de l'ouragan en i et les centres $C_1, C_2, \dots, C_n$ des perturbations en $i + 1$. Le seuil K est le même que celui précédemment choisi.

A l'issue de cette étape de suivi, nous disposons donc de l'empreinte de chaque ouragan présent dans nos données. Cependant, le nombre annuel d'ouragans détectés sur les simulations climatiques aux conditions historiques est bien trop élevé par rapport au nombre annuel observé historiquement (provenant de la base HurDat, détaillée par la suite) :

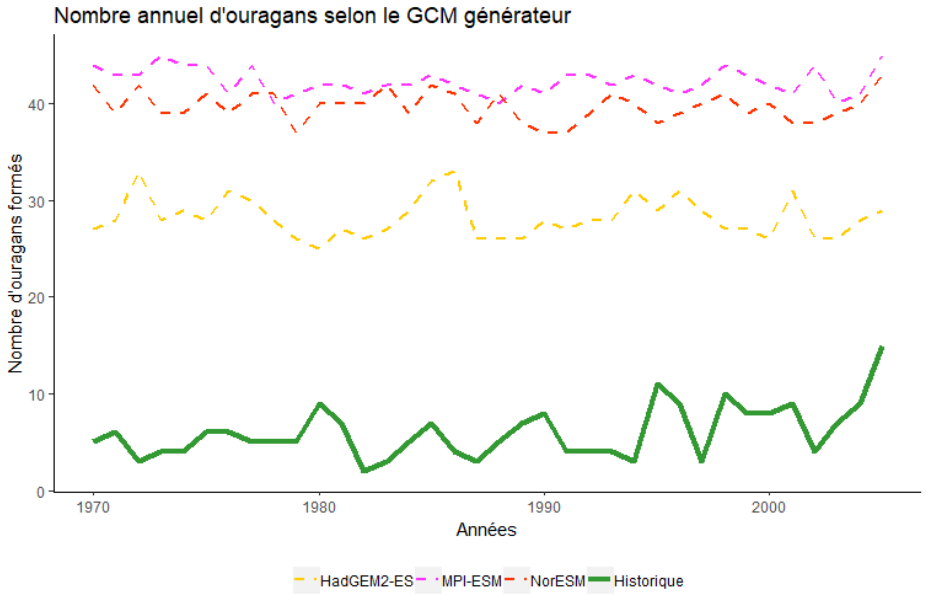


Figure 28 – Nombre annuel d'ouragans détectés selon le GCM générateur (courbes en pointillés, conditions historiques utilisées) par rapport à l'historique (courbe verte en trait plein)

Quel que soit le GCM utilisé, le nombre d'ouragans détectés chaque année est au moins 3 fois plus élevé que le nombre historique.

Voyons maintenant quelles sont les raisons de cette surestimation et comment l'éviter.

4. Retraitements

a) Lieu de formation des ouragans

Pour rappel, voici un aperçu de la zone couverte par le domaine CAM :

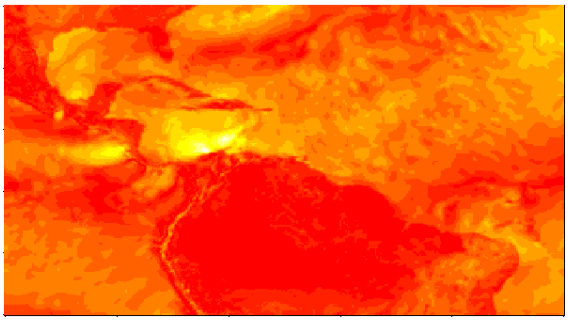


Figure 29 - Visualisation d'une matrice de vitesses de vent tronquée

Comme nous pouvons le voir, malgré le fait que nous ayons tronqué la matrice en enlevant certaines colonnes à gauche, des coefficients continuent d'être associés à des vents soufflant au-dessus de l'océan Pacifique. C'est un problème puisque cette partie du Pacifique est un bassin d'ouragans : une partie des ouragans détectés ne se forment donc pas dans les Caraïbes et ne concernent donc pas la zone Antilles.

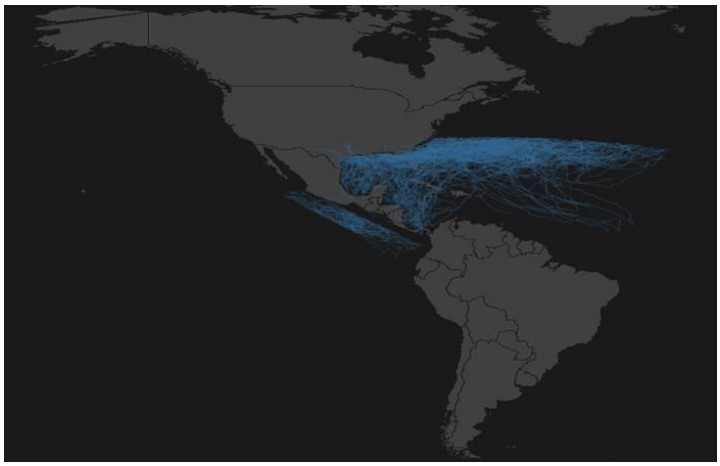


Figure 30 - Empreintes issues du modèle HadGEM2-ES dans le cadre du scénario RCP 8.5 sur la période 2050-2070

Nous devons donc retirer des ouragans détectés tous ceux se formant dans l'océan Pacifique.

b) Durée des ouragans

La surestimation du nombre d'ouragans vient en grande partie de la définition même du phénomène d'ouragan et de la façon dont nous transposons cette définition à nos données. Pour rappel, un ouragan est une perturbation tropicale engendrant des vents dont la vitesse dépasse, à un moment donné, le seuil de 33 m/s.

La méthode que nous avons mise en place permet bien de détecter les ouragans de notre jeu de données. Cependant, diverses perturbations de faible intensité sont comptabilisées comme des ouragans à cause d'un critère trop souple. En effet, sur une matrice de vent, un seul et unique coefficient supérieur à 33 m/s suffit pour donner naissance à un nouvel ouragan selon notre méthode d'extraction. Il nous faut alors trouver d'autres critères afin de caractériser plus précisément un ouragan.

Étudier la littérature scientifique existante permet d'avoir un aperçu de critères supplémentaires. Citons par exemple : une température à la surface de l'océan d'au moins 26°C lors de la formation de la perturbation ou encore la présence d'un cercle de pression minimale fermé (Diro *et al.*). Or, dans notre cas, ces critères ne sont pas applicables puisque nous ne disposons que d'informations sur les vitesses de vent.

Les perturbations de faible intensité que nous avons détectées et que nous cherchons à exclure se différencient des ouragans par leur durée de vie, c'est-à-dire le laps de temps écoulé entre la formation de la perturbation et sa disparition. Nous allons donc utiliser cette durée de vie comme critère : si la perturbation a une durée de vie trop faible (quelques jours seulement), nous allons considérer qu'il ne s'agit pas d'un ouragan. La durée de vie minimale pour qu'une perturbation soit considérée comme un ouragan va être donnée par l'étude des ouragans historiques.

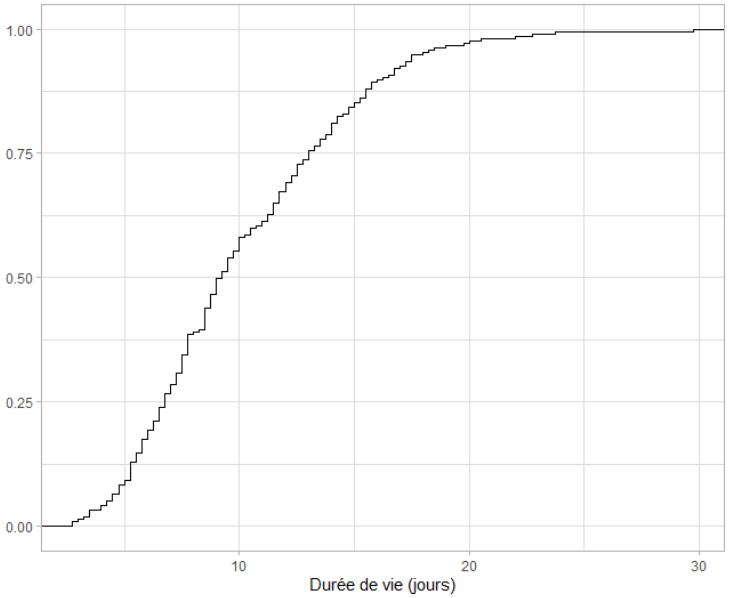


Figure 31 - Fonction de répartition empirique de la durée de vie des ouragans historiques (période 1970-2005)

Nous pouvons voir que la majorité des ouragans durent plus de 5 jours. Plus précisément, le quantile 5% de la distribution est 4.5 jours et nous allons alors utiliser cette durée comme critère.

En conclusion, une perturbation détectée dans nos données va être considérée comme un ouragan si elle engendre, à un moment donné, des vents de vitesse supérieure ou égale à 33 m/s et si sa durée de vie excède 4.5 jours.

c) Vitesse des vents

Comme vu précédemment, les vitesses de vent les plus élevées générées par un ouragan se trouvent généralement à proximité de l'œil de ce dernier. Obtenir les vitesses des vents générées par les ouragans présents dans nos données est alors simple : l'empreinte d'un ouragan nous fournit la position de l'œil sur les matrices de vent à chaque instant et il suffit alors de s'intéresser aux coefficients de ces matrices situés à proximité de l'œil.

Tout d'abord, précisons que l'empreinte d'une perturbation, donc la trajectoire de cette dernière, est déterminée par une succession de localisations et que, dans le cadre de ce mémoire, nous appellerons ces localisations des *positions*. Pour une empreinte donnée, il est commun de ne retenir que la vitesse maximale générée pour chaque position de l'empreinte. Ainsi, dans toute la suite de ce mémoire, les vitesses étudiées correspondront donc à des séries de vitesse maximale par position et par *vitesses de vent associées à une empreinte*, il faudra comprendre *série des vitesses maximales générées en chaque position de l'empreinte*.

Étudions les vitesses de vent générées par les 3 GCMs aux conditions historiques :

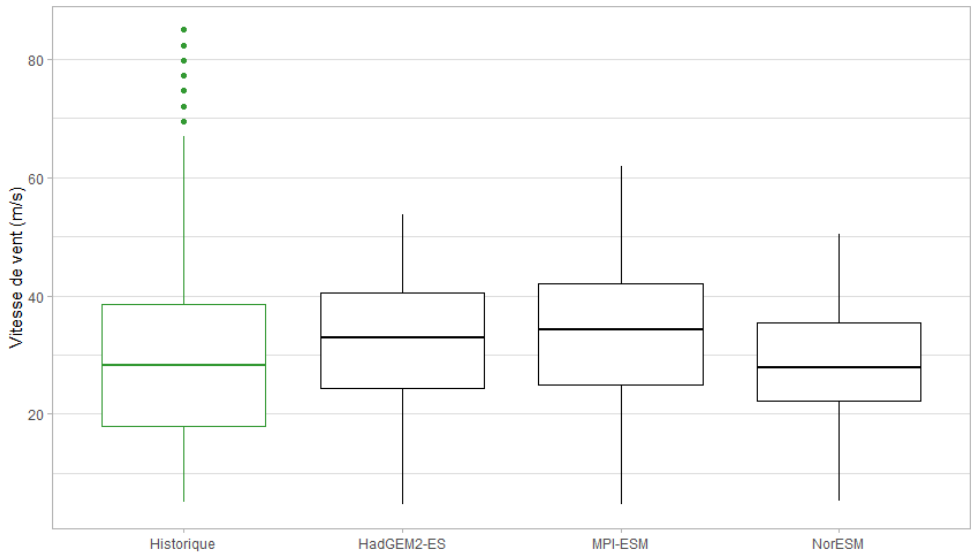


Figure 32 - Boîtes à moustaches construites sur les vitesses associées aux empreintes historiques (période 1970-2005, en vert) et sur celles associées aux empreintes générées par les 3 GCMs aux conditions historiques (en noir)

Nous pouvons voir ici que pour deux des trois GCMs, les vitesses générées ne dépassent pas 55 m/s : le seuil de vitesse de la catégorie 4 de l'échelle de Saffir-Simpson étant 59 m/s, aucun ouragan de catégorie 4 ou 5 n'est alors généré par ces 2 GCMs. Le modèle MPI génère quant à lui des vitesses allant jusqu'à 62 m/s environ : il ne génère alors aucun ouragan de catégorie 5.

Il y a donc manifestement une inadéquation entre vitesses historiques et vitesses générées par les modèles climatiques. Ce problème de sous-estimation des vitesses de vent extrêmes est courant et n'est pas propre aux données CORDEX : bien que la maille choisie soit fine, elle ne l'est pas assez pour étudier précisément des phénomènes aussi locaux et complexes que les ouragans. Les experts du premier groupe de travail du cinquième rapport du GIEC le rappellent d'ailleurs dans leurs conclusions sur les cyclones tropicaux : « En général, certains modèles haute résolution de l'atmosphère […] sont capables de saisir les caractéristiques générales des trajectoires des tempêtes et des cyclones […] mais les modèles […] sous-estiment l'intensité des cyclones. ». Dans une publication de 2016 portant sur la sensibilité de l'activité cyclonique simulée sur le domaine CAM à différentes paramétrisations du GCM générateur, Fuentes-Franco *et al.* ont mis en évidence la même sous-estimation des vitesses de vent extrêmes. Ils remarquent notamment le lien fort existant entre cette sous-estimation et la modélisation des forces de frottement agissant sur le vent à la surface de l'océan : lorsque la vitesse de vent est

élevée, ces forces augmentent de façon trop importante et brident la vitesse du vent. Les ouragans simulés et le couple GCM/RCM utilisés dans l'étude ne sont cependant pas les mêmes que ceux utilisés dans ce mémoire : les conclusions de l'étude de Fuentes-Franco *et al.* doivent donc être considérées avec prudence.

Les ouragans de catégories 4 et 5 étant les plus dévastateurs et donc les plus coûteux, nous ne pouvons continuer l'étude sans corriger les vitesses générées par les GCMs. La méthode employée est une méthode de correction courante en météorologie : la méthode quantile-quantile.

- Correction quantile-quantile

La méthode de correction quantile-quantile est utilisée pour corriger une distribution donnée à partir d'une distribution initiale : corriger des données simulées à partir de données historiques observées par exemple.

Dans notre cas, nous allons corriger les vitesses simulées à partir des vitesses historiques. Nous allons donc corriger les vitesses de vent maximales générées en chaque position de chaque empreinte simulée par les vitesses de vents maximales générées en chaque position de chaque empreinte historique. La distribution des vitesses de vent étant dépendante du GCM ayant généré ces vents, nous procéderons séparément pour les 3 GCMs.

Considérons la série des vitesses de vent associées aux empreintes historiques observées et la série des vitesses de vent associées aux empreintes simulées par un GCM. La méthode de correction quantile-quantile consiste à associer le quantile des données simulées avec le quantile du même ordre obtenu sur données historiques. Par exemple, une vitesse de vent de 33 m/s correspond au quantile 67% de la distribution des données simulées. Cette vitesse, une fois corrigée, sera le quantile 67% de la distribution des données historiques.

Plus précisément, notons q_{histo}^{k} le quantile d'ordre k de la distribution des vitesses de vent associées aux empreintes historiques et notons q_{sim}^{k} le quantile d'ordre k de la distribution des vitesses de vent associées aux empreintes simulées par un GCM. Notons V_{sim} une vitesse de vent provenant de la série des vitesses de vent associées aux empreintes simulées par un GCM et V_{corr} la vitesse V_{sim} corrigée par la méthode quantile-quantile.

Supposons que $q_{sim}^{k} \leq V_{sim} \leq q_{sim}^{k+1}$. Alors :

$$V_{corr} = V_{sim} * \left(\frac{q_{histo}^{k}}{q_{sim}^{k}} \right)$$

Si V_{sim} est inférieure au quantile 1%, la même transformation est appliquée en utilisant les quantiles 1% : q_{sim}^{k} et q_{histo}^{k} sont les quantiles 1% de la distribution simulée et de la distribution historique. Si V_{sim} est supérieure au quantile 99%, q_{sim}^{k} et q_{histo}^{k} sont les quantiles 99% de la distribution simulée et de la distribution historique.

Le quantile q_{sim}^{k} étant dépendant de la distribution des vitesses et donc du GCM générateur, trois corrections différentes sont réalisées : une par GCM. Une fois les données générées

aux conditions historiques corrigées, les projections selon les scénarii RCP 2.6 et RCP 8.5 sont corrigées avec les mêmes séries de coefficients $\left(\frac{q^k_{histo}}{q^k_{sim}}\right)_k$, propres à chaque GCM.

Voyons maintenant comment sont distribuées les vitesses corrigées par rapport aux vitesses historiques :

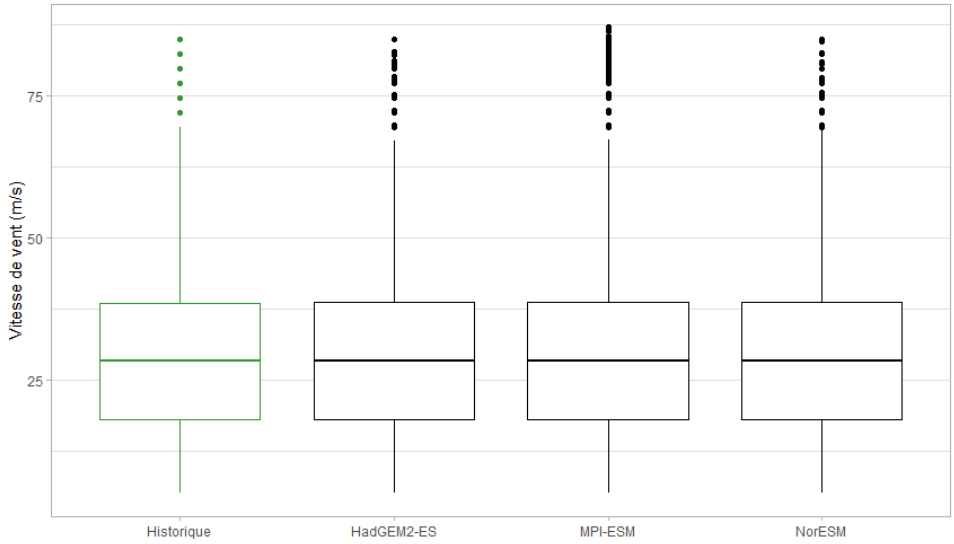


Figure 33 - Boîtes à moustaches construites sur les vitesses associées aux empreintes historiques (période 1970-2005, en vert) et sur celles associées aux empreintes générées par les 3 GCMs aux conditions historiques après correction (en noir)

Nous remarquons sur le graphique ci-dessus qu'une fois corrigées, les vitesses de vent associées aux empreintes générées par les 3 GCMs aux conditions historiques sont davantage réalistes : des ouragans de catégories 4 et 5 sont bien générés. Nous analyserons plus en détails ces simulations dans la suite de ce mémoire.

Les distributions de vitesses de vent générées selon les scénarii RCP étant corrigées avec les mêmes coefficients que celles générées aux conditions historiques, il est très important que ces distributions soient similaires. La méthode quantile-quantile repose donc sur une hypothèse forte de stationnarité que nous ferons ici.

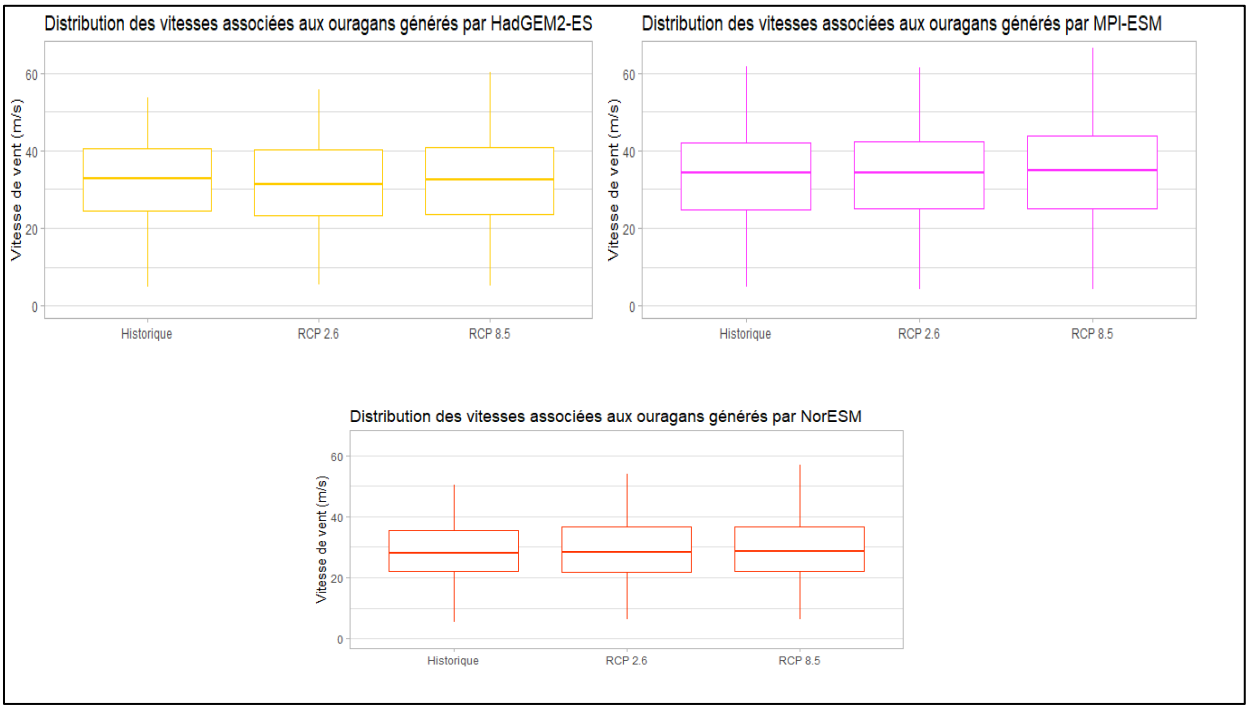


Figure 34 – Distributions des vitesses associées aux ouragans générés par chacun des trois GCMs aux conditions historiques (boîte à moustaches de gauche), sous scénario RCP 2.6 (milieu) et sous scénario RCP 8.5 (droite)

Nous pouvons voir sur la figure ci-dessus que les distributions de vitesses associées à des ouragans générés aux conditions historiques sont très similaires aux vitesses enregistrées sur des ouragans générés sous les scénarii RCP.

Lorsque nous parlerons des vitesses de vent simulées par un GCM, nous ferons à partir de maintenant référence aux vitesses de vent corrigées par la méthode quantile-quantile.

Maintenant que nous avons extrait les ouragans des données issues du projet CORDEX, comparons-les aux ouragans historiques et voyons si ces données permettent de constater des changements de comportement à l'horizon 2050 ou 2100.

IV. Étude des ouragans extraits

Rappelons que les projections climatiques du projet CORDEX proviennent, dans notre cas, de 3 GCMs différents. Dans un premier temps, nous allons étudier les données produites aux conditions historiques puis nous nous intéresserons aux simulations obtenues dans le cadre des scénarii RCP 2.6 et RCP 8.5

Afin d'évaluer la qualité des données produites aux conditions historiques, nous allons nous appuyer sur la base de données de référence du bassin Atlantique : la base HurDat.

1. Présentation de la base de données historiques

La base HurDat est une base de données gérée par le National Hurricane Center situé à Miami. Ce centre de recherche public étudie les ouragans du bassin Atlantique, c'est d'ailleurs le Centre

Météorologique Régional Spécialisé (CMRS) de ce bassin. HurDat rassemble des informations diverses sur les ouragans et tempêtes du bassin Atlantique de 1850 à nos jours : l'empreinte, la vitesse des vents générés ou la pression au centre sont données. Ces informations sont disponibles avec un pas de temps de 6 heures. Cette base est la base utilisée par la communauté scientifique et les agences de modélisation : la plupart des modélisations de tempêtes ou d'ouragans dans le bassin Atlantique s'appuient sur cette base historique.

	Date et heure	Latitude	Longitude	Vitesse de vent maximale (moyenne sur 1 minute à 10m de hauteur)	Pression au centre
Position 1					
Position 2					
Position 3					
…					

Tableau 5 - Aperçu des informations disponibles dans la base HurDat pour une perturbation donnée

Il est important de noter que les informations présentes dans la base HurDat ne proviennent pas directement de relevés météorologiques. En effet, la récolte de données météorologiques est récente et présente de nombreux biais, bien que ces derniers tendent à s'amenuir avec les avancées technologiques récentes. De plus, la mer des Caraïbes est un territoire majoritairement marin : les stations météorologiques ne couvrent donc qu'une très faible partie de ce territoire. Les données de la base HurDat sont en fait issues de réanalyses.

La réanalyse combine les observations avec des données issues de simulations dans le but de produire une description plus fine des phénomènes météorologiques. Plus ces derniers sont anciens, moins il existe d'observations météorologiques fiables les décrivant : Henry Diaz et José Fernández-Partagás, météorologues américains, se sont par exemple fait connaître pour avoir recoupé des documents d'archives permettant de décrire les ouragans de la mer des Caraïbes entre 1850 et 1910. Ce sont sur leurs travaux que s'appuient les données de la base HurDat concernant cette période. Les événements plus récents sont, eux, décrits par bien plus de données : elles peuvent provenir de stations météorologiques, de navires, d'avions ou encore de satellites.

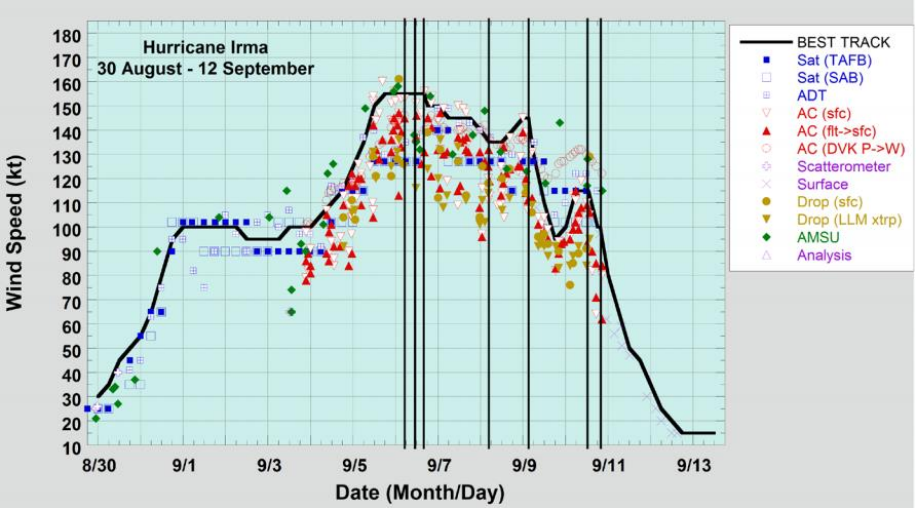


Figure 35 - Vitesse de vent moyenne sur une minute mesurée par différentes sources à différentes dates pour l'ouragan Irma en 2017 (source : NHC)

Notons sur le graphique ci-dessus la courbe noire correspondant à la *Best Track* : c'est l'estimation de la vitesse de vent réelle faite à partir des observations. Cette courbe est donc issue d'une réanalyse et nous retrouverons cette vitesse de vent estimée dans la base HurDat.

En règle générale, les agences de modélisation n'utilisent que les informations de la base HurDat décrivant des événements ayant eu lieu après 1950.

2. Études des ouragans générés aux conditions historiques

Voyons dans un premier temps quels sont les mois les plus propices à la formation d'ouragans :

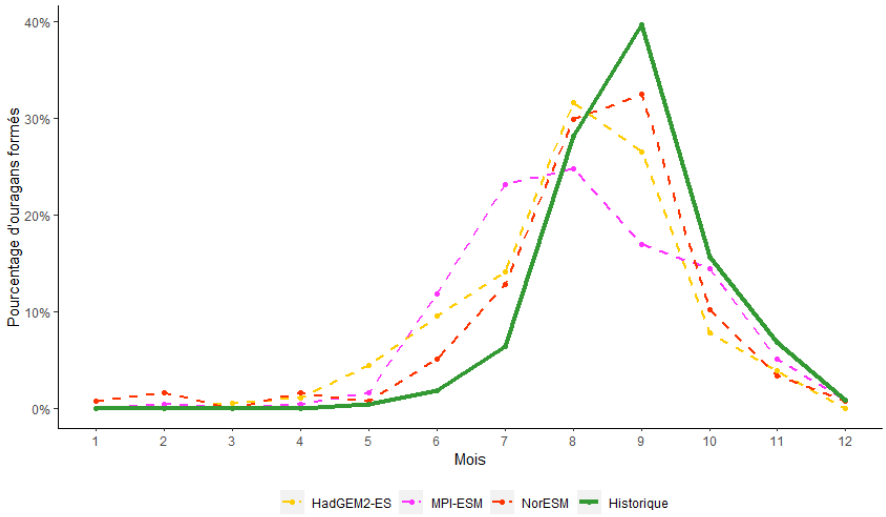


Figure 36 - Pourcentage d'ouragans formés par mois sur données historiques observées (courbe verte) et sur données simulées aux conditions historiques (courbes en pointillés)

Comme expliqué précédemment, la saison cyclonique démarre en juin et termine fin novembre. Nous pouvons voir que les 3 GCMs en tiennent compte puisque qu'une très faible proportion d'ouragans est simulée en dehors de cette période. Cependant, cette saison des ouragans démarre plus ou moins tôt suivant le GCM utilisé. En effet, si la proportion mensuelle d'ouragans simulés est très semblable à l'historique pour NorESM, les deux autres modèles génèrent une saison des ouragans démarrant plus tôt que celle observée historiquement : notons par exemple les 11,9% d'ouragans simulés au mois de juin pour MPI-ESM contre 1,8% historiquement.

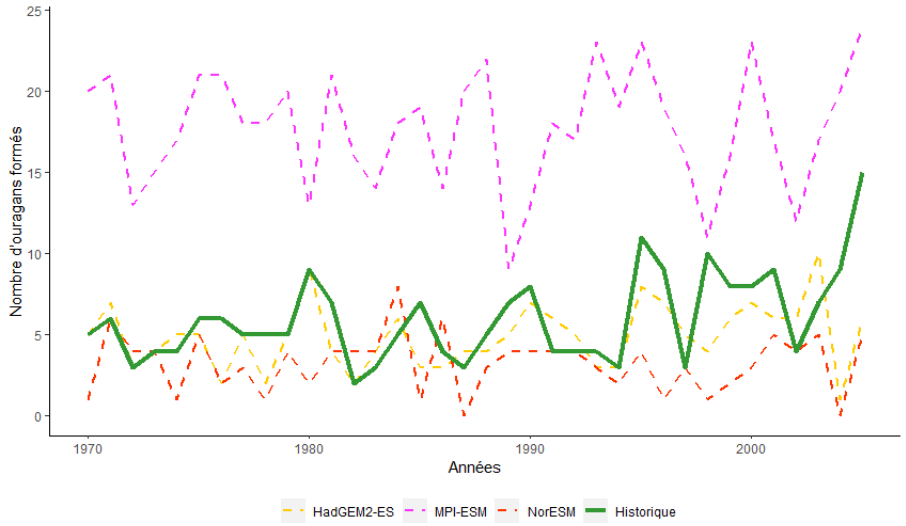


Figure 37 - Nombre d'ouragans formés par an (période 1970-2005, domaine CAM) sur données historiques observées (courbe verte) et sur données simulées aux conditions historiques (courbe en pointillés)

Remarquons ici que le nombre d'ouragans formés historiquement par an reste stable sur la période 1970-2005. Cette stabilité est bien observable sur les simulations climatiques. Notons tout de même que le nombre d'ouragans simulés chaque année par le modèle MPI-ESM est bien plus élevé que le nombre d'ouragans observés annuellement dans l'historique (de l'ordre de 4 fois plus élevé). Quant aux deux autres GCMs, ils simulent chaque année un nombre d'ouragans semblable à celui observé historiquement.

Avant d'étudier les empreintes des ouragans simulés, rappelons que l'empreinte d'un ouragan correspond à sa trajectoire. Cette dernière est donnée par une succession de positions et 3 heures s'écoulent entre chacune d'elles. Nous ferons, à partir de maintenant, l'hypothèse que l'empreinte d'un ouragan correspond à l'interpolation linéaire de ses différentes positions. Cette hypothèse de trajectoire linéaire entre deux positions est réaliste étant donnée la faible distance parcourue par l'ouragan en 3 heures : 120 km au maximum, environ 60 km de manière générale.

Voici les empreintes des ouragans historiques sur la période 1970-2005 :

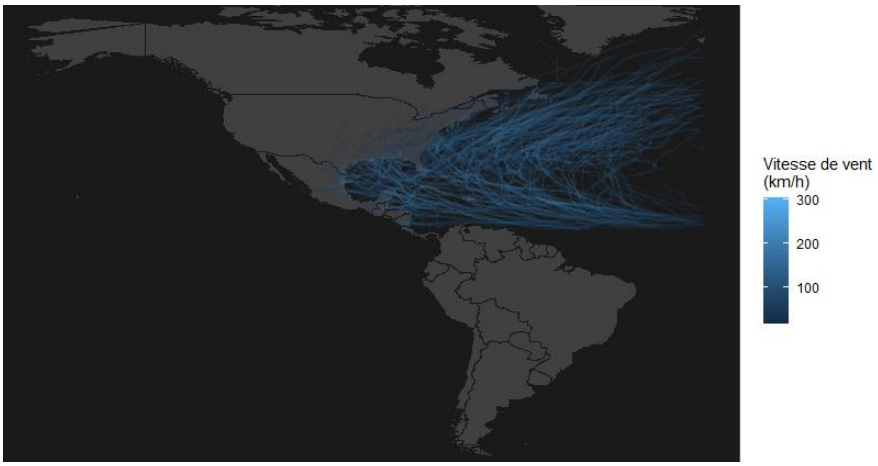


Figure 38 - Empreintes des ouragans historiques observés dans le bassin Atlantique (période 1970-2005)

Notons que les ouragans partent du Sud et remontent vers le Nord dans une trajectoire en arc de cercle. Cette trajectoire s'explique par la présence de la crête subtropicale de l'hémisphère Nord : la crête subtropicale est une zone de haute pression comprise entre les latitudes 25°N et 35°N pour l'hémisphère Nord, 25°S et 35°S pour l'hémisphère Sud. L'air se déplaçant des zones de haute pression vers les zones de basse pression, la crête va engendrer des vents soufflant vers le Sud et d'autres soufflant vers le Nord :

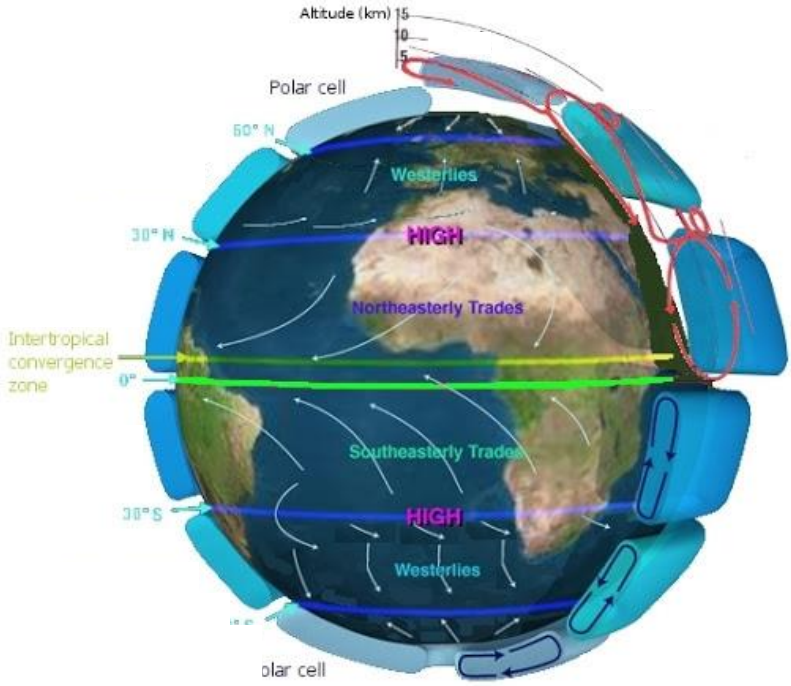


Figure 39 - Crête subtropicale (visible à 30°N et 30°S) et vents engendrés (source : NASA)

Ces vents expliquent la trajectoire d'Est en Ouest des ouragans dans la première partie de leur vie, puis la déviation vers le Nord-Est dans la seconde partie de leur vie, dès lors qu'ils franchissent la crête subtropicale.

Notons que les vitesses de vent les plus faibles sont enregistrées en début et en fin de trajectoire, en début et fin de vie des ouragans.

La force de Coriolis étant quasi nulle au voisinage de l'Equateur, les ouragans ne peuvent se former plus au Sud que ceux observés ci-dessus.

Intéressons-nous maintenant aux empreintes produites par les 3 GCMs aux conditions historiques de la période 1970-2005 :

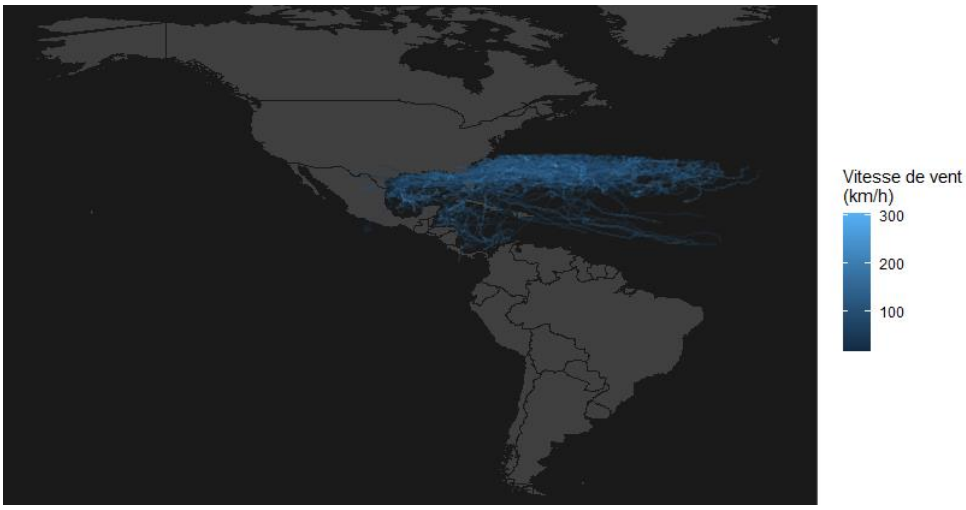


Figure 40 - Empreintes issues du GCM HadGEM2-ES avec conditions historiques (période 1970-2005)

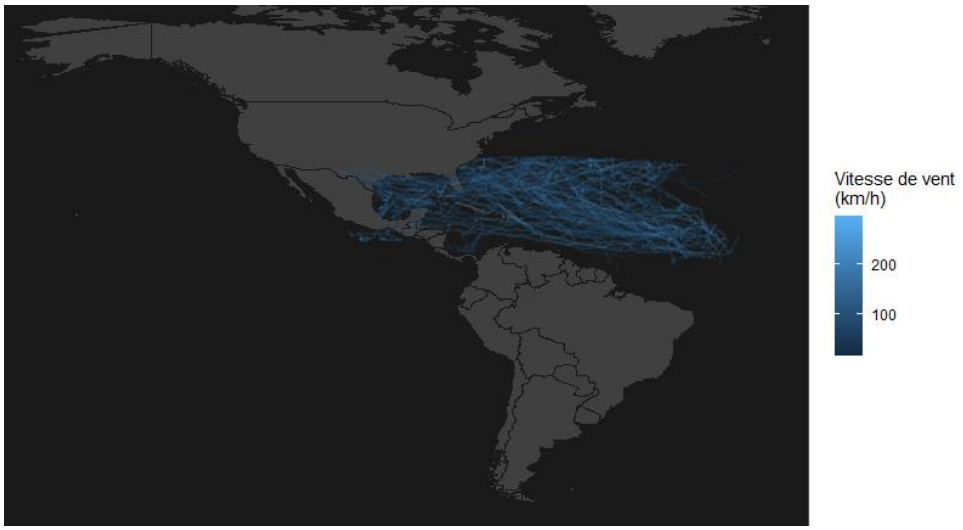


Figure 41 - Empreintes issues du GCM MPI-ESM avec conditions historiques (seule la période 1970-1976 est affichée pour une question de lisibilité)

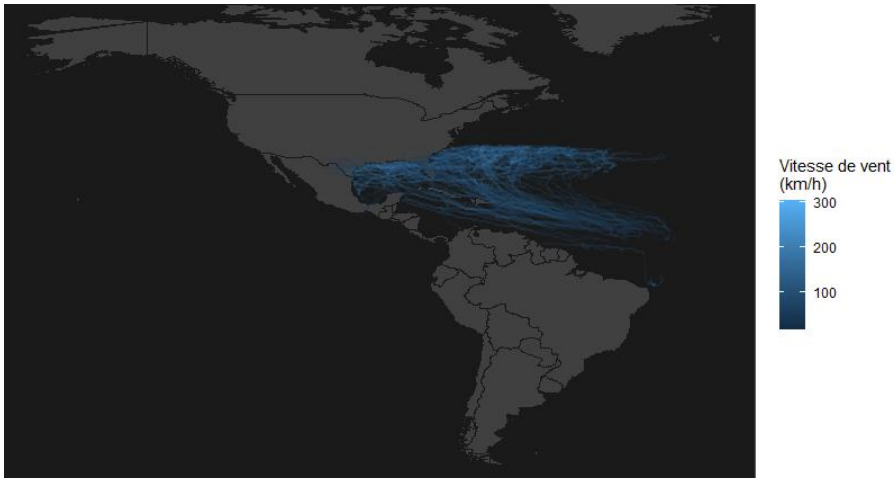


Figure 42 - Empreintes issues du GCM NorESM avec conditions historiques (période 1970-2005)

Notons tout d'abord que les empreintes ont été tronquées au Nord : c'est une conséquence du domaine CORDEX choisi qui limite l'étude des empreintes à une zone précise.

Nous pouvons remarquer que les empreintes varient fortement suivant le GCM générateur : si les empreintes obtenues avec HadGEM2-ES sont peu fidèles aux empreintes historiques, celles obtenues avec NorESM et MPI-ESM sont, à première vue, de meilleure qualité : bien que les empreintes générées par MPI-ESM n'aient pas la forme en arc de cercle caractéristique des ouragans de la mer des Caraïbes, les zones traversées par les ouragans simulés sont celles traversées par les ouragans observés historiquement.

Voyons maintenant plus en détails comment sont exposées les différentes zones du bassin. Dans un premier temps, définissons deux portes : une porte Nord et une porte Sud. Une porte est une ligne droite s'étendant sur plusieurs centaines de kilomètres et dont le rôle est de caractériser l'exposition d'un territoire en comptant les ouragans la franchissant. Afin de caractériser précisément l'exposition d'un territoire, nous allons également compter le nombre d'ouragan frappant ce dernier. Déterminer si un ouragan frappe ou non un territoire à partir de son empreinte et des vitesses de vent associées n'est pas chose facile : nous considérerons que si un ouragan se trouve à moins d'une certaine distance (132 km) du territoire et que les vents engendrés dépassent les 33 ms/s (le seuil de vitesse d'un ouragan de catégorie 1), le territoire est frappé. Voyons comment est calculée cette distance de 132 km.

A partir de l'année 2004, une variable supplémentaire est renseignée dans la base historique HurDat : en chaque position de chaque empreinte, le radius de vent au seuil de 33 m/s est donné. Ce radius est le rayon du cercle dont le centre est l'ouragan et indiquant jusqu'à quelle distance de l'ouragan les vitesses de vent sont supérieures à 33 m/s.

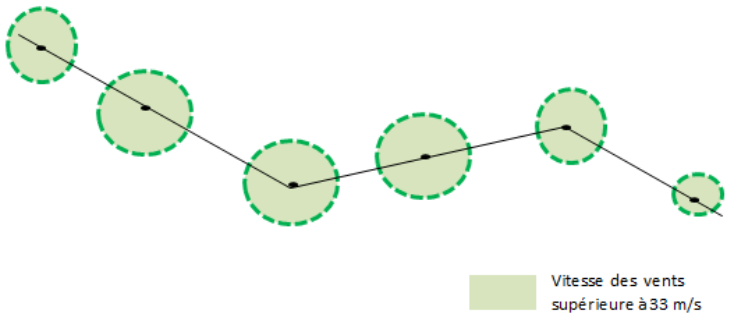


Figure 43 - Empreinte d'un ouragan avec cercles de vitesse de vent au seuil 33 m/s en chaque position

Cette variable s'avère être très utile puisque le seuil de 33 m/s est le seuil de vitesse des vents d'un ouragan de catégorie 1. Ces vents peuvent occasionner des dégâts et nous considérerons qu'un territoire est frappé par un ouragan si ce territoire est couvert, totalement ou partiellement, par un des cercles de vitesse de vent de l'ouragan au seuil de 33 m/s.

Ces radius apparaissent dans la base de données historiques mais nous ne pouvons les obtenir précisément sur les simulations climatiques à cause de la maille trop large. Un modèle prédictif avait alors été construit et entraîné sur les données historiques : à partir de la vitesse maximale des vents générés en une position et de la latitude de cette même position, le modèle estimait le radius de vent au seuil de 33 m/s. Appliquer ce modèle aux

ouragans simulés aurait permis d'obtenir une estimation du radius de vent en chaque position de chaque empreinte.

L'estimation n'était cependant pas assez précise et il a été alors choisi de procéder de façon conservative. Quand les vitesses de vent générées dépassent 33 m/s, le radius du cercle de vitesse 33 m/s est considéré comme constant sur la période 1970-2005 et égal au quantile 95% de la série des radius de la base HurDat sur la période 2004-2018, à savoir 132 km.

Voyons maintenant comment sont exposées les différentes zones du bassin.

➢ Portes Nord et Sud

Les portes Nord et Sud couvrent les zones d'entrée et de sortie des ouragans dans la mer de Caraïbes :

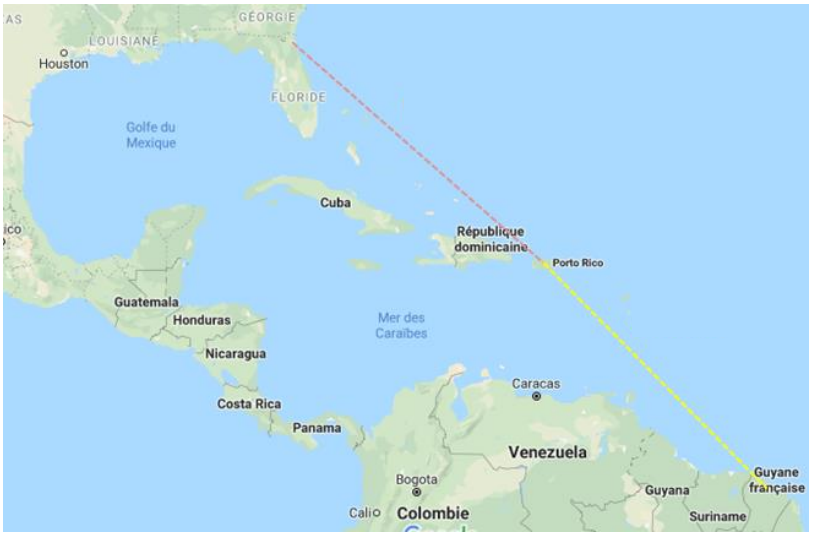


Figure 44 - Visualisation des portes Nord (en rose) et Sud (en jaune)

Pour chacune de ces deux portes, les fréquences de passage suivantes sont observées :

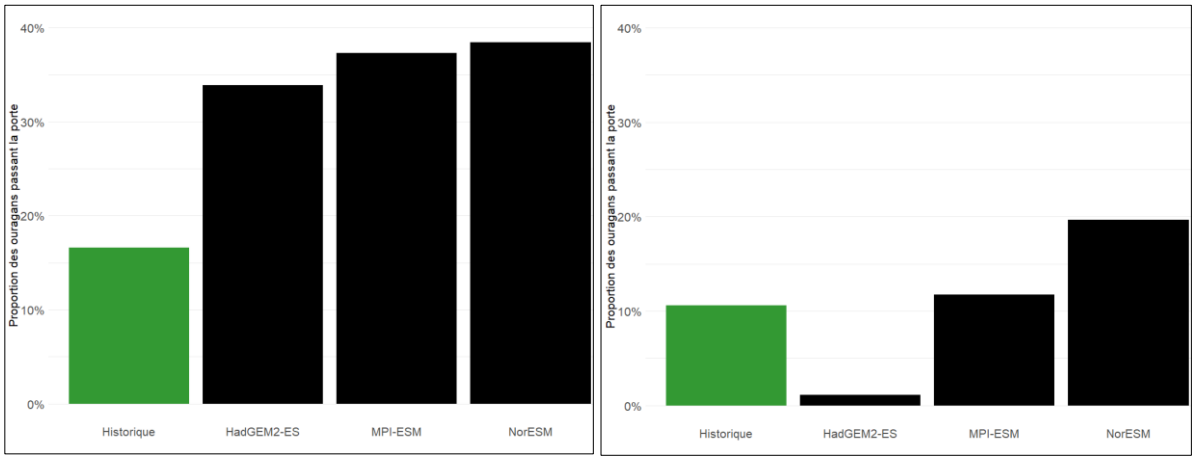


Figure 45 - Proportion des ouragans passant la porte Nord (gauche) et la porte Sud (droite) selon qu'ils soient générés par un GCM (barres noires) ou observés historiquement sur la période 1970-2005 (barre verte)

Nous pouvons remarquer que les 3 modèles surestiment la proportion d'ouragans franchissant la porte Nord : ceci provient du fait qu'historiquement, une partie des ouragans observés entament leur virage vers le Nord avant cette porte ou après, une fois dans les terres américaines. Or, nous pouvons par exemple voir sur les figures 41, 42 et 43 que les ouragans générés ne passent que très peu sur ces terres américaines. Les expositions de la caisse Antilles-Guyane se trouvant dans le Sud du bassin, cette surestimation n'impactera pas directement notre étude.

La porte Sud, quant à elle, donne un aperçu du début de vie des ouragans générés. Si la proportion d'ouragans produits par MPI-ESM passant par cette porte est semblable à la proportion historique (de l'ordre de 11%), cette dernière est sous-estimée par HadGEM2-ES, surestimée par NorESM. La sous-estimation par le modèle du Hadley Centre était prévisible étant donné le décalage des empreintes produites par ce modèle vers le Nord (voir figure 41).

- <u>Fréquence à laquelle les territoires de la caisse sont frappés</u>

Remarquons que la fréquence à laquelle les ouragans générés frappent les territoires de la caisse dépend fortement du GCM générateur :

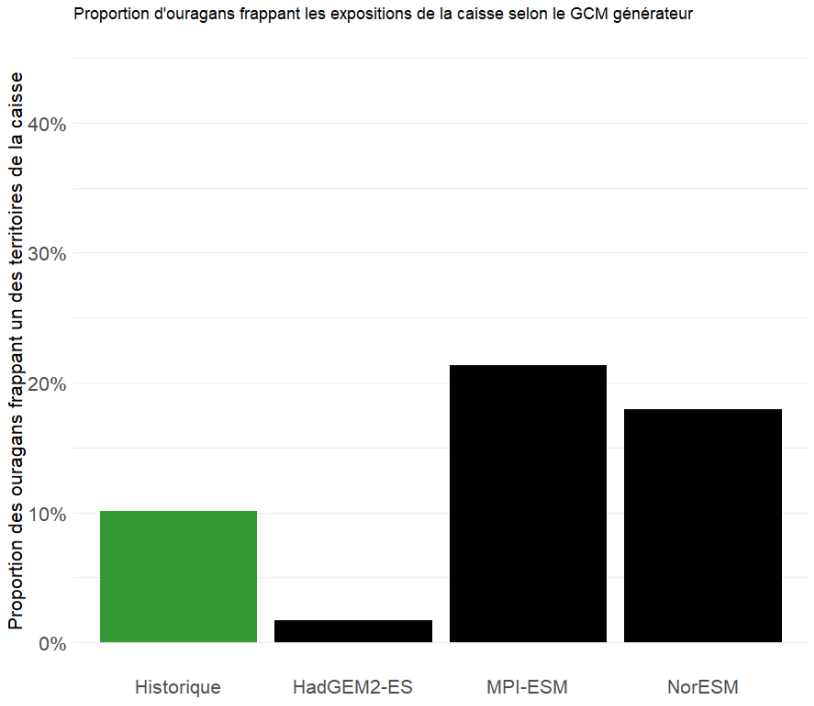


Figure 46 - Proportion des ouragans générés frappant les territoires de la caisse Antilles-Guyane selon qu'ils soient générés par un GCM (barres noires) ou observés historiquement sur la période 1970-2005 (barre verte)

La figure ci-dessus nous confirme l'aperçu obtenu avec la porte Sud : le modèle du Hadley Centre génère une proportion trop faible d'ouragans passant dans le Sud de la mer des Antilles et donc près de nos expositions, alors que la proportion générée par NorESM est trop importante. Pour MPI-ESM, cette zone a permis d'affiner notre diagnostique : bien que la proportion d'ouragans générés passant par la porte Sud soit

réaliste, plus de 20% des ouragans générés frappent les expositions de la caisse contre seulement 10% historiquement.

Ces fortes variations selon le GCM générateur sont également observables lorsque l'on étudie chacun des 4 territoires séparément.

Au-delà de la caractérisation des empreintes des ouragans générés par nos 3 modèles climatiques, ces observations permettent de donner un premier aperçu de la sinistralité engendrée par ces modèles. Sous réserve que nous utilisions ces empreintes et que nous puissions correctement estimer la sinistralité associée à un ouragan, le modèle du Hadley Centre semble conduire à une sous-estimation de la sinistralité historique, contrairement aux deux autres modèles qui semblent conduire à une surestimation de cette sinistralité.

Nous disposons pour chaque ouragan de la vitesse de vent maximale enregistrée en chaque position de l'empreinte. Nous pouvons alors facilement déterminer la catégorie (sur l'échelle de Saffir-Simpson) de chaque ouragan.

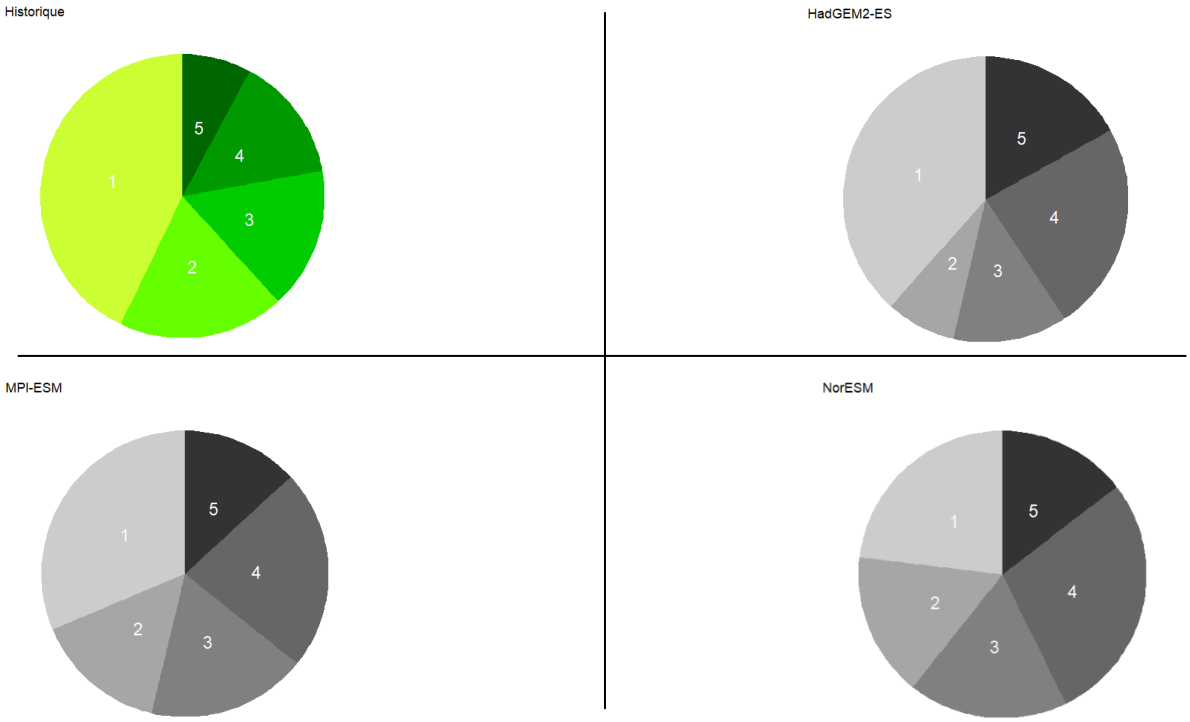


Figure 47 - Répartition selon les 5 catégories de l'échelle de Saffir-Simpson des ouragans observés historiquement (période 1970-2005) et simulés sous conditions historiques

Les ouragans les plus puissants sont également les moins fréquents et nous pouvons le voir ci-dessus. Si le modèle MPI-ESM génère des ouragans dont la répartition en catégories de Saffir-Simpson est semblable à celle observée historiquement, HadGEM2-ES et NorESM sont quant à eux plus sévères : la proportion d'ouragans de catégories 4 et 5 est d'environ 41% pour ces deux GCMs contre 22% historiquement.

Nous pouvons donc constater que ces 3 modèles, plus ou moins sensibles au changement climatique, génèrent des ouragans dont le comportement n'est pas en adéquation totale avec les données historiques. Les points d'adéquation et d'inadéquation peuvent être résumés de façon sommaire par le tableau suivant :

	HadGEM2-ES	**MPI-ESM**	**NorESM**
Fréquence annuelle et mensuelle	Adéquation	Inadéquation	Adéquation
Empreintes	Inadéquation	Adéquation	Adéquation
Vitesses de vent	Adéquation partielle	Adéquation	Inadéquation

Tableau 6 - Points d'adéquation/inadéquation entre ouragans observés historiquement et ouragans simulés selon le GCM utilisé

Ces points d'inadéquation peuvent surprendre au premier abord étant donnée la popularité de ces modèles climatiques. Cependant, ces derniers projettent de très nombreuses variables (une centaine) jusqu'en 2100. L'adéquation partielle avec les données historiques d'un phénomène aussi local et extrême que les ouragans, observé à l'aide d'une seule de ces variables, ne doit donc pas surprendre.

Rappelons également que le RCM utilisé joue un rôle important. Le modèle REMO est utilisé ici mais l'utilisation d'un autre RCM, comme RegCM par exemple, pourrait conduire à des résultats bien différents. Au moment de la réalisation de cette étude, il n'était cependant pas disponible pour le domaine CAM.

3. Études des ouragans générés selon le scénario RCP 2.6

Il s'agit ici d'observer les ouragans générés par nos 3 modèles climatiques dans le cadre du scénario RCP 2.6. Dans cette partie et dans la suivante où nous nous intéresserons au scénario RCP 8.5, on définit la période historique comme la période 1970-2005, l'horizon 2050 comme la période 2031-2050 et l'horizon 2100 comme la période 2081-2100.

Notons dans un premier temps qu'aussi bien à l'horizon 2050 qu'à l'horizon 2100, les ouragans continuent d'être majoritairement formés pendant la saison cyclonique, entre juin et novembre. Les proportions d'ouragans générés pendant cette période correspondent aux proportions observées sur les données générées aux conditions historiques.

Voyons maintenant comment évolue le nombre annuel d'ouragans tout au long du XXI^{ème} siècle :

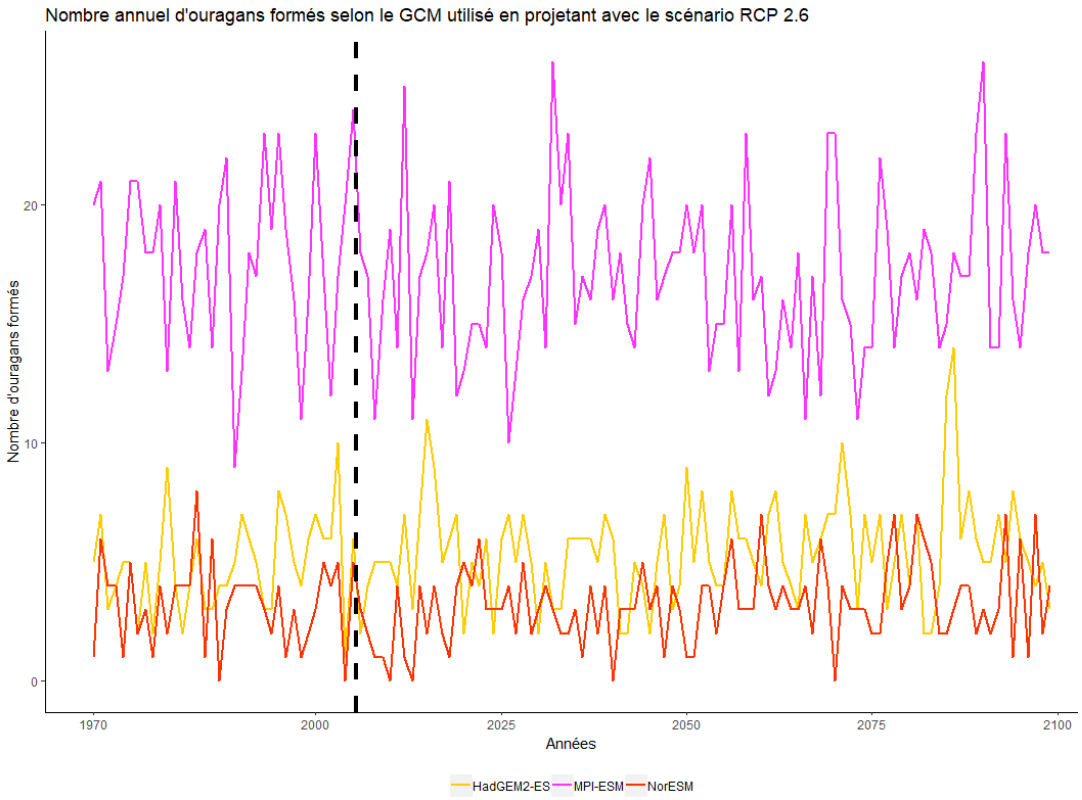


Figure 48 – Nombre annuel d'ouragans formés sur la période 1970-2100 pour chacun des 3 GCMs. Les conditions historiques sont utilisées avant 2005 (trait vertical noir), le scénario RCP 2.6 est utilisé après 2005.

Nous pouvons constater ci-dessus que dans le cadre du scénario RCP 2.6, le nombre annuel d'ouragans reste stable tout au long du XXI^{ème} siècle quel que soit le GCM utilisé, ce qui concorde avec les prévisions du GIEC.

Aussi bien à l'horizon 2050 qu'à l'horizon 2100, les empreintes issues des simulations dans le cadre du scénario RCP 2.6 ne diffèrent pas des simulations aux conditions historiques : il n'y a, visuellement, aucune différence. En particulier, l'étude de la latitude moyenne à laquelle les ouragans atteignent leur pic d'intensité ne montre pas de déplacement vers le Nord de cette dernière.

Voyons si la fréquence à laquelle sont frappés les territoires de la caisse Antilles-Guyane augmente au cours du XXI^{ème} siècle :

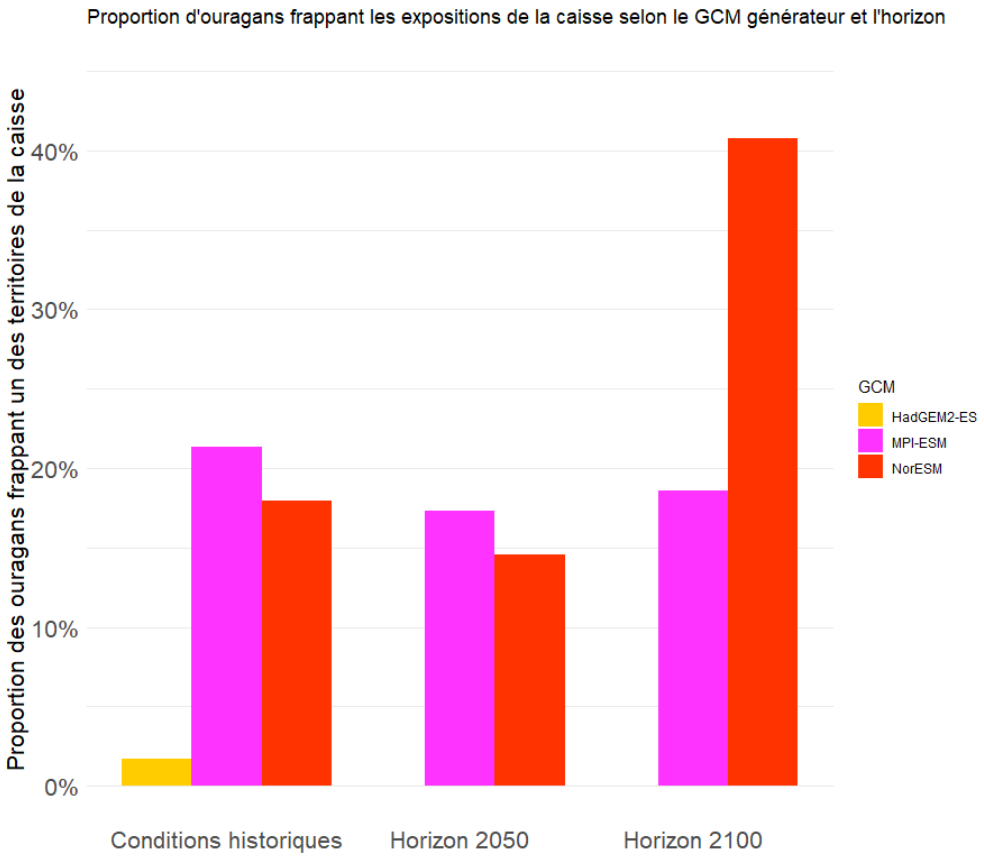


Figure 49 - Proportion des ouragans générés frappant les territoires de la caisse Antilles-Guyane selon le GCM générateur et l'horizon (projections réalisées sous scénario RCP 2.6)

Comme vu précédemment, les empreintes générées par le modèle HadGEM2-ES ne sont pas réalistes : nous voyons même qu'à horizon 2050 et horizon 2100, aucun des ouragans générés par ce modèle ne frappe les territoires de la caisse. Remarquons également une augmentation de la fréquence à laquelle les territoires de la caisse sont frappés entre l'horizon 2050 et l'horizon 2100 : si cette augmentation est moindre pour le modèle allemand, elle est considérable pour le modèle moins sensible au changement climatique. En effet, 4 ouragans sur 10 générés par le modèle norvégien à l'horizon 2100 frappent les territoires de la caisse. Étant donnée la fréquence annuelle stable d'ouragans générés par le modèle NorESM au cours du XXI^{ème} siècle, nous pouvons conclure que selon ce modèle, le nombre d'ouragans frappant les territoires de la caisse va doubler d'ici 2100. En revanche, pour MPI-ESM, ce nombre resterait stable.

En ce qui concerne la fréquence à laquelle sont touchées les îles de Guadeloupe et de Martinique :

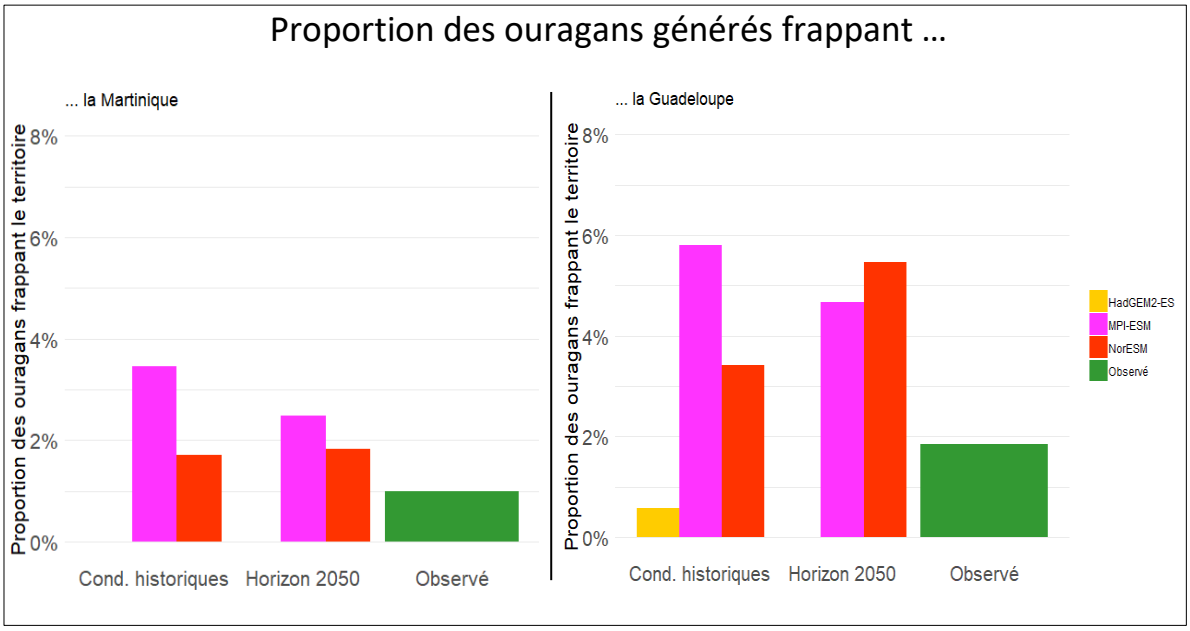


Figure 50 - Proportion des ouragans générés frappant la Martinique (gauche) et la Guadeloupe (droite) selon la nature des ouragans (simulés par GCM ou observés historiquement) et l'horizon (projections réalisées sous scénario RCP 2.6)

Nous avons pu constater plus haut que MPI-ESM et NorESM surestiment la proportion d'ouragans frappant la caisse Antilles-Guyane. Nous voyons ci-dessus que cette constatation est toujours valable pour la Martinique et la Guadeloupe prises séparément.

Aussi, les projections faites à l'aide des modèles MPI-ESM et NorESM permettent de confirmer les prévisions de la CCR pour la Martinique, qui avançait une variation de la fréquence à laquelle les ouragans frappent ce territoire comprise entre -35% et +51%. La variation avancée pour la Guadeloupe, comprise entre +17% et +68%, n'est quant à elle pas observable dans le cadre du scénario RCP 2.6.

Remarquons que pour MPI-ESM, la baisse de fréquence est très similaire pour les deux îles. La même observation est faisable pour NorESM mais pour l'augmentation de fréquence. Cette similarité de variation de la fréquence entre les deux îles s'explique par la proximité géographique de ces dernières.

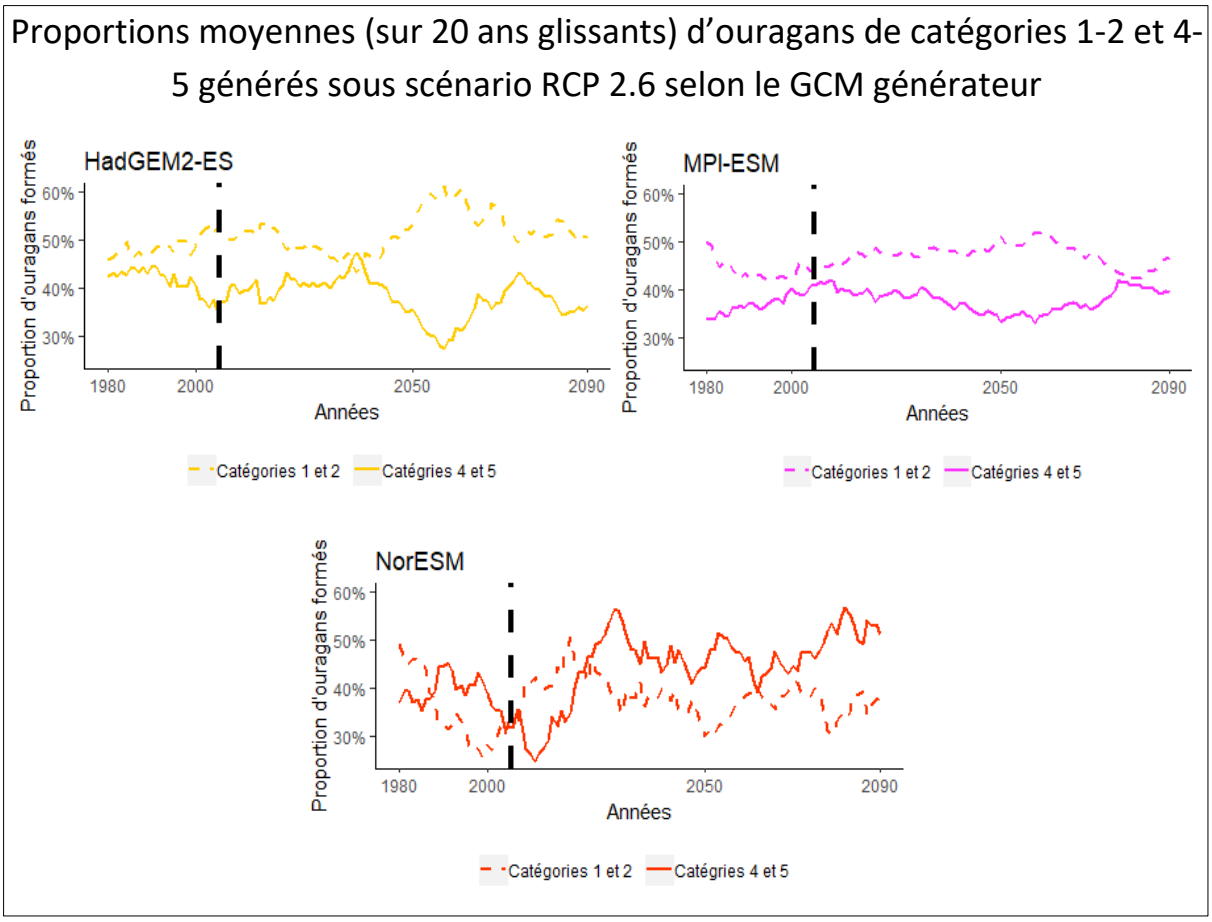
L'échelle de Saffir-Simpson permettant de décrire simplement la puissance d'un ouragan à partir de la vitesse maximale de vent qu'il engendre, voyons comment évoluent les proportions d'ouragans générés par catégorie tout au long du siècle :

	Catégorie 1	Catégorie 2	Catégorie 3	Catégorie 4	Catégorie 5
HadGEM2-ES					
Historique	38,4%	7,9%	13,0%	23,7%	16,9%
Horizon 2050	43,7%	3,1%	13,5%	16,7%	22,9%
Horizon 2100	39,8%	11,0%	14,4%	17,8%	16,9%
Écart histo/2050	+5,3	-4,8	+0,5	-7	+6
Écart histo/2100	+1,4	+3,1	+1,4	-5,9	-
MPI-ESM					
Historique	31,3%	14,9%	18,0%	22,6%	13,1%
Horizon 2050	31,6%	14,6%	15,9%	23,3%	14,6%
Horizon 2100	31,7%	14,5%	14,2%	22,0%	17,5%
Écart histo/2050	+0,3	-0,3	-2,1	+0,7	+1,5
Écart histo/2100	+0,4	-0,4	-3,8	-0,6	+4,4
NorESM					
Historique	23,1%	16,2%	17,9%	28,2%	14,5%
Horizon 2050	32,7%	10,9%	12,7%	18,2%	25,4%
Horizon 2100	27,6%	14,5%	6,6%	14,5%	36,8%
Écart histo/2050	+9,6	-5,3	-5,2	-10	+10,9
Écart histo/2100	+4,5	-1,7	-11,3	-13,7	+22,3

Tableau 7 – Proportions d'ouragans générés par catégorie de Saffir-Simpson selon le GCM générateur (projections faites avec le scénario RCP 2.6)

Précisons que les proportions pour la période « historique » sont obtenues avec des données générées aux conditions historiques et que les proportions pour les horizons 2050 et 2100 sont obtenues avec des données générées dans le cadre du scénario RCP 2.6.

Dans un premier temps, nous pouvons remarquer que les proportions de catégories générées par les modèles HadGEM2-ES et MPI-ESM ne varient que très peu en milieu et fin de siècle par rapport à la période historique. Si les proportions d'ouragans de catégorie 1 et 5 augmentent très légèrement à l'horizon 2100, les proportions de catégories 2 et 4 baissent tout aussi légèrement et il y a alors compensation. De plus, si certaines tendances semblent apparaître à l'horizon 2050, par exemple la variation de -4,8% pour la catégorie 2 avec HadGEM2-ES, elles sont atténuées ou parfois même inversée à l'horizon 2100. En étudiant précisément l'évolution des proportions d'ouragans faiblement et fortement intenses générés, nous constatons que cette faible variabilité des proportions est en fait observable tout au long du siècle :



Note : l'axe des abscisses donne le centre de la fenêtre de 20 années sur laquelle sont calculées les deux proportions moyennes. Par exemple, l'abscisse « 1980 » donne les deux proportions moyennes sur la période 1970 – 1989.

Figure 51 – Proportions moyennes (sur 20 ans glissants) d'ouragans de catégories 1-2 (trait en pointillés) et 4-5 (trait plein) générés sous scénario RCP 2.6 (après 2005, conditions historiques utilisées avant 2005) selon le GCM générateur

Le même compensation entre catégories que pour les deux modèles précédents est observable pour NorESM mais avec des variations de proportions plus importantes : la proportion d'ouragans de catégorie 5 augmente significativement à l'horizon 2100 par rapport aux proportions dites « historiques », alors que les proportions d'événements de catégories 3 et 4 diminuent elles aussi significativement à l'horizon 2100. Encore une fois, l'étude précise de l'évolution des proportions d'ouragans faiblement et fortement intenses générés confirme les tendances observées à l'horizon 2050 et 2100.

Rappelons que le modèle norvégien est le moins sensible au changement au climatique et qu'il est donc étonnant de trouver des variations de proportions aussi marquées alors même que HadGEM2-ES et MPI-ESM, modèles plus sensibles, génèrent des proportions par catégories bien plus stables au cours du siècle. Concernant les catégories les plus basses, si certaines tendances semblent à apparaître à l'horizon 2050, elles sont atténuées à l'horizon 2100, comme pour les deux GCMs précédents.

Ainsi, aussi bien à l'horizon 2050 qu'à l'horizon 2100, les proportions générées d'ouragans les plus puissants, de catégories 4 et 5, et les moins puissants, de catégorie 1 et 2, restent stables tout au long du XXI^{ème} siècle dans le cadre du scénario RCP 2.6.

Une certaine stabilité transparaît donc des projections faites dans le cadre du scénario RCP 2.6 : stabilité du nombre annuel d'ouragan générés pour les 3 GCMs, ainsi que des empreintes et de la puissance des ouragans générés par les deux modèles les plus sensibles au changement climatique.

4. Études des ouragans générés selon le scénario RCP 8.5

Il s'agit ici d'observer les ouragans générés par nos 3 modèles climatiques dans le cadre du scénario RCP 8.5.

Si les ouragans continuent d'être majoritairement générés pendant la saison cyclonique traditionnelle pour les modèles MPI-ESM et NorESM, le modèle anglais, modèle le plus sensible au changement climatique, génère en fin de siècle des ouragans sur une période plus longue qu'historiquement :

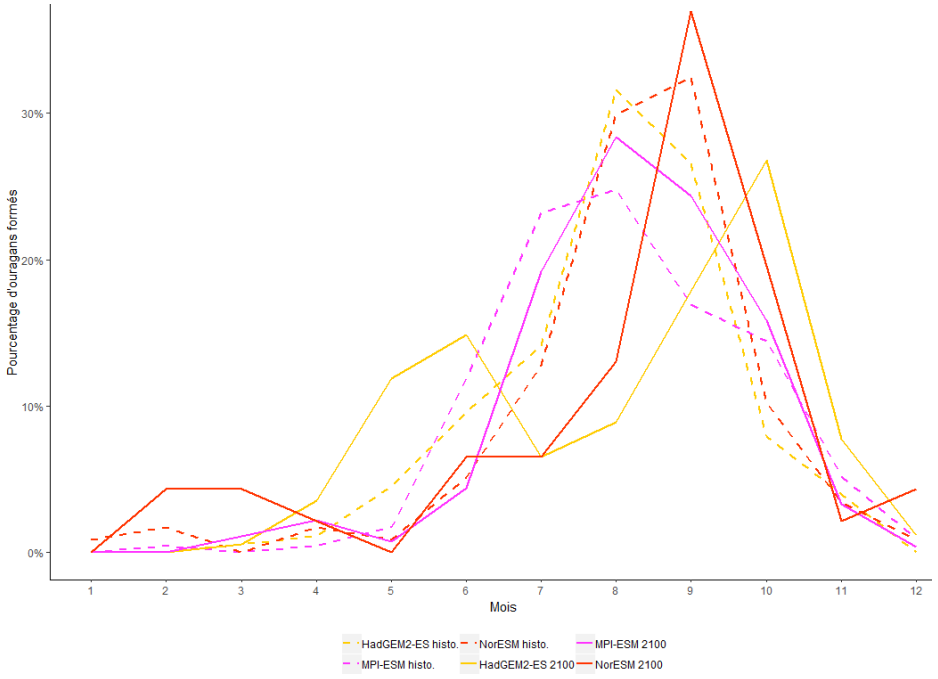


Figure 52 - Proportion d'ouragans formés chaque mois sur données simulées aux conditions historiques (courbes en pointillés) et sur données simulées dans le cadre du scénario RCP 8.5 à horizon 2100 (courbes en trait plein) par GCM

Nous pouvons voir qu'historiquement, HadGEM2-ES génère la majorité des ouragans au mois d'août. Cependant, à l'horizon 2100, ce maximum est décalé au mois d'octobre. Remarquons également des proportions non négligeables, supérieures à 10%, d'ouragans générés au mois de mai et juin. Ceci laisse penser que par l'action du réchauffement climatique, la saison cyclonique traditionnelle évolue et s'élargit. Ce phénomène pourrait s'expliquer par l'augmentation de la température à l'horizon 2100 : historiquement, les ouragans se forment majoritairement à la fin de l'été, lorsque les fortes chaleurs ont pu faire augmenter la température de l'océan, ce dernier devenant alors propice à la formation d'ouragans. A l'horizon 2100, les températures étant plus élevées, l'océan pourrait atteindre la température

propice à la formation d'ouragans bien plus tôt dans l'année et la conserver plus longtemps, d'où l'élargissement de la saison cyclonique.

Voyons ensuite comment évolue le nombre d'ouragans générés chaque année :

Figure 53 – Nombre annuel d'ouragans formés sur la période 1970-2100 pour chacun des 3 GCMs. Les conditions historiques sont utilisées avant 2005 (trait vertical noir), le scénario RCP 8.5 est utilisé après 2005.

Contrairement au scénario RCP 2.6, les modèles ne sont ici pas unanimes sur la tendance que va suivre le nombre annuel d'ouragans dans le bassin des Caraïbes. Si ce nombre reste stable pour le modèle le moins sensible au changement climatique, MPI-ESM et HadGEM2-ES, respectivement moyennement sensible et très sensible au changement climatique, font apparaître des tendances contraires. A partir de l'année 2050, MPI-ESM génère de moins en moins d'ouragans alors que HadGEM2-ES en génère de plus en plus. Les modèles norvégiens et allemands s'accordent donc avec les prévisions du GIEC relatives au changement climatique alors que HadGEM2-ES, pourtant le modèle le plus sensible à ce phénomène, est en désaccord avec ces dernières.

Ainsi, même dans le scénario le plus pessimiste, les modèles s'accordent sur une stabilité du nombre d'ouragans générés à l'horizon 2050. Ce nombre est impacté à l'horizon 2100 par le réchauffement climatique mais les modèles divergent quant à la nature de cet impact.

De même que pour le scénario RCP 2.6, aucune différence n'est visible entre les empreintes historiques et celles obtenues dans le cadre du scénario RCP 8.5. Aucun déplacement vers le Nord de la latitude moyenne à laquelle les ouragans atteignent leur pic d'intensité n'est à noter.

Ainsi, même dans le cadre du scénario le plus pessimiste concernant le réchauffement climatique, les données CORDEX utilisées dans le cadre de cette étude ne permettent pas de constater un déplacement vers le Nord ou le Nord-Est de la position moyenne à laquelle les ouragans atteignent leur pic d'intensité.

Voyons maintenant si la fréquence à laquelle sont frappés les territoires de la caisse Antilles-Guyane augmente au cours du XXI^{ème} siècle en se projetant avec le scénario RCP 8.5 :

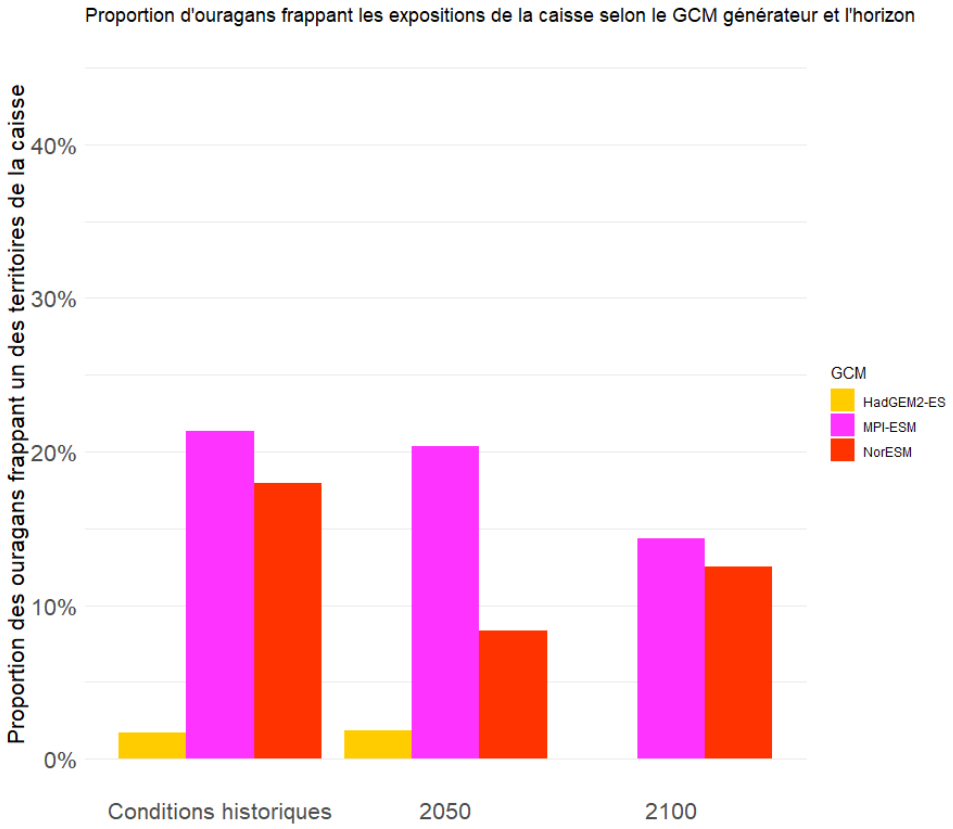


Figure 54 - Proportion des ouragans générés frappant les territoires de la caisse Antilles-Guyane selon le GCM générateur et l'horizon (projections réalisées sous scénario RCP 8.5)

Comme dans le cadre du scénario RCP 2.6, les empreintes générées ici par le modèle anglais ne sont pas exploitables. Bien que le scénario RCP 8.5 soit le plus pessimiste à l'égard du réchauffement climatique, le nombre d'ouragans frappant les territoires de la caisse reste stable au cours du XXI^{ème} siècle. En effet, bien que les ouragans du modèle allemand touchent moins les territoires de la caisse en fin de siècle, le nombre d'ouragans générés diminue lui aussi, comme nous l'avons vu en figure 54. Même si la proportion d'ouragans frappant les territoires de la caisse est divisée par 2 à l'horizon 2050 pour NorESM, cette proportion se stabilise autour de 13% à l'horizon 2100. Ce dernier modèle générant un nombre annuel d'ouragans stable au cours du siècle, le nombre d'ouragans frappant les territoires de la caisse reste stable lui aussi.

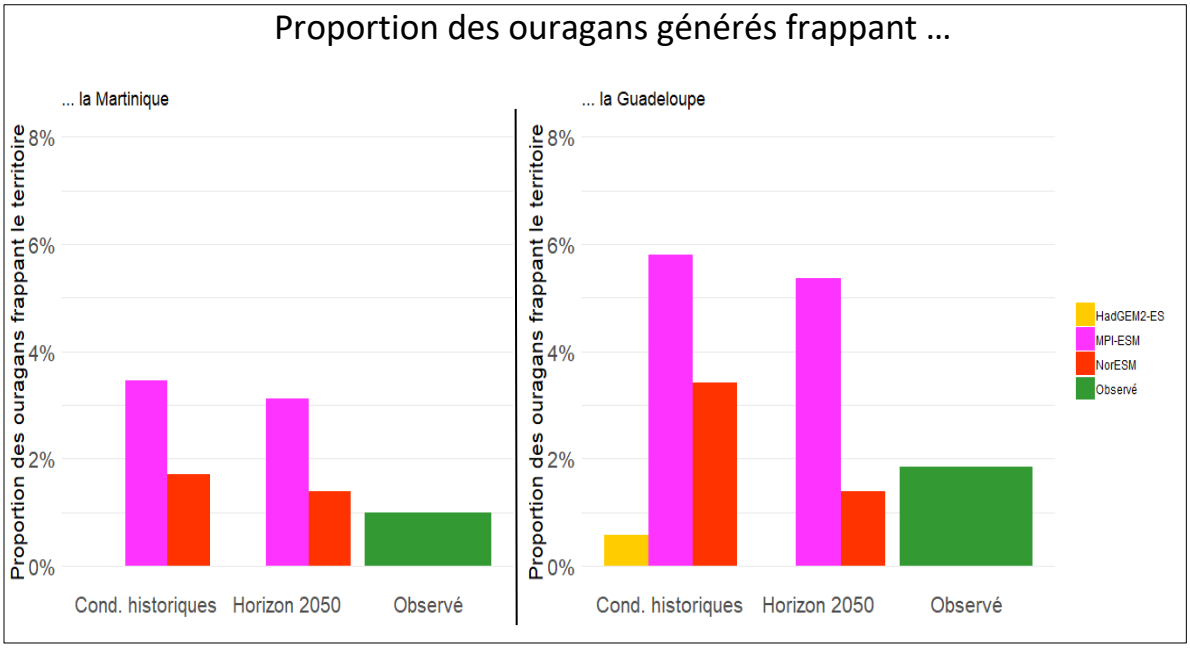


Figure 55 - Proportion des ouragans générés frappant la Martinique (gauche) et la Guadeloupe (droite) selon la nature des ouragans (simulés par GCM ou observés historiquement) et l'horizon (projections réalisées sous scénario RCP 8.5)

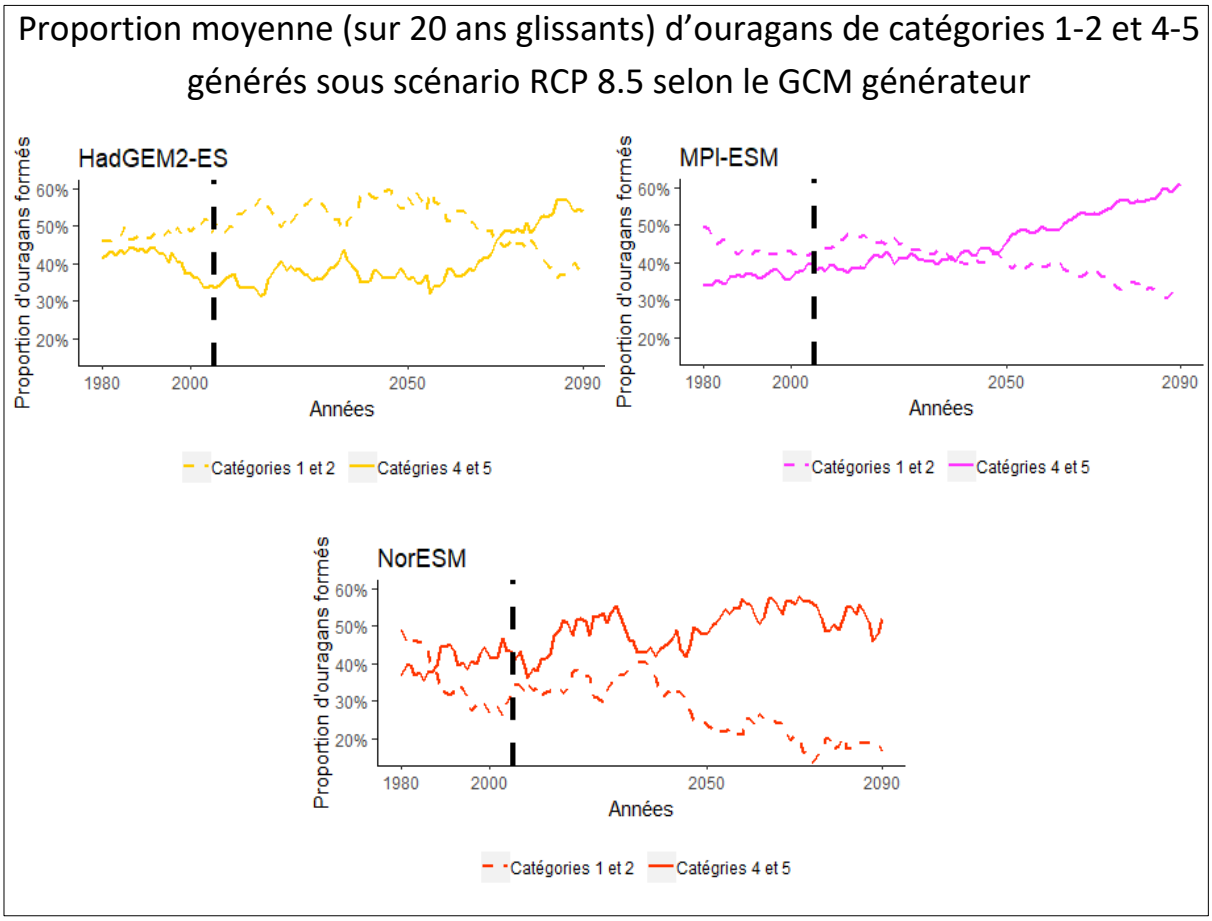
Là encore, les projections permettent de confirmer les prévisions de la CCR concernant la Martinique mais pas celles concernant la Guadeloupe. Les variations de fréquence observées sur les deux îles sont ici encore très semblables du fait de leur proximité géographique.

Comme pour le scénario RCP 2.6, voyons comment évoluent les proportions d'ouragans générés par catégorie tout au long du siècle :

	Catégorie 1	Catégorie 2	Catégorie 3	Catégorie 4	Catégorie 5
HadGEM2-ES					
Historique	38,4%	7,9%	13,0%	23,7%	16,9%
Horizon 2050	42,7%	14,5%	6,3%	18,2%	18,2%
Horizon 2100	25,2%	13,0%	6,9%	27,8%	26,9%
Écart histo/2050	+4,3	+6,6	-6,7	-5,5	+1,3
Écart histo/2100	-13,2	+5,1	-6,1	+4,1	+10
MPI-ESM					
Historique	31,3%	14,9%	18,0%	22,6%	13,1%
Horizon 2050	25,7%	13,0%	16,9%	25,7%	18,6%
Horizon 2100	23,1%	6,6%	8,7%	21,3%	40,2%
Écart histo/2050	-5,6	-1,9	-1,1	+3,1	+5,5
Écart histo/2100	-8,2	-8,3	-9,3	-1,3	+27,1
NorESM					
Historique	23,1%	16,2%	17,9%	28,2%	14,5%
Horizon 2050	23,6%	15,3%	16,7%	18,1%	26,4%
Horizon 2100	10,4%	10,4%	16,7%	16,7%	45,8%
Écart histo/2050	+0,5	-0,9	-1,2	-10,1	+11,9
Écart histo/2100	-12,7	-5,8	-1,2	-11,5	+31,3

Tableau 8 - Proportions d'ouragans générés par catégorie de Saffir-Simpson selon le GCM générateur (projections faites avec le scénario RCP 8.5)

Quel que soit le GCM utilisé, nous pouvons remarquer une tendance à la diminution des proportions d'ouragans faiblement puissants et à l'augmentation des proportions d'ouragans les plus puissants :



Note : l'axe des abscisses donne le centre de la fenêtre de 20 années sur laquelle sont calculées les deux proportions moyennes. Par exemple, l'abscisse « 1980 » donne les deux proportions moyennes sur la période 1970 – 1989.

Figure 56 - Proportions moyennes (sur 20 ans glissants) d'ouragans de catégories 1-2 (trait en pointillés) et 4-5 (trait plein) générés sous scénario RCP 8.5 (après 2005, conditions historiques utilisées avant 2005) selon le GCM générateur

Ce constat concorde avec les prévisions du GIEC. Remarquons que l'amplitude des variations de proportions est plus faible pour HadGEM2-ES alors que ce modèle est le plus sensible des trois au changement climatique.

Notons également qu'hormis pour la catégorie 5, aucune tendance n'est clairement observable dès 2050.

Ainsi, grâce à l'étude des projections faites dans le cadre du scénario RCP 8.5, nous pouvons constater que le changement climatique impacte le nombre annuel d'ouragans générés dans la deuxième partie du siècle ainsi que la puissance de ces derniers. Si les modèles sont unanimes sur la diminution des proportions d'ouragans faiblement puissants et l'augmentation des proportions d'ouragans les plus puissants générés, l'impact sur le nombre annuel d'ouragans est quant à lui moins tranché. Deux des trois modèles sont en accord avec les prévisions du GIEC, à savoir un nombre d'ouragans annuel stable ou en légère baisse, mais le modèle le plus sensible au changement climatique indique, lui, une augmentation légère de ce nombre passé 2050. Les modèles ne sont pas non plus unanimes quant à l'évolution de la fréquence à laquelle sont frappés les territoires de la caisse Antilles-Guyane.

V. Modèle catastrophe utilisé pour l'estimation des pertes financières

L'objectif de ce mémoire étant d'évaluer l'impact financier des ouragans observés sur les simulations issues du projet CORDEX, il faut pouvoir transformer cette activité cyclonique en pertes financières. Pour ce faire, nous allons utiliser un modèle catastrophe.

1. Modèles catastrophes

Les modèles catastrophes sont utilisés notamment par les assureurs et les réassureurs dans le but d'évaluer leur exposition au risque catastrophe et de le gérer. Ce risque peut provenir de périls très divers : aussi bien naturels comme les inondations, les ouragans ou les tremblements de terre, que man-made comme les cyber-attaques ou le terrorisme.

Les modèles catastrophes sont structurés en 4 modules différents :

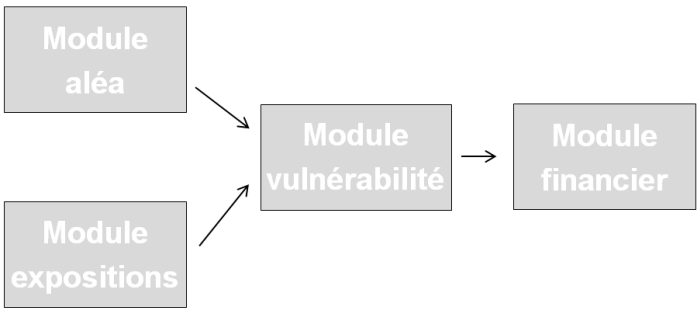


Figure 57 - Schéma de la structure d'un modèle catastrophe

Le rôle du module aléa est de simuler le péril en question et donc dans notre cas, les ouragans. La fréquence, la sévérité ainsi que la localisation de ces événements simulés sont déterminées à partir de lois de probabilité calibrées sur un historique. Le module expositions décrit quant à lui les expositions de l'assureur, c'est-à-dire les différents bâtiments assurés dans le cas du péril ouragan. La localisation, le type de bâtiment et le coût de reconstruction sont les principales caractéristiques des expositions.

Ces deux modules sont ensuite couplés grâce au module vulnérabilité. Ce dernier a pour rôle de déterminer les dommages causés par les événements simulés sur les expositions. En pratique, des courbes d'endommagement sont utilisées : elles relient sévérité du péril et taux de dommage pour le bâtiment considéré. La détermination de ces courbes est complexe puisqu'elles doivent tenir compte des caractéristiques des bâtiments.

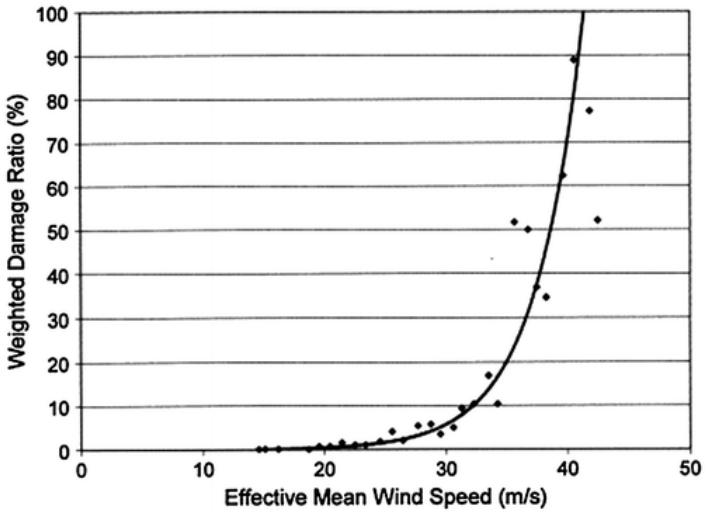


Figure 58 - Exemple de courbe de vulnérabilité établie avec le péril vent (Pita et al., 2014)

Enfin, le module financier a pour but de transformer les dommages constatés en pertes financières pour l'utilisateur du modèle.

Dans notre cas, les événements du module aléa sont ceux observés sur les simulations climatiques. Les expositions sont celles de la caisse Antilles-Guyane. En revanche, nous ne disposons pas des différentes courbes de vulnérabilité : nous allons donc utiliser un modèle catastrophe existant en y intégrant les ouragans détectés sur les simulations climatiques. La qualité du modèle catastrophe utilisé, et notamment la justesse des courbes de vulnérabilité utilisées, joue donc un rôle crucial dans l'estimation des pertes financières occasionnées par l'activité cyclonique observée sur les données issues du projet CORDEX.

Nous allons donc agir sur le module aléa d'un modèle catastrophe fourni par un modélisateur A sur la base des expositions de la caisse Antilles-Guyane au 31 décembre 2019. Ce modèle catastrophe nous est restitué sous la forme d'un catalogue d'événements : 10 000 années d'activité cyclonique ont été simulées et les ouragans qui en résultent composent le catalogue. L'empreinte de chaque ouragan y est décrite et nous disposons de la perte engendrée par l'événement :

		Latitude & longitude	Année de simulation	Jour et heure	Vitesse de vent maximale	Pression minimale au centre	Vitesse de déplacement	Perte nette engendrée
Ouragan i	Position 1							
	Position 2							
	Position 3							
	…							
	Position n_i							

Tableau 9 - Informations fournies par la table d'événements sur l'ouragan i (dont l'empreinte est constituée de n_i positions)

Précisons que pour un ouragan donné, nous disposons de la répartition de la perte nette engendrée sur chacun des 5 territoires de la caisse Antilles-Guyane. Cette perte est brute de réassurance, seules les franchises et limites des contrats ont été appliquées.

2. Étude de la table d'événements

Les ouragans se trouvant dans la table d'événements ci-dessus ont été générés par le modélisateur A en utilisant la base HurDat sur la période 1950-2007. Le nombre annuel d'ouragans, les empreintes, les vitesses de déplacement et les vitesses de vent générées sont déterminés de façon probabiliste à partir de l'historique. Nous ne détaillerons pas ici les méthodes utilisées.

Un détail est important concernant les événements générés : la table ne contient que des ouragans passant à proximité des expositions de la caisse Antilles-Guyane. Les ouragans dont l'empreinte est trop éloignée de nos expositions pour engendrer des pertes n'apparaissent pas dans la table. Notons également que les empreintes ont été tronquées : seule la partie proche de nos expositions a été conservée.

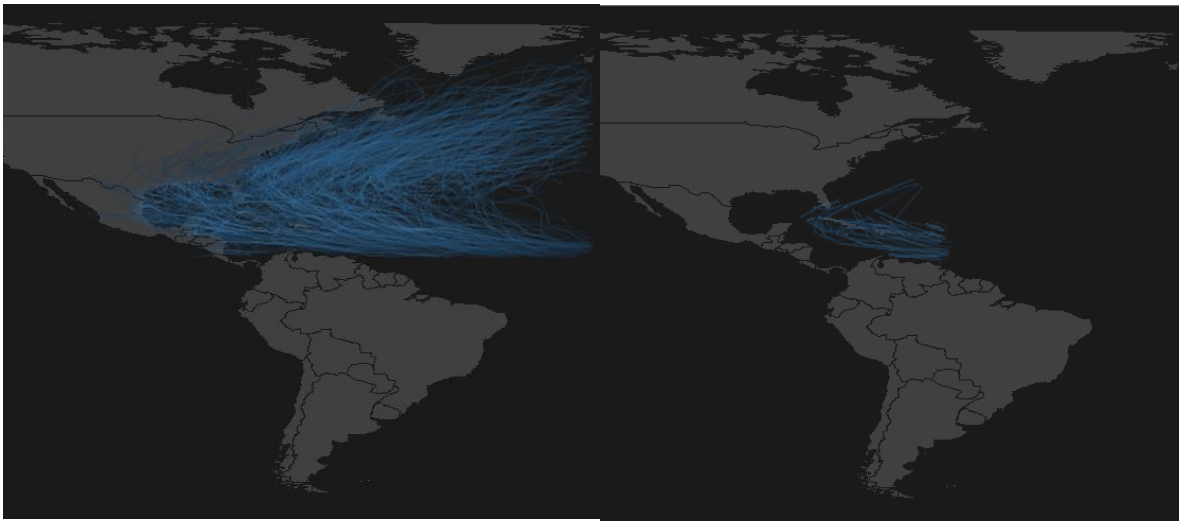


Figure 59 - Empreintes des ouragans historiques sur la période 1950-2007 (gauche) et empreintes d'une partie des ouragans de la table d'événements (droite)

Par conséquent, la fréquence annuelle des ouragans de la table correspond à une fréquence annuelle d'ouragans passant à proximité de nos expositions et non à une fréquence annuelle historique comme nous pouvons le voir ci-dessous :

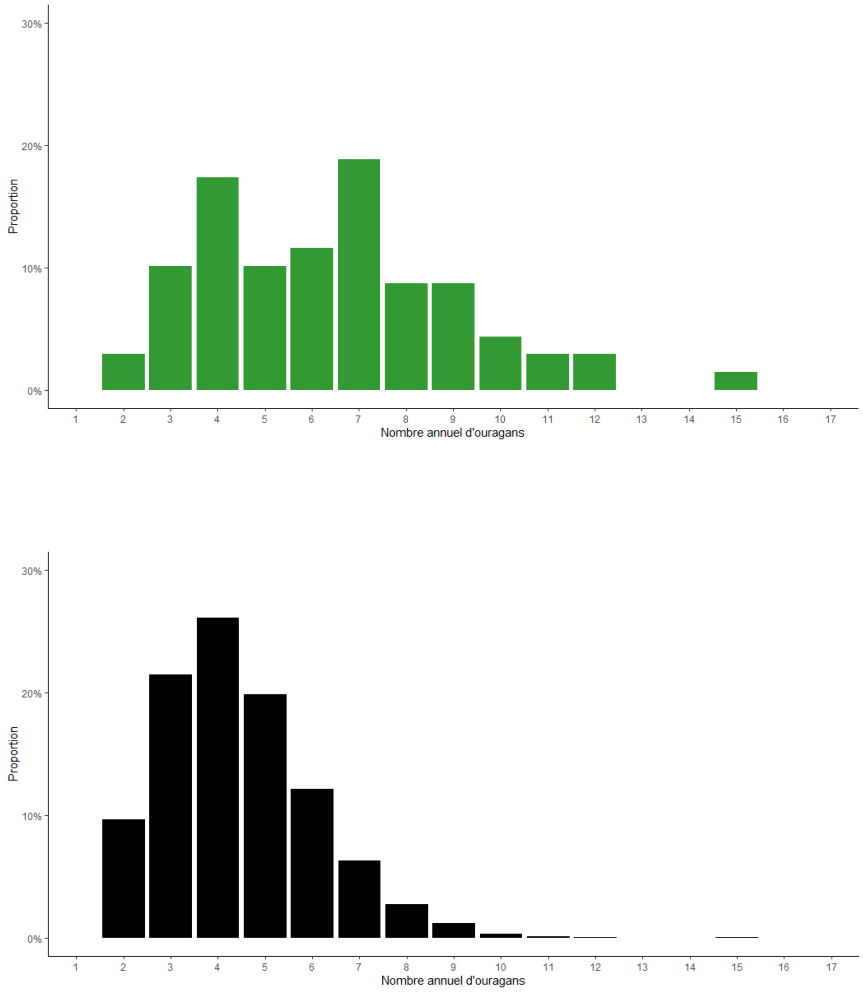


Figure 60 - Distribution du nombre annuel d'ouragans historiques sur la période 1950-2007 (haut) et d'ouragans simulés sur les 10 000 années de la table (bas)

Nous voyons bien sur la figure ci-dessus que la fréquence annuelle historique est plus élevée que la fréquence annuelle de la table d'événements.

Nous pouvons observer une proportion d'ouragans violents plus faible dans les ouragans de la table :

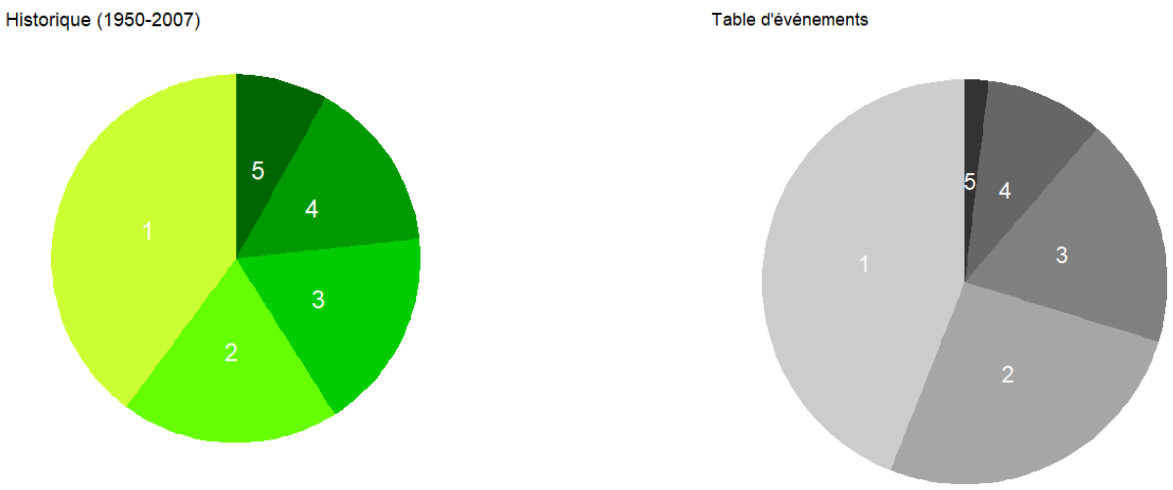


Figure 61 - Répartition selon les 5 catégories de l'échelle de Saffir-Simpson des ouragans historiques (gauche) et des ouragans simulés (droite)

Les ouragans de catégories 4 et 5 représentent respectivement 15% et 8% des ouragans historiques sur la période 1950-2007 alors qu'ils représentent respectivement 9% et 2% des ouragans de la table.

La perte nette engendrée par un ouragan de la table correspond aux dégâts causés par le vent mais aussi par les fortes précipitations. Notons que la perte causée par ces deux périls peut être amplifiée : dans le cas où un ouragan cause des dégâts très importants, la demande en main d'œuvre et en matières premières augmente brusquement. Les conséquences peuvent alors être multiples : hausse temporaire des prix ou coûts supplémentaires liés à l'achat et au stockage des matériaux. Aussi, l'offre ne peut immédiatement s'adapter à cette demande soudaine et des pertes supplémentaires sont alors causées par les délais imposés. Dans notre cas, les pertes données par le modélisateur intègrent cette amplification.

Si les spécificités propres à un territoire peuvent amplifier la perte causée par un ouragan sur ce même territoire, et donc agir à la fin du mécanisme d'estimation de la perte, elles peuvent aussi agir plus en amont : dès la valorisation des biens assurés.

- <u>Spécificités économiques et géographiques impactant la valorisation en Outre-Mer</u>

  Les caractéristiques géographiques ainsi que les spécificités économiques de l'Outre-Mer font augmenter le coût de construction, et donc la valorisation, des biens assurés sur ces territoires. Dans un rapport datant d'octobre 2018 relatif au marché des matériaux de construction dans les DOM de Mayotte et de la Réunion, l'Autorité de la Concurrence note que le prix des matériaux de construction est 39% plus élevé à la Réunion qu'en métropole, 35% à Mayotte, alors que le niveau général des prix sur ces territoires n'est que 7% plus élevé. Rappelons que les territoires de la caisse Antilles-Guyane comptant le plus d'expositions (à savoir Guadeloupe, Martinique et Guyane) sont, tout comme Mayotte et la Réunion, des DOM. Les territoires de Saint-Barthélemy et de Saint-Martin bénéficient quant à eux d'une plus grande autonomie et sont considérés comme des collectivités d'Outre-Mer (ou COM).

Bien que l'octroi de mer, une taxe à l'importation existant dans les DOM, fasse augmenter le prix des produits importés, une TVA bien plus basse qu'en France métropolitaine limite l'impact sur les prix pour le consommateur.

	TVA Métropole	TVA DOM excepté Mayotte et Guyane	TVA Guyane et Mayotte
Taux normal	20%	8,5%	TVA non applicable
Taux intermédiaire	10%	2,1%	
Taux réduit ou taux particulier	5,5% ou 2,1%	1,05% ou 1,75%	

Note : la TVA correspond à un pourcentage du prix de vente du produit

Tableau 10 – Taux de TVA en vigueur en France métropolitaine et dans les DOM (source : Direction Générale des Finances Publiques)

De plus, le régime de TVA en vigueur dans les DOM prévoit une exonération de TVA pour la plupart des matériaux de construction. La fiscalité n'explique donc pas à elle seule les coûts de construction plus élevés en Outre-Mer.

Dans son rapport, l'Autorité de la Concurrence pointe dans un premier temps des niveaux de marge commerciale « très importants » dans les prix proposés au consommateur pour la plupart des matériaux de construction. Ces marges seraient liées à une concurrence très faible pour la production des matériaux de construction. Le nombre réduit d'acteurs sur ces territoires résulte principalement d'un marché étroit (faible nombre de consommateurs potentiels), d'espaces de stockage très réduits ou encore de normes coûteuses à respecter.

Concernant les espaces de stockage, la taille réduite des territoires de l'Outre-Mer, leur relief accidenté ainsi qu'une volonté de préservation de l'environnement rend le stockage des matériaux et des machines coûteux.

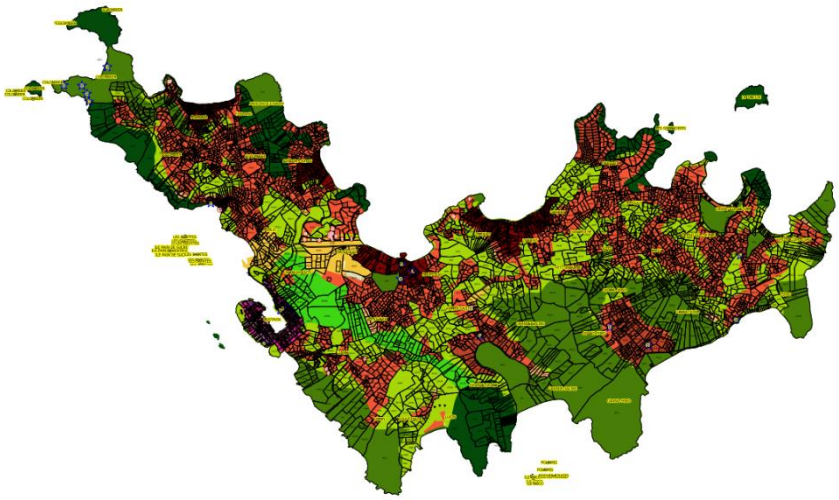


Figure 62 - Ile de Saint-Barthélemy et ses zones protégées visibles en vert (source : Collectivité de Saint-Barthélemy)

Nous pouvons voir ci-dessus que la majorité des terres de l'île de Saint-Barthélemy sont protégées : sur ces terres protégées, il est impossible de stocker des matériaux ou des machines.

Les normes jouent également un rôle important dans les coûts élevés des matériaux de construction. Contrairement à certaines certifications obligatoires comme la certification « CE », les normes sont d'application facultative. Cependant, respecter les normes d'usage dans le secteur du bâtiment s'avère être incontournable : elles permettent de garantir un niveau de qualité et de sécurité mais aussi de ne pas invalider les contrats d'assurance dommages-ouvrage réclamant, généralement, que les matériaux utilisés soient conformes aux normes. Les organismes vérifiant la conformité des matériaux étant majoritairement situés en Europe, les matériaux importés transitent systématiquement par le continent. Là où en métropole les surcoûts engendrés par cette pratique sont moindres, ils deviennent beaucoup plus importants en Outre-Mer. De plus, la production locale de matériaux de construction s'en trouve freinée : les coûts liés au transit des matériaux par l'Europe pour vérifier leur conformité et ainsi s'insérer sur le marché des matériaux de construction découragent toute initiative locale.

Ces spécificités propres à l'Outre-Mer impactent donc la valorisation des biens de la caisse Antilles-Guyane. Les méthodes de valorisation développées par GMA ne seront pas détaillées ici mais elles ont été adaptées pour tenir compte de ces surcoûts.

Comme nous pouvons nous en douter, les pertes nettes varient grandement suivant les catégories de Saffir-Simpson des ouragans les ayant causées :

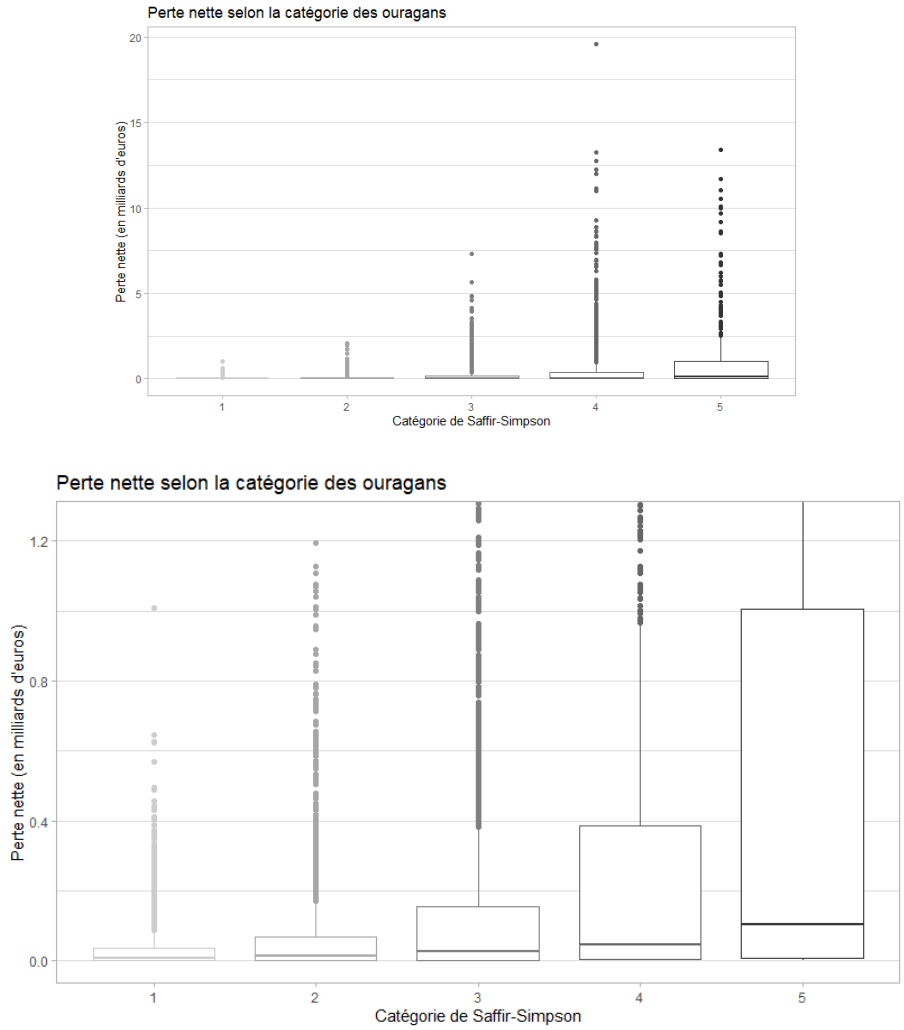


Figure 63 - Boîtes à moustaches créées sur les séries de pertes nettes par événement (pertes sur l'ensemble du portefeuille) selon la catégorie de l'événement (haut) et agrandissement de la partie basse du même graphique (bas)

Dans un premier temps, nous pouvons fort heureusement remarquer que tous les montants de perte sont inférieurs à l'engagement total de la caisse Antilles-Guyane. Aussi, nous pouvons voir que les 5 distributions représentées sont très asymétriques : les queues de distribution associées aux valeurs élevées sont particulièrement longues et cette longueur augmente avec l'intensité de l'ouragan considéré.

Notons cependant que bien qu'un ouragan soit particulièrement intense, de catégories 4 ou 5 par exemple, ce dernier peut, dans certains cas, engendrer une perte nette inférieure à certaines pertes associées à des ouragans de première catégorie. Ceci provient du fait qu'un ouragan n'atteint pas forcément son pic d'intensité proche des expositions de la caisse Antilles-Guyane. Nos expositions se trouvant dans le Sud-Ouest de la mer des Caraïbes, il est très probable que certains ouragans de catégorie élevée frappent peu de temps après leur genèse : les vitesses de vent générées sont alors plus faibles.

3. Évaluation des pertes proposées par la table

Nous allons nous appuyer sur la table du modélisateur A pour transformer les ouragans observés en pertes financières. Il est donc important de comparer les pertes financières données par cette table aux pertes historiques afin de déterminer si la table sous-estime les pertes auxquelles le Groupe est exposé ou si elle les surestime. Introduisons d'abord les notions d'AAL, de courbe AEP et de courbe OEP afin d'étudier les pertes proposées par le modélisateur.

- AAL

  L'AAL ou Annual Average Loss est la perte annuelle moyenne. En notant $X_1, X_2, \dots, X_n$ les variables aléatoires décrivant les pertes annuelles de n années et en les considérant comme indépendantes et identiquement distribuées, on a :

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

  L'AAL est donc la moyenne empirique de la série des n pertes annuelles. Par la loi des grands nombres, c'est un estimateur convergent de l'espérance de perte annuelle.

- Courbe AEP et OEP

  Les courbes AEP (Aggregate Excedance Probability) et OEP (Occurrence Excedence Probability) sont couramment utilisées pour donner un aperçu des pertes financières issues d'un modèle catastrophe. Ces deux courbes associent une période de retour avec un niveau de sinistralité : ce niveau de sinistralité est une perte annuelle totale dans le cas de la courbe AEP et une perte annuelle maximale par événement dans le cas de la courbe OEP.

  La période de retour peut s'interpréter de cette façon : une période de retour de n années pour un phénomène donné signifie que ce phénomène se produit en moyenne une fois sur n années consécutives. Le montant de perte associé à une période de retour de n années est le quantile d'ordre $1 - \frac{1}{n}$ de la série des pertes (pertes annuelles agrégées dans le cas de la courbe AEP, pertes annuelles maximales dans le cas de la courbe OEP). Un montant de perte équivalent ou supérieur est donc observable, en moyenne, une fois toutes les n années.

Les trois pertes historiques dont nous disposons sont les plus importantes enregistrées depuis la création de l'ancêtre de la caisse Antilles-Guyane au début des années 1990. Elles sont associées à Irma (2017), Dean (2007) et à Maria (2017), ouragans de catégorie 5, ayant causé respectivement 705 millions d'euros, 179 millions d'euros et 42 millions d'euros de pertes sur notre portefeuille. Ces pertes correspondent à des pertes as-if à la date du 31/12/2019 : la mise en as-if est faite grâce au nombre de risque des années 2007 et 2017. L'inflation est également prise en compte en utilisant l'indice FFB, indice produit par la Fédération Française du Bâtiment donnant le coût de reconstruction moyen d'une pièce en France métropolitaine. Les pertes sont brutes de réassurance et seules les franchises et limites des contrats ont été appliquées.

Voyons avec quelle fréquence la table du modélisateur génère des pertes équivalentes ou supérieures aux 3 montants de perte présentés :

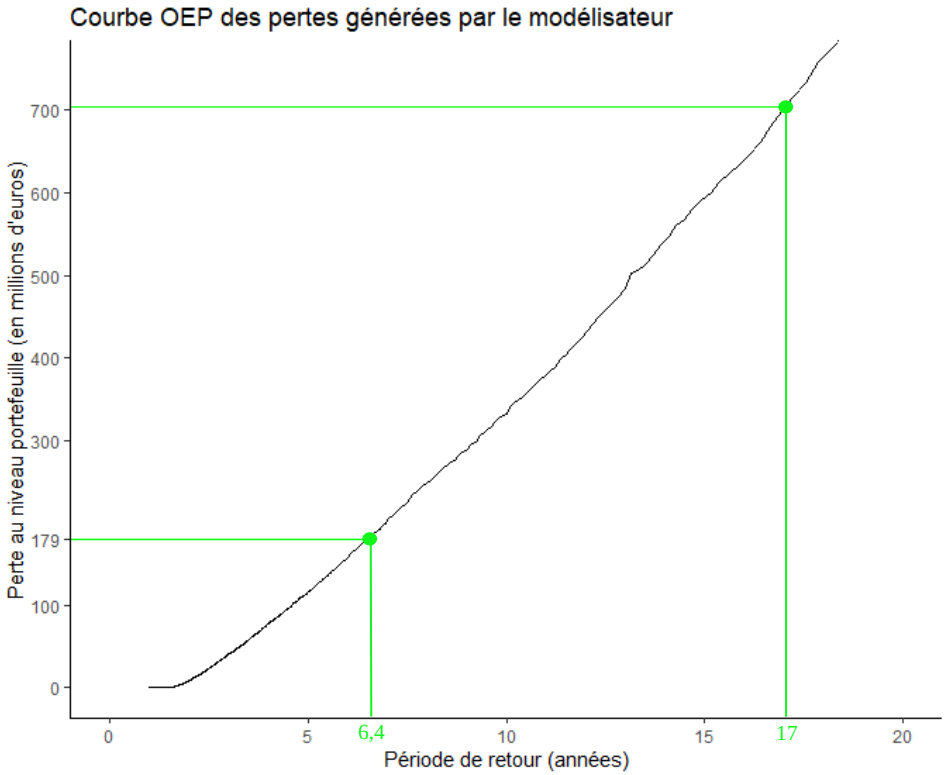


Figure 64 – Courbe OEP calculée sur la série des pertes par événement (événements de la table)

Nous pouvons voir sur le graphique ci-dessus que la perte correspondant à l'événement Dean, à savoir 179 millions d'euros, est associée à une période de retour de 6,4 années : un événement générant une perte équivalente ou supérieure arrive donc, selon le modélisateur, en moyenne une fois tous les 6 ans environ. Bien que l'historique dont nous disposons ne soit que très peu profond et qu'il soit donc difficile d'estimer historiquement une période de retour, la période de retour de 6 ans proposée par le modélisateur est bien trop courte au regard des pertes observées en 30 ans d'historique. La période de retour proposée par le modélisateur pour l'événement Irma est, quant à elle, de 17 ans. Étant donnée la faible profondeur de l'historique, il est difficile de se prononcer sur l'adéquation de cette période de retour avec la sinistralité historique.

Le modèle catastrophe proposé par le modélisateur A surestime donc les pertes subies par la caisse Antilles-Guyane. Voyons maintenant comment coupler ce modèle avec les projections climatiques du projet CORDEX afin d'estimer l'impact du réchauffement climatique sur le risque cyclonique dans les Antilles.

VI. Estimation des pertes financières

Nous allons coupler la table d'événements avec les caractéristiques des événements observés sur les données du projet CORDEX. Différentes approches sont alors possibles et elles se

différencient principalement par l'importance accordée aux données issues de nos simulations climatiques : certaines ne se s'appuient que partiellement sur ces simulations alors que d'autres reposent davantage dessus. Dans ce deuxième cas, la qualité des données extraites des simulations du projet CORDEX joue un rôle crucial. Précisons que dans les deux cas, l'homogénéité entre données de la table d'événements et données issues du projet CORDEX jouent également un rôle important.

Comme indiqué précédemment, la qualité du modèle catastrophe utilisé, ici proposé par le modélisateur A, est déterminante dans la justesse des pertes produites par les deux méthodes présentées ci-dessous.

Voyons dans un premier temps comment coupler cette table d'événements avec nos projections climatiques en utilisant au maximum les informations présentes dans les projections climatiques du projet CORDEX.

1. Méthode de recherche d'ouragans voisins

a) Description générale de la méthode

Cette première méthode de transformations des ouragans observés en pertes financières pour le Groupe s'appuie fortement sur les caractéristiques des ouragans identifiés dans les projections climatiques du projet CORDEX. L'ensemble des informations que nous avons pu obtenir sur ces projections est utilisé.

Cette méthode consiste à associer un ouragan observé dans nos projections climatiques avec des ouragans semblables provenant de la table d'événements. Cette association se fait territoire par territoire : la perte engendrée sur un territoire est obtenue grâce aux pertes engendrées par les ouragans de la table qui ont été associés. Cette méthode s'inspire de l'algorithme des k plus proches voisins (détaillé dans les annexes).

Pour un ouragan observé sur les projections climatiques, nous disposons des informations suivantes en chaque position de l'empreinte :

- Latitude
- Longitude
- Vitesse maximale des vents générés
- Vitesse de déplacement de l'ouragan

En procédant territoire par territoire, nous pouvons remplacer les variables latitude et longitude par une seule variable indiquant la distance de l'ouragan au territoire. En faisant ce remplacement, nous faisons une hypothèse importante : deux ouragans de même intensité situés à égale distance de l'île engendrent les mêmes pertes, peu importe leur localisation précise. Faire cette hypothèse permet de réduire la dimension du problème et donc de faciliter la recherche des ouragans voisins. Étant données la taille réduite des territoires de la caisse Antilles-Guyane par rapport à la taille d'un ouragan et la présence uniforme de Groupama sur

ces territoires, il semble raisonnable de considérer que les dégâts générés par deux ouragans d'intensités semblables situés à égale distance d'un territoire sont semblables.

En considérant un territoire i, un ouragan peut être vu comme un nuage de points dans $\mathbb{R}^3$. Si l'empreinte de l'ouragan est composée de n positions, alors le nuage de points est composé de n points. Chaque point est décrit par 3 coordonnées : distance au territoire i, vitesse maximale de vent et vitesse de déplacement. La caisse Antilles-Guyane étant présente sur 5 territoires, un ouragan est alors décrit par 5 nuages de points dans $\mathbb{R}^3$.

Nous allons ensuite associer à chacun de ces 5 nuages, des nuages voisins issus de la table d'événements. Nous observons le comportement de l'ouragan par rapport à un territoire et nous trouvons des ouragans issus de la table ayant le même comportement vis-à-vis de ce territoire.

Si nous considérons le nuage associé au territoire i, la perte moyenne issue des nuages voisins est alors considérée comme la perte causée par l'ouragan au territoire i. Les 5 pertes obtenues sont ensuite sommées pour obtenir la perte engendrée par l'ouragan sur l'ensemble des 5 territoires.

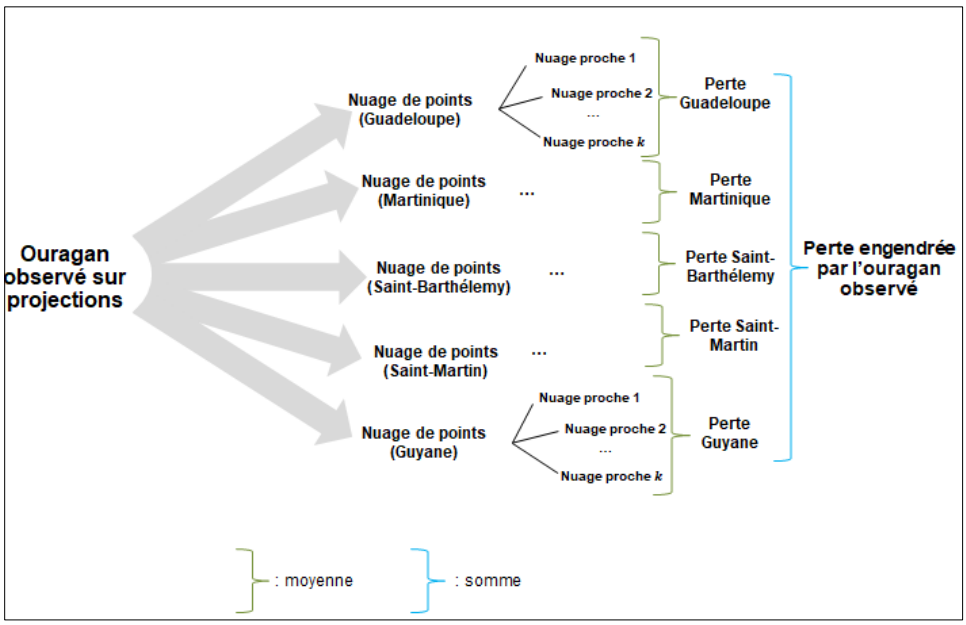


Figure 65 - Schéma de calcul de la perte nette engendrée par un ouragan observé sur les données issues du projet CORDEX avec la méthode de recherche d'ouragans voisins

Notons que depuis la création de la caisse Antilles-Guyane, les expositions situées en Guyane n'ont jamais été frappées par un ouragan. Nous pouvons effectivement remarquer en figure 39 que les ouragans historiques ne passent pas à proximité de la Guyane. Il en va de même pour les ouragans extraits des projections climatiques ainsi que pour ceux présents dans la table du modélisateur. La probabilité de survenance d'un événement dans cette zone étant très faible, la perte causée par un ouragan en Guyane est donc toujours considérée comme nulle.

Le nombre k de nuages proches cherchés doit être suffisamment élevé pour pouvoir tenir compte du caractère incertain de la perte : deux ouragans semblables peuvent en effet causer des montants de perte légèrement différents. Il ne doit cependant pas être trop élevé car nous

souhaitons que les nuages proches soient suffisamment proches du nuage initial. Si ces derniers sont trop éloignés, les pertes qui leur sont associées correspondent à des ouragans se comportant différemment vis-à-vis du territoire considéré. La perte moyenne calculée sur les nuages proches est alors faussée. Le nombre k est alors choisi égal à 3.

b) Recherche des nuages proches

La démarche est la même pour chaque territoire mais pour l'exemple, considérons le territoire de Saint-Martin.

Observer l'ouragan issu des projections climatiques nous a permis de créer un nuage de points associé à Saint-Martin. Ce nuage décrit le comportement de l'ouragan par rapport à cette île et repose sur 3 variables : la distance à l'île, la vitesse maximale des vents générés et la vitesse de déplacement.

Une fois le nuage de points dédié à Saint-Martin créé, nous devons chercher les nuages proches. Ces nuages correspondent à des ouragans de la table du modélisateur ayant un comportement similaire à proximité de l'île. Cette recherche se fait à l'aide du barycentre de chaque nuage.

- Notion de barycentre

  Soient $A_1, A_2, \dots, A_n$ des points de $\mathbb{R}^d$ et $a_1, a_2, \dots, a_n$ des nombres réels de somme non nulle. Le barycentre des points $A_1, A_2, \dots, A_n$ affectés des coefficients $a_1, a_2, \dots, a_n$ est l'unique point G de $\mathbb{R}^d$ tel que :

  $$\sum_{i=1}^{n} a_i \overrightarrow{GA_i} = \vec{0}$$

  L'expression « centre de gravité » peut aussi être employée pour faire référence à la notion de barycentre. En effet, supposons par exemple que les points $A_1, A_2, \dots, A_n$ possèdent chacun une masse (au sens physique du terme) et que les coefficients $a_1, a_2, \dots, a_n$ soient égaux à ces masses. Le barycentre associé correspond alors au centre de gravité (au sens physique du terme) du nuage de points.

  Dans le cas où les coefficients $a_1, a_2, \dots, a_n$ sont égaux, on parle d'isobarycentre.

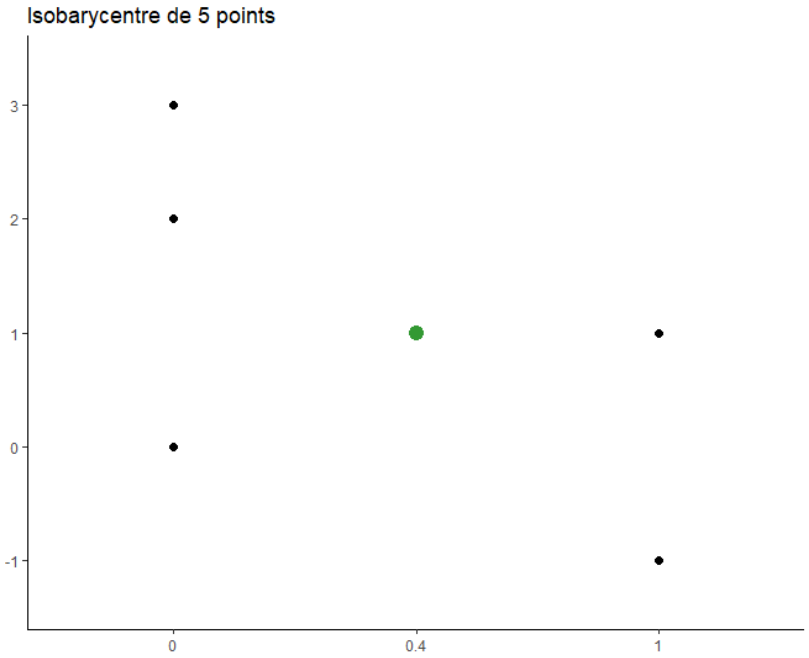


Figure 66 - Isobarycentre (en vert) de 5 points (en noir) dans $\mathbb{R}^2$

Dans notre cas, $d = 3$ et les points $A_1, A_2, \dots, A_n$ sont ceux composant les nuages de points. Nous allons donc chercher les nuages proches du nuage associé à Saint-Martin *via* le barycentre de ces nuages. Trouver les 3 nuages proches va donc consister à chercher les 3 barycentres les plus proches de celui du nuage associé à Saint-Martin. La distance utilisée est celle associée à la norme euclidienne sur $\mathbb{R}^3$.

Précisons que les barycentres de nos nuages de points ne sont pas des isobarycentres. En effet, affecter le même poids à tous les points signifierait accorder la même importance à chaque position de l'ouragan associé au nuage. Or, il est nécessaire d'accorder plus d'importance aux points dont la distance à l'île est faible par rapport aux autres points. Les poids du barycentre doivent refléter la contribution de chaque position à la perte engendrée. Si par exemple tous les points ont la même distance à l'île, les poids associés à ces points sont égaux, peu importe la valeur de la distance. Ainsi, seule la contribution de chaque distance à la somme des distances du nuage compte.

Considérons un point A_i (avec $1 \leq i \leq n$) du nuage de points $(A_1, \dots, A_n)$ et a_i son poids. Notons d_i la coordonnée de distance au territoire du point A_i et $c_i := \frac{d_i}{\sum_j d_j}$ la contribution de la distance d_i à la somme des distances du nuage de points. Nous définissons le poids du point A_i de la façon suivante :

$$a_i := \frac{\left(\frac{1}{c_i}\right)^2}{\sum_j \left(\frac{1}{c_j}\right)^2}$$

Plus la contribution du point A_i est faible dans la somme des distances, plus il est proche de l'île par rapport aux autres points et plus le poids doit être important, d'où l'inverse de la

contribution c_i. Cette contribution est mise au carré pour renforcer l'importance des points à distance d_i faible par rapport aux autres distances du nuage. Diviser par la somme de ces fractions sur l'ensemble des points du nuage permet ensuite de normaliser le poids.

Une fois les barycentres construits, la recherche des nuages proches peut être effectuée. Prenons l'exemple d'un ouragan généré par le modèle climatique NorESM et voyons quels ouragans de la table lui ont été associés lorsque que nous nous concentrons sur le territoire de la Guadeloupe :

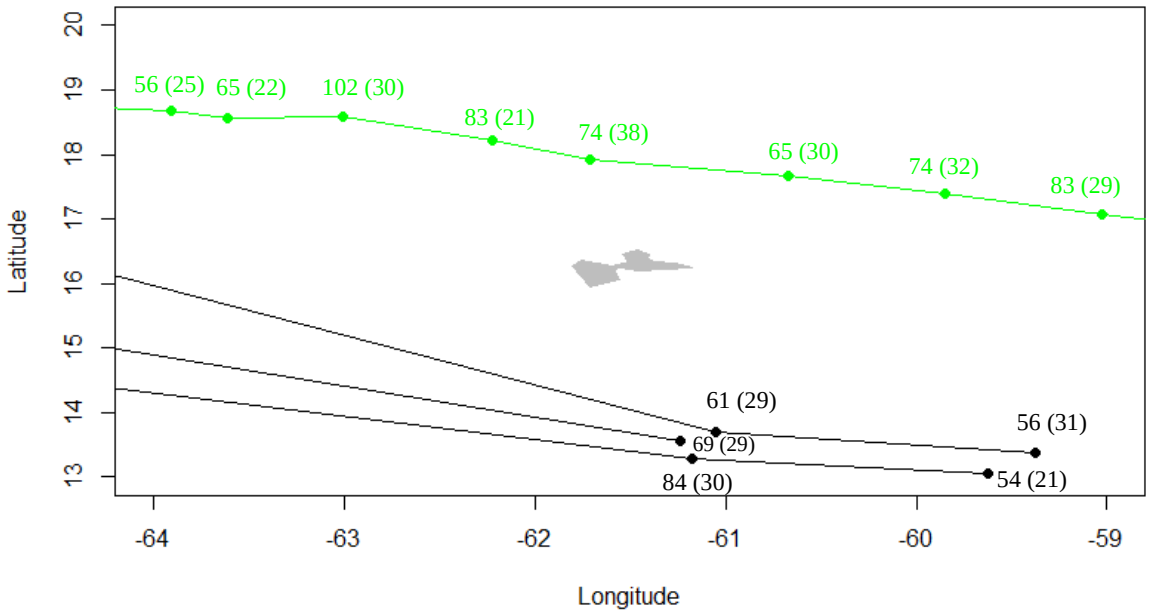


Figure 67 - Empreinte de l'ouragan issu du modèle (vert), empreintes des 3 ouragans de la table qui lui ont été associés (noir) avec vitesses de vent maximale en chaque position (km/h), vitesse de déplacement de l'ouragan (entre parenthèses, en km/h) et île de Guadeloupe (gris)

Nous pouvons voir sur ce schéma que les ouragans correspondant aux nuages proches pour le territoire de la Guadeloupe sont situés à la même distance de l'île que l'ouragan issu des projections climatiques. Les vitesses de vent ainsi que les vitesses de déplacement sont également semblables. Dans ce cas précis, le comportement de l'ouragan simulé par NorESM vis-à-vis de la Guadeloupe est donc très semblable au comportement des 3 ouragans de la table qui lui ont été associés. Les vitesses de vent étant ici très faibles, les pertes engendrées par les ouragans de la table sur la Guadeloupe sont nulles. La perte engendrée par l'ouragan simulé par le modèle climatique est donc considérée comme nulle également.

c) Étude des pertes générées

- HadGEM2-ES et NorESM : inexploitables avec cette méthode d'estimation des pertes
  La perte annuelle engendrée par ces deux modèles est très faible voire nulle :

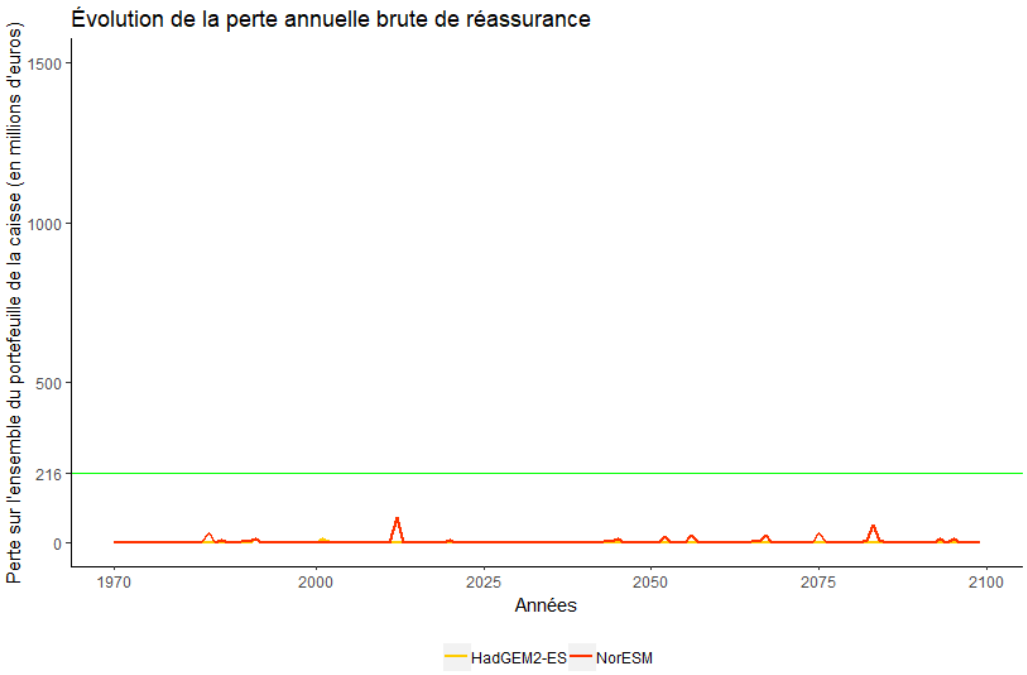


Figure 68 - Évolution des pertes annuelles agrégées (brutes de réassurance) causées par des ouragans sur l'ensemble du portefeuille de la caisse par GCM générateur (conditions historiques utilisées avant 2005, scénario RCP 2.6 après) et comparaison avec l'AAL proposée par le modélisateur (216M€, en vert)

Si les pertes engendrés par le modèle norvégien sont rarement non-nulles, celles engendrées par HadGEM2-ES ne le sont jamais. Bien que ces pertes soient inexploitables, tâchons de comprendre comment elles ont pu être obtenues.

Les ouragans générés par ces deux modèles ont été associées à des ouragans donnés par le modélisateur engendrant des pertes très souvent nulles. Il y a alors deux explications possibles : soit les ouragans générés par ces deux GCMs causeraient effectivement des dégâts mais les ouragans associés n'en causent pas (les nuages proches trouvés pour l'ouragan sont trop éloignés), soit les ouragans générés ne causeraient aucun dégât et sont associés à des ouragans n'en générant pas non plus.

Voyons d'abord si les ouragans générés par les deux modèles climatiques sont bien associés à des ouragans semblables issus de la table du modélisateur A. Rappelons que pour chaque ouragan créé par un de nos modèles, 5 barycentres sont créés (un pour chaque territoire de la caisse Antilles-Guyane). A chacun de ces 5 barycentres sont associés 3 barycentres : ceux des nuages proches. Ainsi, 15 associations de barycentres sont faites pour un seul ouragan généré. Considérons ces associations pour l'ensemble des ouragans générés et analysons les séries d'écarts absolus entre coordonnées des barycentres associés deux à deux. La méthode pour générer les 3 séries est la suivante :

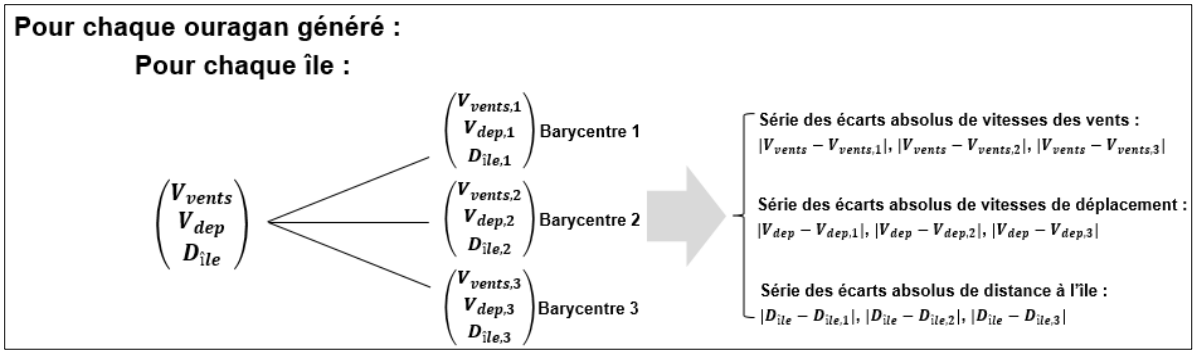


Figure 69 - Schéma détaillant la méthode utilisée pour créer chacune des trois séries servant à évaluer la qualité des associations deux à deux des barycentres (pour un GCM donné)

Si nous générons ces trois séries d'écarts absolus pour le modèle HadGEM2-ES, celui donnant lieu à des pertes annuelles nulles, nous obtenons les quantiles 80% suivants pour chacune des trois séries :

	Écarts absolus de vitesses des vents	Écarts absolus de vitesses de déplacement	Écarts absolus de distance aux îles
Quantile 80%	19,4 km/h	7,4 km/h	10,3 km

Tableau 11 – Quantile 80% des séries d'écarts absolus entre coordonnées des barycentres associés deux à deux (GCM : HadGEM2-ES)

Ainsi, dans 80% des associations entre ouragan généré par le modèle et ouragan issu de la table du modélisateur, l'écart de vitesse de vent entre les deux barycentres est au maximum de 19,4 km/h, ce qui est faible au regard des vitesses de vent engendrées par un ouragan. Cet écart est plus conséquent en ce qui concerne les vitesses de déplacement : selon la figure 26, les vitesses de déplacement des ouragans dans les Antilles varient entre 17 et 30 km/h. Une variation de 7,4 km/h est alors importante mais gardons à l'esprit que la majorité des écarts absolus sont inférieurs à ce seuil.

La dernière série, celle des distances à l'île, est la plus importante puisque les poids en jeux dans le calcul des barycentres sont obtenus à partir de cette variable. Nous pouvons constater que le quantile 80% est extrêmement faible : dans 80% des associations deux à deux de barycentres, l'écart entre les coordonnées de distance à l'île ne dépasse pas 10 km alors même que les distances en question sont de l'ordre de plusieurs centaines de kilomètres.

Par conséquent, les ouragans générés par le modèle HadGEM2-ES sont bien associés à des ouragans semblables issus de la table proposée par le modélisateur et ce peu importe l'île considérée. Nous arrivons à la même conclusion pour le modèle NorESM, générant lui aussi des pertes annuelles très faibles voire nulles.

La table du modélisateur contenant suffisamment d'événements pour que certains d'entre eux soient semblables à ceux détectés sur les données issues du projet CORDEX, il apparaît donc que les caractéristiques des ouragans détectés laissent penser qu'ils n'engendrent aucune perte.

L'étude de la série de coordonnées « Distance à l'île » des barycentres liés aux ouragans générés par HadGEM2-ES nous permet de voir que les ouragans générés passent bien trop loin des territoires de la caisse :

	Série de coordonnées « Distance à la Guadeloupe »	Série de coordonnées « Distance à la Martinique »	Série de coordonnées « Distance à Saint-Martin »	Série de coordonnées « Distance à Saint-Barthélemy »
Quantile 5%	1235 km	1421 km	935 km	975 km

Tableau 12 - Quantile 5% des séries de coordonnées « Distance à l'île » des barycentres liés aux ouragans générés par HadGEM2-ES

La taille d'un ouragan variant entre 500 et 1000 km selon Météo-France, l'étude de la coordonnée « Distance à l'île » des barycentres nous indique clairement que les ouragans sont bien trop éloignés des côtes des îles contenant les expositions de la caisse Antilles-Guyane. Ils ne peuvent donc engendrer des pertes.
Si cet éloignement est la cause des pertes nulles engendrées par le modèle anglais, nous avons pu voir précédemment, en figure 47, que les ouragans générés par NorESM passent bien plus près des côtes. Les pertes engendrées sont pourtant très faibles : nous aurions pu nous attendre à des pertes non négligeables, particulièrement à l'horizon 2100 comme l'illustre la figure 47.
Cette divergence illustre bien la complexité d'une question simple au premier abord : savoir si un ouragan simulé frappe ou non un territoire. Il est possible de répondre de façon simple et rapide, en utilisant peu de données, comme cela a été fait pour créer la figure 47. Nous voyons cependant ici que cette approche produit parfois des résultats trompeurs et que faire appel à des connaissances plus poussées, *via* un modèle catastrophe, amène à des résultats bien différents.

Tous ces constats, réalisés sur des projections faites dans le cadre du scénario RCP 2.6, restent valables dans le cadre du scénario RCP 8.5.

- <u>Modèle MPI-ESM</u>

  Si le modèle MPI-ESM génère bien plus d'ouragans que les modèles anglais et norvégien, la proportion d'ouragans causant des pertes est elle aussi plus importante. Les pertes annuelles engendrées sont donc plus réalistes, aussi bien en scénario RCP 2.6 qu'en scénario RCP 8.5 :

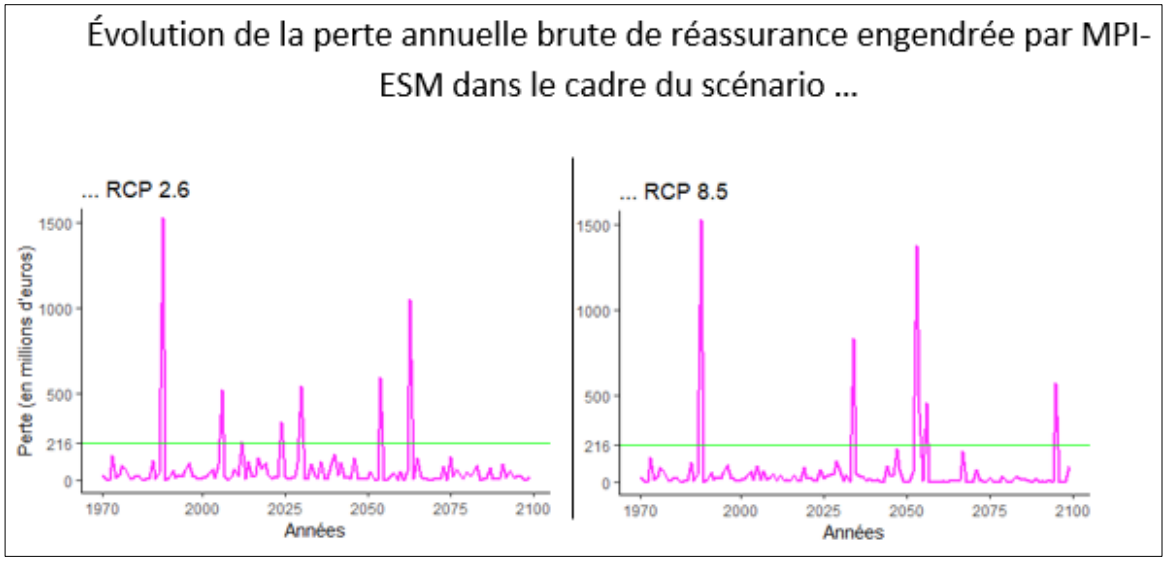


Figure 70 - Évolution de la perte annuelle au niveau portefeuille brute de réassurance engendrée par MPI-ESM, conditions historiques utilisées avant 2005 et RCP 2.6 (gauche) ou RCP 8.5 (droite) utilisé après. L'AAL donnée par le modélisateur (216M€) apparaît en vert.

Nous pouvons voir ci-dessus que le scénario RCP 8.5 accentue l'ampleur des pertes extrêmes alors même que le nombre annuel d'ouragans diminue au cours du XXI^{ème} siècle, comme vu en figure 54. A l'inverse, les pertes plus faibles, inférieures à l'AAL proposée par le modélisateur, sont réduites. Tout ceci est explicable à l'aide du tableau 8 : moins d'ouragans de faible intensité et davantage d'ouragans puissants sont générés par le modèle. Ainsi, lorsque les empreintes produites par le modèle sont exploitables, le lien entre des événements davantage extrêmes et une sinistralité aux extrêmes accentués peut être observé en utilisant la méthode de recherche d'ouragans voisins développée ici.

Cette méthode de transformation des ouragans observés en pertes financières présente l'avantage d'exploiter totalement les données disponibles. Cependant, reposer en grande partie sur les données peut être risqué si leur qualité n'est pas garantie, comme nous l'avons vu à travers les modèles HadGEM2-ES et NorESM. Si le modèle allemand génère quant à lui des résultats interprétables permettant d'observer une tendance, ils ne sont cependant pas exploitables en pratique. Comme vu précédemment, le projet CORDEX ne rend accessible qu'une seule simulation du XXI^{ème} siècle, chaque année n'est simulée qu'une seule fois. Bien que le *run* rendu disponible soit considéré comme central, aucune inférence statistique n'est possible : le nombre d'années simulées est bien trop faible pour permettre le calcul de périodes de retour robustes.

2. Méthode de resimulation à partir de la table

Contrairement à la méthode de recherche d'ouragans voisins, cette seconde méthode repose peu sur les données issues du projet CORDEX : seules les vitesses de vent des ouragans générés sont utilisées. Nous espérons ainsi obtenir des résultats plus vraisemblables en évitant les

problèmes liés aux données issues du projet CORDEX. Cette seconde méthode présente également l'avantage de produire des périodes de retour robustes.

Nous avons pu voir précédemment que le nombre annuel de cyclones ne semblait pas suivre une tendance haussière ou baissière au cours du XXI^{ème} siècle. Pour cette raison, nous allons nous concentrer sur les vitesses de vents générées par les simulations. Plus précisément, nous allons exploiter les proportions d'ouragans de catégories 1 à 5 générés.

Chaque année, un nombre aléatoire d'ouragans, noté N, est simulé. La loi de probabilité suivie par N est calibrée sur l'historique, c'est-à-dire la base HurDat : notons cette loi $\mathcal{L}$. Ensuite, pour chacun des N ouragans, sa catégorie est déterminée à partir des proportions observées dans les données. Supposons par exemple que nous souhaitions simuler l'activité cyclonique issue du modèle MPI-ESM à l'horizon 2100. Si la proportion d'ouragans de catégorie 1 est de 23% à l'horizon 2100, nous allons simuler, en moyenne, 23% d'ouragans de cette catégorie.

Il est important de noter que la loi $\mathcal{L}$ est choisie constante tout au long du siècle. En scénario RCP 2.6, nous avons pu voir que le nombre annuel d'ouragans générés dans le bassin est stable. En scénario RCP 8.5, nous avons pu constater le caractère incertain de l'impact du réchauffement climatique sur ce nombre. Rappelons cependant que le GIEC juge *probable* que ce nombre diminue ou reste stable au niveau mondial.

Voici le pseudo code de la démarche dans le cas où nous souhaitons simuler une activité cyclonique dont la proportion d'ouragans dans les catégories 1 à 5 est donnée par $p = (p_1, p_2, p_3, p_4, p_5)$:

Pour chacune des 10 000 années simulées :

    Tirage de N (N suit une loi $\mathcal{L}$)

    Pour chacun des N ouragans de l'année :

        Tirage de U (U suit une loi Uniforme sur $[0,1]$)

        Transformation du tirage de U en une catégorie (entier de 1 à 5) notée C à partir du vecteur p

        Tirage d'un ouragan de catégorie C dans la table d'événements

        Ajout de la perte associée à l'ouragan tiré à la perte annuelle

La transformation du tirage de U, noté u, en catégorie à partir de p se fait de la façon suivante : si $u \leq p_1$, alors l'ouragan est de catégorie 1, si $p_1 \leq u \leq p_1 + p_2$ alors l'ouragan est de catégorie 2 et ainsi de suite.

a) Loi de fréquence $\mathcal{L}$

La loi $\mathcal{L}$ donne le nombre d'ouragans tirés chaque année et est calibrée sur l'historique, à savoir la période 1950-2007, également utilisée par le modélisateur A pour obtenir la table d'événements.

Comme nous l'avons vu précédemment, la table d'événements du modélisateur ne contient que des ouragans passant proche des expositions de la caisse Antilles-Guyane. Par conséquent, la proportion d'ouragans frappant les expositions de la caisse est bien plus importante dans la table qu'historiquement. Ainsi, calibrer une loi de fréquence sur l'historique brut et tirer des ouragans dans la table du modélisateur aboutit à des montants de pertes annuelles aberrants : la perte annuelle moyenne est d'environ 1,3 milliards d'euros alors que la plus grosse perte annuelle enregistrée pour la caisse ne dépasse pas 800 millions d'euros. Apparaît alors la nécessité de ne travailler que sur les ouragans frappant nos expositions : la loi $\mathcal{L}$ va être calibrée sur un historique retraité et les ouragans seront tirés parmi l'ensemble des ouragans de la table causant des pertes.

L'historique retraité ne doit contenir que des ouragans ayant touché les territoires dépendant de la caisse Antilles-Guyane. L'historique remontant jusqu'en 1950, la documentation interne à Groupama ne permet pas de déterminer quels sont les ouragans ayant touché les côtes des territoires dépendant de la caisse. Certains documents disponibles en ligne donnent une liste d'ouragans ayant frappé ces territoires mais la liste n'est pas exhaustive, elle ne contient en général que les phénomènes ayant causé le plus de dégâts. Nous allons donc réutiliser ici la distance de 132 km calculée précédemment à partir des radius du cercle de vitesse 33 m/s : sur la période 1950-2004, nous considérerons qu'un territoire est frappé par un ouragan si ce territoire est situé à moins de 132 km du centre de l'ouragan. Sur la période 2004-2007, nous utiliserons les radius renseignés dans la base historique HurDat.

La fréquence à laquelle les ouragans frappent les territoires de la caisse Antilles-Guyane est alors distribuée de cette façon sur la période 1950-2007 :

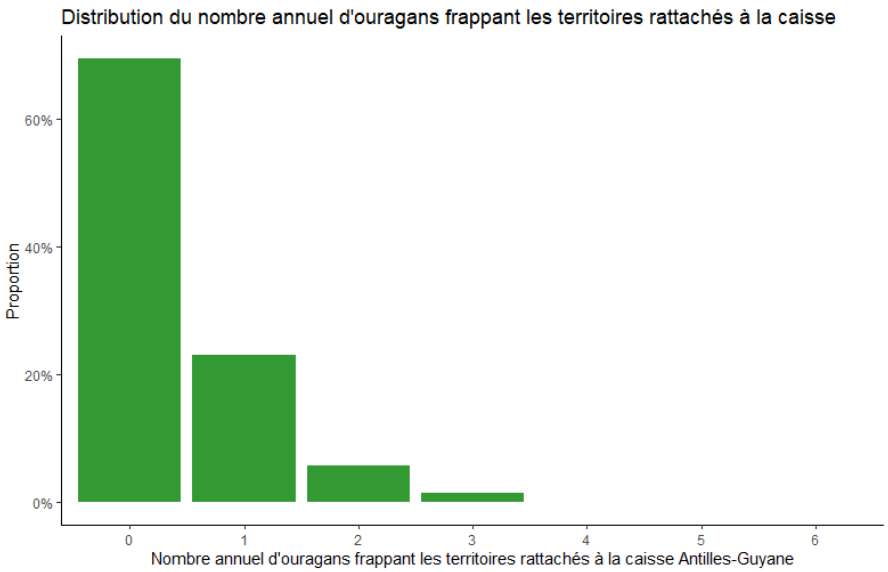


Figure 71 - Histogramme de la distribution du nombre annuel d'ouragans frappant la Guadeloupe, la Martinique, Saint-Barthélemy, Saint-Martin ou la Guyane sur la période 1950-2007

Nous faisons ici l'hypothèse que les nombres annuels d'ouragans frappant un des 5 territoires de la caisse Antilles-Guyane sur la période 1950-2007 sont indépendants. Aussi, nous pouvons voir en figure 73 que le nombre annuel d'ouragans formés dans l'Atlantique ne présente pas de tendance haussière ou baissière sur la période 1950-2007. Nous pouvons donc raisonnablement faire l'hypothèse que les nombres annuels d'ouragans frappant les territoires de la caisse sont identiquement distribués sur cette période.

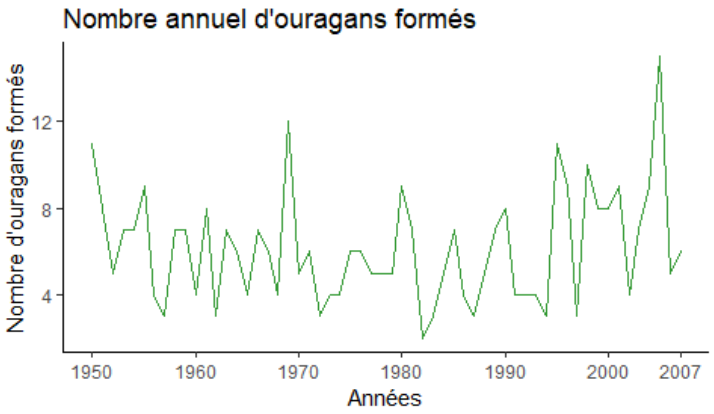


Figure 72 - Évolution du nombre annuel d'ouragans se formant dans l'Atlantique entre 1950 et 2007

La distribution présentée en figure 72 est unimodale et légèrement sur-dispersée : la variance, en l'occurrence égale à 0.45, excède la moyenne de la série, égale à 0.39. Deux lois de probabilité sont alors choisies comme candidates pour modéliser la fréquence à laquelle sont frappés les territoires de la caisse : une loi de Poisson et une loi Binomiale Négative. La première est utilisée pour modéliser un phénomène sans sur-dispersion (variance et espérance de la loi sont égales), alors que la deuxième permet de modéliser avec sur-dispersion.

Rappelons que si une variable aléatoire discrète notée X est distribuée selon une loi de Poisson de paramètre $\lambda > 0$, alors pour tout $k \geq 0$:

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Si X est distribuée selon une loi Binomiale Négative de paramètres r et p (réels tels que $0 < r$ et $0 < p < 1$), alors pour tout $k \geq 0$:

$$\mathbb{P}(X = k) = \frac{\Gamma(r+k)}{k!\,\Gamma(r)}p^r(1-p)^k$$

En notant $\hat{\lambda}$ l'estimateur par la méthode du maximum de vraisemblance du paramètre λ de la loi de Poisson ajustée sur les données de la figure 72 et $\hat{p}, \hat{r}$ les estimateurs par la même méthode des paramètres p, r de la loi Binomiale Négative ajustée sur les mêmes données, nous obtenons :

Binomiale Négative	**Poisson**
$(\hat{p}, \hat{r}) = (0.88, 2.76)$	$\hat{\lambda} = 0.39$
Log-vraisemblance : -56.63	Log-vraisemblance : -56.90

Visualisons l'histogramme de ces deux distributions ainsi que celui de la distribution empirique :

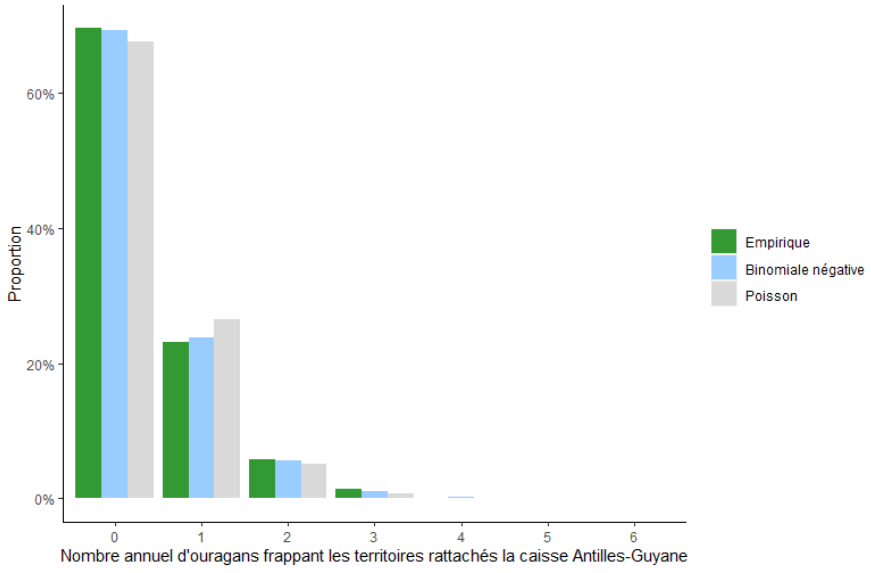


Figure 73 - Histogrammes de la distribution empirique (vert) et des distributions candidates (Binomiale Négative en bleu, Poisson en gris)

La log-vraisemblance associée à la modélisation par loi Binomiale Négative est légèrement plus élevée que celle associée à la modélisation par loi de Poisson mais nous pouvons voir ci-dessus que l'ajustement est, dans les deux cas, très bon.

Retenons la loi Binomiale Négative mais il est important de noter que les données empiriques ne sont composées que de 58 observations (les années allant de 1950 à 2007) et que ce nombre s'avère être bien trop faible pour réaliser une modélisation robuste.

b) Choix des proportions de catégories de Saffir-Simpson

Les proportions de catégories 1 à 5 sur l'échelle de Saffir-Simpson sont observées historiquement ainsi qu'à deux horizons : horizon 2050 et horizon 2100.

Historiquement, les catégories sont observées sur l'ensemble de la période 1970-2005. A l'horizon 2050, les catégories sont observées pendant 20 ans sur la période 2031-2050. A l'horizon 2100, les catégories sont observées pendant 20 ans sur la période 2081-2100.

c) Étude des pertes générées

Voyons quels sont les montants de pertes générés dans le cadre des scénarii RCP 2.6 et RCP 8.5 et comparons les avec les montants générés sous conditions historiques.

Afin d'obtenir des montants de pertes robustes et des intervalles de confiance, 1 000 courbes AEP et OEP ont été simulées : pour un horizon et un GCM donnés, la simulation des 10 000 années de sinistralité a été effectuée 1 000 fois.

➢ Scénario RCP 2.6

	Périodes de retour					
	5 ans	10 ans	20 ans	50 ans	100 ans	200 ans
HadGEM2-ES						
Historique	**8.8** 7 11	**92.6** 80 107	**379.5** 327 446	**1864.6** 1551 2304	**3779.8** 3253 4192	**5933.5** 5056 7204
Horizon 2050	**9.6** 8 12	**106.1** 92 119	**477.9** 389 571	**2400** 1900 2916	**4337.4** 3868 5057	**6914.4** 5769 8201
Horizon 2100	**9.2** 7 11	**97** 84 111	**389.9** 335 460	**1874.6** 1571 2323	**3780.8** 3255 4213	**6052.3** 5075 7218
Écart histo/2050	+9.1%	+14.6%	+29.5%	+28.7%	+14.8%	+16.5%
Écart histo/2100	+4.5%	+4.8%	+2.7%	-	-	+2%
MPI-ESM						
Historique	**10.3** 8 13	**111.8** 98 126	**465** 393 554	**1999.7** 1647 2427	**3840.3** 3316 4278	**5998.7** 5056 7205
Horizon 2050	**9.9** 8 12	**107.5** 94 121	**439.3** 373 520	**1962** 1587 2358	**3844.3** 3252 4278	**6043.5** 5075 7204
Horizon 2100	**10** 8 12	**107.8** 94 121	**440.4** 378 523	**1967.8** 1610 2387	**3846.1** 3265 4279	**6038.7** 5056 7218
Écart histo/2050	-3.9%	-3.8%	-5.5%	-1.9%	+0.1%	+0.7%
Écart histo/2100	-2.9%	-3.6%	-5.3%	-1.6%	+0.1%	+0.7%
NorESM						
Historique	**12.3** 9 15	**139.6** 119 162	**594.4** 506 704	**2360.7** 1922 2755	**4147.4** 3693 4813	**6484.9** 5502 7354
Horizon 2050	**10.7** 8 13	**122.1** 109 141	**575.6** 487 697	**2694.2** 2304 3053	**4668.7** 4087 5410	**7301.8** 6226 8585
Horizon 2100	**11.8** 9 15	**152.2** 126 177	**824.8** 675 1008	**3260.4** 2937 3722	**5645.4** 4937 6258	**8396.1** 7218 9186
Écart histo/2050	-13%	-12.5%	-3.2%	+14.1%	+12.6%	+12.6%
Écart histo/2100	-4.1%	+9%	+38.8%	+38.1%	+36.1%	+29.5%

Note de lecture : pour HadGEM2-ES sous conditions historiques, la moyenne (sur les 1 000 courbes AEP simulées) des montants associés à une période de retour de 5 ans est 8.8 millions d'euros. L'intervalle de confiance au seuil de 95% est [7M€, 11M€].

Tableau 13 - Détail des 1 000 courbes AEP obtenues par la méthode de resimulation (10 000 années simulées) pour différents GCMs et différents horizons sous scénario RCP 2.6 (pertes au niveau portefeuille, brutes de réassurance, exprimées en millions d'euros). Les montants moyens figurent en gras, en-dessous figure l'intervalle de confiance au seuil de 95%.

Rappelons dans un premier temps que dans le cadre de la méthode de resimulation développée dans cette partie, seules les proportions d'ouragans simulés par catégorie de Saffir-Simpson influent sur les pertes obtenues. La loi donnant le nombre annuel d'ouragans est en effet la même aussi bien en vision historique, qu'à l'horizon 2050 ou 2100.

Nous avons précédemment observé que les proportions d'ouragans générés par catégorie étaient très stables pour le modèle moyennement sensible au changement climatique, MPI-ESM. Cette stabilité se traduit par des courbes AEP à horizon 2050 et 2100 très similaires à celles obtenues sous conditions historiques. Aucune tendance liée à la longueur de la période de retour n'est observée : il n'y a par exemple pas d'augmentation des pertes liées aux périodes de retour élevées.

Tout comme le modèle allemand, HadGEM2-ES génère lui aussi des proportions par catégorie de Saffir-Simpson globalement stables au cours du siècle. Cependant, bien que les montants obtenus en fin de siècle soient très similaires à ceux obtenus sous conditions historiques, les pertes générées à l'horizon 2050 sont nettement supérieures à celles générées sous conditions historiques. Ceci provient de l'augmentation brusque, à l'horizon 2050, de la proportion d'ouragans de catégorie 5 générés.

Concernant le modèle le moins sensible au changement climatique, NorESM, les variations de pertes à horizon 2050 et 2100 sont plus marquées. Si les pertes associées à des périodes de retour courtes, de 5 et 10 ans, varient peu, les pertes associées aux périodes de retour plus longues augmentent quant à elles de façon significative à l'horizon 2100. Elles s'expliquent par la part importante d'ouragans de catégorie 5 générés par le modèle norvégien : ils représentent un quart des ouragans générés à l'horizon 2050, plus d'un tiers à l'horizon 2100, alors même qu'ils ne représentent que 14% des ouragans générés sous conditions historiques.

➢ Scénario RCP 8.5

	Périodes de retour					
	5 ans	10 ans	20 ans	50 ans	100 ans	200 ans
HadGEM2-ES						
Historique	**8.8** 7 11	**92.6** 80 107	**379.5** 327 446	**1864.6** 1551 2304	**3779.8** 3253 4192	**5933.5** 5056 7204
Horizon 2050	**8.8** 7 11	**91.6** 79 106	**365.6** 312 427	**1914.1** 1563 2370	**3914.2** 3274 4431	**6280.1** 5287 7305
Horizon 2100	**12.2** 10 15	**154.1** 130 177	**788.3** 642 937	**3102.5** 2769 3389	**5231.9** 4518 5805	**7861.7** 6929 8644
Écart histo/2050	-	-1.1%	-3.7%	+2.6%	+3.5%	+5.8%
Écart histo/2100	+38.6%	+66.4%	+107.7%	+66.4%	+38.4%	+32.5%
MPI-ESM						
Historique	**10.3** 8 13	**111.8** 98 126	**465** 393 554	**1999.7** 1647 2427	**3840.3** 3316 4278	**5998.7** 5056 7205
Horizon 2050	**11.3** 9 14	**129.3** 112 149	**577.2** 488 682	**2488.1** 2029 2917	**4333.5** 3887 5055	**6821.6** 5689 7829
Horizon 2100	**14.2** 11 18	**211.6** 172 251	**1128.5** 955 1352	**3722.9** 3252 4128	**6197.4** 5501 7026	**8782.1** 7830 9681
Écart histo/2050	+9.7%	+15.6%	+24.1%	+24.4%	+12.8%	+13.7%
Écart histo/2100	+37.9%	+89.3%	+142.7%	+86.2%	+61.4%	+46.4%
NorESM						
Historique	**12.3** 9 15	**139.6** 119 162	**594.4** 506 704	**2360.7** 1922 2755	**4147.4** 3693 4813	**6484.9** 5502 7354
Horizon 2050	**11.7** 9 15	**138.6** 118 164	**649** 544 787	**2798.2** 2420 3143	**4797.8** 4130 5565	**7411.4** 6499 8585
Horizon 2100	**14.2** 11 18	**211.6** 172 251	**1128.5** 955 1352	**3722.9** 3252 4128	**6197.4** 5501 7026	**8782.1** 7830 9681
Écart histo/2050	-4.9%	-0.7%	+9.2%	+18.5%	+15.7%	+14.3%
Écart histo/2100	+15.5%	+51.6%	+89.8%	+57.7%	+49.4%	+35.4%

Note de lecture : pour HadGEM2-ES sous conditions historiques, la moyenne (sur les 1 000 courbes AEP simulées) des montants associés à une période de retour de 5 ans est 8.8 millions d'euros. L'intervalle de confiance au seuil de 95% est [7M€, 11M€].

Tableau 14 - Détail des 1 000 courbes AEP obtenues par la méthode de resimulation (10 000 années simulées) pour différents GCMs et différents horizons sous scénario RCP 8.5 (pertes au niveau portefeuille, brutes de réassurance, exprimées en millions d'euros). Les montants moyens figurent en gras, en-dessous figure l'intervalle de confiance au seuil de 95%.

Remarquons dans un premier temps que les modèles sont unanimes sur l'augmentation des pertes associées à toutes les périodes de retour, particulièrement en fin de siècle. Ceci provient des proportions élevées d'ouragans puissants générés à l'horizon 2100. A l'horizon 2050, bien que certaines augmentations de pertes soient déjà perceptibles pour deux des trois GCMs, le modèle anglais, le plus sensible au changement climatique, génère des pertes semblables à celles obtenues sous conditions historiques. Ce phénomène s'explique par des proportions d'ouragans générés par catégorie de Saffir-Simpson stables jusqu'en milieu de siècle.

A l'horizon 2100, nous pouvons constater fortes divergences entre les modèles quant à l'augmentation des pertes associées à certaines périodes de retour : par exemple, concernant la période de retour 20 ans, notons l'augmentation de 89,8% de la perte moyenne pour le modèle NorESM et l'augmentation de 142,7% de la perte moyenne pour le modèle MPI-ESM. Les pertes associées aux périodes de retour plus longues sont en revanche davantage stables : notons que l'augmentation de la perte moyenne de période de retour 100 ans varie entre 38% et 61%, celle de période de retour 200 ans varie entre 33% et 46%.

Comme dans le cadre du scénario RCP 2.6, le modèle le plus sensible au changement climatique n'engendre pas les augmentations de pertes les plus grandes.

Voyons comment varie la période de retour d'événements comme Irma, Maria et Dean à l'horizon 2100 à l'aide des courbes OEP :

	Irma	Dean	Maria
HadGEM2-ES			
Historique	26.2	13.9	7.6
Horizon 2100	20.1	11.3	6.5
Écart histo/2100	-23.3%	-18.7%	-14.5%
MPI-ESM			
Historique	29	12.9	7
Horizon 2100	17.1	10.2	6.4
Écart histo/2100	-41%	-20.9%	-8.6%
NorESM			
Historique	21.4	11.1	6.5
Horizon 2100	14.4	8.6	5.9
Écart histo/2100	-48.6%	-22.5%	-9.2%

Tableau 15 – Périodes de retour moyennes (sur les 1 000 courbes OEP simulées) associées aux événements Irma, Dean et Maria selon l'horizon et le GCM générateur (projections réalisées sous scénario RCP 8.5, périodes de retour en années)

Nous pouvons voir que plus l'événement est extrême, plus la réduction de période de retour à l'horizon 2100 est importante. Les pertes extrêmes sont donc plus fréquentes et ceci provient des proportions élevées d'ouragans puissants générés à l'horizon 2100.

Ainsi, dans le cadre du scénario RCP 2.6, les pertes financières générées par la méthode de resimulation sont stables au cours du XXI^{ème} siècle. En revanche, dans le cadre du scénario

RCP 8.5, la proportion grandissante d'ouragans puissants générés par les modèles engendre des pertes financières bien plus élevées en fin de siècle que sous conditions historiques. Si les pertes associées aux périodes de retour courtes et moyennes augmentent de façon très variable suivant le modèle utilisé, cette augmentation est plus claire pour les périodes de retour longues. Les événements majeurs que sont Irma, Dean et Maria voient alors leur période de retour baisser significativement à l'horizon 2100. Gardons cependant à l'esprit que nous avons fait une hypothèse forte de stationnarité du nombre d'ouragans générés tout au long du siècle.

Conclusion générale

Les données issues du projet CORDEX sont particulièrement adaptées à l'étude de phénomènes globaux comme l'augmentation des températures par exemple. Cependant, nous avons pu voir que la modélisation de phénomènes aussi locaux, rares et complexes que les ouragans comporte de nombreux biais et que ces derniers dépendent du modèle climatique générateur. Un de ces biais, celui des vitesses de vent, concerne une variable déterminante pour cette l'étude et il est alors nécessaire de mettre en place une correction. Cette dernière a un impact considérable sur toute la suite de l'étude, il est important de le garder à l'esprit.

Afin de transformer les événements observés sur les simulations climatiques en pertes financières pour Groupama, un modèle catastrophe proposé par un modélisateur extérieur au Groupe a été utilisé. Ce modèle repose sur des hypothèses et des paramètres, notamment les courbes de vulnérabilité, fixés par le modélisateur. Nous avons pu constater que ce paramétrage n'est pas optimal puisque les périodes de retour associées aux pertes historiques des ouragans Dean, Maria et Irma sont trop courtes par rapport à celles constatées historiquement.

Bien que ce modèle soit conservateur et ne tienne pas compte de l'augmentation à venir des dégâts liés aux fortes précipitations induite par le réchauffement climatique, nous l'avons utilisé en proposant de le coupler aux données issues des projections climatiques de deux façons différentes.

La première méthode de couplage, ou méthode de « recherche des ouragans voisins », repose grandement sur les données issues des simulations. Cette méthode consiste à associer un ouragan observé sur les simulations climatiques à 3 ouragans issus du catalogue du modélisateur. La moyenne des 3 pertes est alors vue comme la perte engendrée par l'ouragan observé sur les simulations climatiques. Bien que l'association des ouragans fonctionne et que cette méthode permette d'exploiter grandement les données extraites, les biais des modèles climatiques, en particulier le biais des trajectoires, impactent lourdement les pertes obtenues et les rendent ainsi inexploitables.

La seconde méthode de couplage, ou méthode de « resimulation », repose quant à elle davantage sur le modèle catastrophe et moins sur les données extraites des projections climatiques. Cette méthode consiste à tirer des ouragans dans le catalogue fourni par le modélisateur dans les proportions de catégories de Saffir-Simson constatées sur les simulations climatiques. Certaines données, comme les empreintes des ouragans, étaient problématiques et

ne sont donc plus utilisées : seule la vitesse des vents générés est utilisée et la correction des vitesses de vent mise en place au début de ce mémoire impacte donc fortement les pertes générées ici. Si les pertes générées par cette méthode en scénario RCP 2.6 sont semblables aux pertes générées en conditions historiques, celles obtenues dans le cadre du scénario RCP 8.5 leur sont supérieures, particulièrement en fin de siècle. Cette augmentation est liée à la plus forte proportion d'ouragans puissants constatés sur les projections.

Nous avons donc pu voir que l'exploitation de projections climatiques accessibles grâce au contexte de données ouvertes permet de fournir, au prix d'approximations et de corrections, une estimation de l'impact du réchauffement climatique sur le risque cyclonique dans les Antilles pour Groupama. Cependant, la structure des projections aussi bien que les projections en elles-mêmes sont amenées à évoluer. Des mailles plus fines, plusieurs *runs* disponibles pour un même GCM ou des modèles moins biaisés pour la simulation des ouragans seront disponibles à l'avenir et permettront une estimation plus rigoureuse. Des avancées pourraient aussi voir le jour avec le 6ème rapport du GIEC publié courant 2022 : davantage d'informations concernant l'influence du changement climatique sur les ouragans pourraient être données et le cadre nous servant dans ce mémoire, à savoir les conclusions du 5ème rapport, pourrait alors grandement évoluer.

Bibliographie

Abadie, A. (2018). Catastrophes naturelles : en Outre-mer, l'assurabilité en question. *Argus de l'Assurance*.

Autorité de la Concurrence (2018). Avis 18-A-09 du 03 octobre 2018 relatif à la situation concurrentielle sur les marchés des matériaux de construction à Mayotte et à la Réunion.

Baldini, L. M., Baldini, J. U. L., McElwaine, J. N., Benoit Frappier, A., Asmerom, Y., Liu, K., Prufer, K. M., Ridley, H. E., Polyak, V., Kennett, D. J., Macpherson, C. G., Aquino, V., Awe, J. & Breitenbach S. F. M. (2016). Persistent northward North Atlantic tropical cyclone track migration over the past five centuries. *Nature*.

Caisse Centrale de Réassurance (2020). Évolution du risque cyclonique en Outre-mer à horizon 2050.

Comaniciu, D. & Meer, P. (2002). Mean shift : a robust approach toward feature space analysis. *Institute of Electrical and Electronics Engineers*.

CORDEX (2014). CORDEX requested variables.

Deparnis, K. G. (2005). Mean Shift Clustering.

Direction Générale des Finances Publiques (2016). Guide de la T.V.A. à l'usage des collectivités locales.

Diro, G. T., Giorgi, F., Fuentes-Franco, R., Walsh, K. J. E., Giuliani, G. & Coppola, E. (2014). Tropical cyclones in a regional climate change projection with RegCM4 over the CORDEX Central America domain. *Climate Change*.

Fuentes-Franco, R., Giorgi, F., Coppola, E. & Zimmermann, K. (2016). Sensitivity of tropical cyclones to resolution, convection scheme and ocean flux parametrization over eastern tropical Pacific and tropical north Atlantic oceans in the RegCM4 model. *Climate Dynamics*.

GIEC (2013). Climate change 2013 : The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the intergovernmental panel on climate change. *IPCC's Fifth Assessment Report*.

GIEC (2013). Technical Summary. *IPCC's Fifth Assessment Report*.

Hebert, C. G., Weinzapfel, R. A. & Chambers, M. A. (2005). Hurricane Severity Index : a new way of estimating hurricane destructive potential.

Hutton, M. (2019). Octroi de mer : où en est-on ? *La Première* (*France Info*).

Kossin, J. P., Emanuel, K. A. & Vecchi, G. A. (2014). The poleward migration of the location of tropical cyclone maximum intensity. *Nature*.

Lopez, O. (2018). Statistique inférentielle. *ISUP*

Luzi, M. (2016). Actuariat IARD, techniques et illusions. *Bulletin français d'actuariat*.

Météo-France. http://www.meteofrance.fr/

Noel, V. (2020). La genèse des chapelets de tempêtes. *Institut des Actuaires*.

Projet DRIAS. http://www.drias-climat.fr/

Tierney, J. E., Jiang, Z., King, J., Malevich, S. B., Hakim, G. J. & Poulsen, C. J. (2020). Glacial cooling and climate sensitivity revisited. *Nature*.

Vishnu, S., Sanjay, J. & Krishnan, R. (2019). Assessment of climatological tropical cyclone activity over the north Indian Ocean in the CORDEX-South Asia regional climate models. *Climate Dynamics*.

Annexes

Suivi des ouragans

Les données issues du projet CORDEX sont des matrices à 225 lignes et 337 colonnes donnant la vitesse des vents en chaque point de la grille de maille 25 km recouvrant le domaine CAM. Deux matrices successives donnent la vitesse des vents à trois heures d'intervalle.

Chaque matrice peut être transformée en image en affectant à chaque coefficient une couleur. Cette dernière peut par exemple être d'autant plus claire que le coefficient associé est élevé :

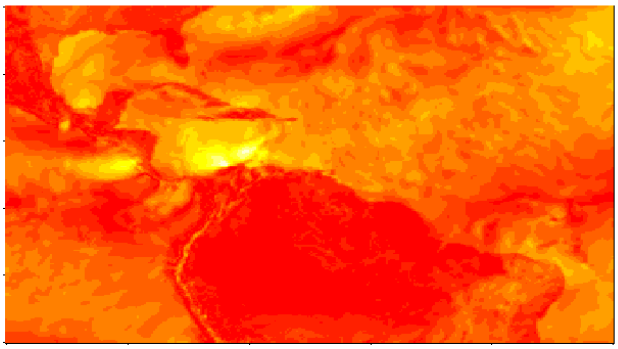


Figure 74 – Visualisation d'une matrice de vitesses de vent

Extraire les ouragans présents dans nos données peut alors être assimilé à la détection et au suivi d'objets de couleur jaune claire sur plusieurs images successives, c'est-à-dire à du *tracking* d'objets.

Afin de faciliter le suivi des ouragans, nous procédons dans un premier temps au repérage des matrices de vents forts : les pics d'intensité des ouragans sont alors détectés. Suivre les ouragans entre ces matrices de vents forts doit permettre de déterminer leur empreinte, de savoir à quel moment ils ont été générés, à quel moment ils ont disparu : c'est là qu'intervient l'algorithme de *tracking*.

Les matrices situées entre deux matrices de vents forts sont filtrées au seuil de 16 m/s. Ainsi, suivre un ouragan s'apparente à suivre une tâche claire sur un fond rouge parmi un ensemble d'autres tâches claires (qui sont des perturbations moins puissantes) :

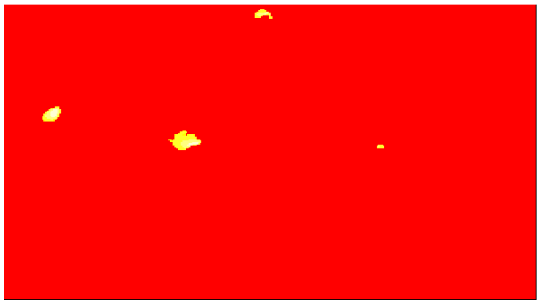


Figure 75 - Visualisation d'une matrice de vitesses de vent filtrée au seuil de 16 m/s

Différents algorithmes de *tracking* existants ont été testés :

- KCF (Kernelized Correlation Filters), CSRT (Channel and Spatial Reliability Tracker) ou encore MOSSE (Minimum Output Sum of Squared Error), algorithmes classiques de *tracking* utilisant l'information disponible sur une image pour rechercher efficacement l'objet sur l'image suivante.
- GOTURN (Generic Object Tracking Using Regression Networks), algorithme de *tracking* faisant appel à de l'apprentissage automatique (ou *machine learning*). Contrairement à d'autres *trackers* fonctionnant par l'apprentissage automatique, il ne nécessite pas d'étape d'entraînement préalable : il a en effet déjà été entraîné à suivre des objets très variés comme des balles, des visages, des mains ou des voitures.

Ces algorithmes sont *detection-free*, c'est-à-dire qu'ils ne détectent pas automatiquement l'objet à suivre, l'utilisateur doit indiquer la position des objets à suivre. Nous n'avons pas besoin de détection automatique puisque ces positions sont déjà connues : ce sont celles des ouragans sur les matrices de vents forts.

Indiquons donc au *tracker* la position de l'objet à suivre, à savoir l'ouragan, sur la matrice de vents forts filtrée au seuil de 16 m/s :

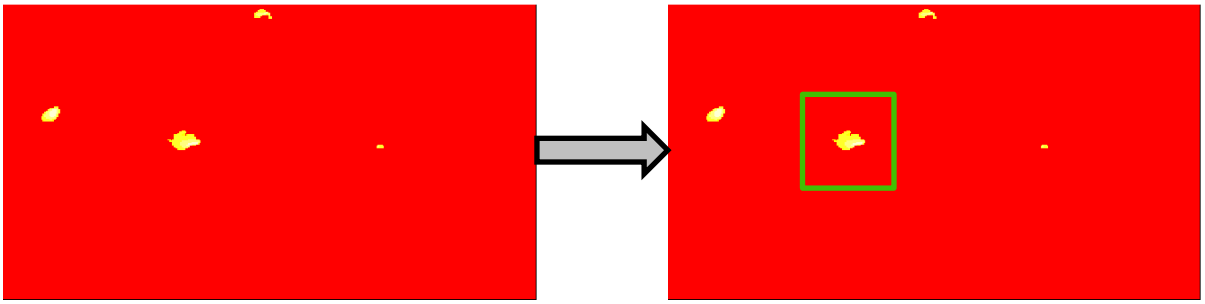


Figure 76 – Visualisation d'une matrice de vents forts filtrée au seuil de 16 m/s (gauche) et de la même matrice avec tracker positionné (droite)

Pour chacun de ces algorithmes, le suivi des ouragans est, au début, satisfaisant. Cependant, lorsque l'ouragan disparaît de l'image (soit parce qu'il quitte le domaine CAM, soit parce qu'il engendre des vents trop faibles), le *tracker* s'emballe :

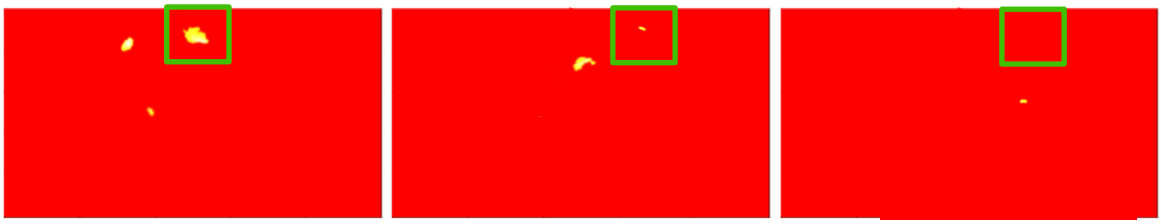


Figure 77 - Suivi d'un ouragan par le tracker

Dans la plupart des cas, juste après la disparition de l'ouragan suivi, le *tracker* reste immobile à l'endroit de la disparition. Puis, il suit généralement une autre dépression n'ayant rien à voir avec l'ouragan initialement suivi.

Face à ces nombreux problèmes, créer de toute pièce un algorithme de *tracking* adapté à nos données en se basant sur un algorithme existant (Mean-Shift) apparaît comme une alternative convenable.

Algorithme Mean-Shift

L'algorithme Mean-Shift est utilisé pour extraire les ouragans des données issues du projet CORDEX : sur une matrice de vent donnée, il permet de déterminer le nombre de perturbations présentes ainsi que leur localisation.

Pour rappel, nous disposons au départ de matrices de vent filtrées à un certain seuil s (33 m/s ou 16 m/s selon le cas) : seuls les coefficients de la matrice supérieurs à ce seuil sont non nuls. Ces N coefficients non nuls sont ensuite vus comme N points de $\mathbb{R}^2$, ils sont alors décrits par leur indice de colonne et leur indice de ligne dans la matrice.

Nous considérons ensuite que ces points, notés $x_1, \dots, x_N$, sont N réalisations de variables aléatoires $X_1, \dots, X_N$ indépendantes et identiquement distribuées de densité commune notée f. Cette densité est estimée par noyau gaussien, notons $\hat{f}_H$ l'estimateur. Il est défini de la façon suivante :

$$\hat{f}_H(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i)$$

Avec :

- x un point quelconque de $\mathbb{R}^2$
- H la matrice de lissage, égale à $\begin{pmatrix} h^2 & 0 \\ 0 & h^2 \end{pmatrix}$ avec h réel positif.
- $K_H(x) = (2\pi)^{-1}|H|^{-1/2} \exp\left(-\frac{1}{2h^2}\|x\|^2\right)$

Avec une matrice de lissage adéquate, $x \longmapsto \hat{f}_H(x)$ admet un maximum local en chaque groupement de points, c'est-à-dire en chaque ouragan.

Afin de déterminer le nombre et la localisation de ces groupements, à chaque point x_i est associée une suite $(x_i^t)_{t\geq 0}$ convergeant vers le maximum local de $\hat{f}_H$ le plus proche. Deux points x_i et x_j appartiennent au même groupement (et donc décrivent un même ouragan) si leurs suites respectives convergent vers le même maximum.

Pour un point x_i, la suite $(x_i^t)_{t\geq 0}$ est construite de la façon suivante :

$$\begin{cases} x_i^1 = x_i \\ x_i^{t+1} = x_i^t + m(x_i^t) \text{ pour } t \geq 1 \end{cases}$$

La fonction notée m est définie de la façon suivante :

$$m(x) = \frac{\sum_i x_i \exp(-\frac{1}{2} x^T H^{-1} x)}{\sum_i \exp(-\frac{1}{2} x^T H^{-1} x)} - x$$

Définie de cette façon, m permet à la suite $(\hat{f}_H(x_i^t))_{t\geq 1}$ d'être strictement croissante et de converger, tout comme $(x_i^t)_{t\geq 0}$.

Notons que :

$$\hat{f}_H(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i) = \frac{c_H}{n}\sum_{i=1}^{n} k\left(\left\|\frac{x - x_i}{h}\right\|^2\right)$$

avec $k(x) = \exp(-\frac{1}{2}x)$ et c_H une constante dépendant de H. Nous pouvons alors en déduire l'expression du gradient de $\hat{f}_H$ évalué en un point x, noté $\nabla\hat{f}_H(x)$:

$$\nabla\hat{f}_H(x) = \frac{2c_H}{nh^2}\sum_{i=1}^{n}(x_i - x)k'\left(\left\|\frac{x - x_i}{h}\right\|^2\right)$$

D'où :

$$\nabla\hat{f}_H(x) = \frac{2c_H}{nh^2}\left[\sum_{i=1}^{n} k'\left(\left\|\frac{x - x_i}{h}\right\|^2\right)\right]\left[\frac{\sum_{i=1}^{n} x_i k'\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} k'\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x\right]$$

Les deux premiers termes de l'expression du gradient ci-dessus étant positifs, la nullité du gradient équivaut à la nullité du troisième terme. C'est ce troisième terme qui est défini comme la fonction m.

> **Théorème** (Theorem 1, Comaniciu & Meer[1]) : si le noyau K_H possède un profil convexe et strictement décroissant, alors les suites $(x_i^t)_{t\geq 0}$ et $(\hat{f}_H(x_i^t))_{t\geq 0}$ convergent et cette dernière suite est strictement croissante.

Le profil de notre noyau étant la fonction k, nous pouvons en déduire que la suite $(x_i^t)_{t\geq 0}$ converge. Par conséquent, par définition de $(x_i^t)_{t\geq 0}$, la suite $(m(x_i^t))_{t\geq 0}$ tend vers le vecteur nul et $(x_i^t)_{t\geq 0}$ converge vers un point annulant le gradient de $\hat{f}_H$. De plus, la suite $(\hat{f}_H(x_i^t))_{t\geq 0}$ convergeant de façon croissante, $\hat{f}_H$ admet donc un maximum local en la limite de la suite $(x_i^t)_{t\geq 0}$.

Algorithme des *k* plus proches voisins

La première méthode d'estimation des pertes financières pour le Groupe à partir des ouragans observés sur les simulations climatiques exploite grandement les informations extraites de ces simulations. La méthode consiste à associer un ouragan observé à des ouragans semblables

[1] Voir *Mean shift : a robust approach toward feature space analysis*

issus de la table d'événements proposée par le modélisateur. La perte générée par l'ouragan observé est alors obtenue grâce aux pertes des ouragans qui lui ont été associés. Cette méthode s'inspire de l'algorithme des k plus proches voisins.

Supposons que nous souhaitions expliquer une variable Y, à valeur dans $\mathbb{R}$, par p variables quantitatives, elles aussi à valeur dans $\mathbb{R}$, et que n individus soient à notre disposition. Un individu i est alors décrit par $(y_i, x_i) \in \mathbb{R} \times \mathbb{R}^p$.

Supposons que nous disposions d'un nouvel individu décrit par (y_{n+1}, x_{n+1}) avec y_{n+1} inconnu. Estimer y_{n+1} par l'algorithme des k plus proches voisins va d'abord consister à ordonner les couples $(y_1, x_1), \dots, (y_n, x_n)$: en notant $\|.\|$ la norme euclidienne sur $\mathbb{R}^p$, notons $(y_{(1)}, x_{(1)}), \dots, (y_{(n)}, x_{(n)})$ les couples ordonnés tel que :

$$\|x_{(1)} - x_{n+1}\| \leq \|x_{(2)} - x_{n+1}\| \leq \cdots \leq \|x_{(n)} - x_{n+1}\|$$

En fixant k préalablement $(1 \leq k \leq n)$, l'estimateur de y_{n+1} par l'algorithme des k plus proches voisins, noté $\widehat{y_{n+1}}$, est alors défini par :

$$\widehat{y_{n+1}} = \frac{1}{k} \sum_{i=1}^{k} y_{(i)}$$

L'algorithme estime donc y_{n+1} par la moyenne des y_i des k individus situés à proximité. Dans notre cas, $p = 3$ et les individus sont les ouragans proposés par le modélisateur. La perte engendrée par l'ouragan correspond à la variable Y.

Estimation par la méthode du maximum de vraisemblance

Dans cette étude, le paramètre de la loi de fréquence utilisée lors de la resimulation de 10 000 années cycloniques est estimé par la méthode du maximum de vraisemblance.

Soient $x_1, \dots, x_n$ des réalisations de variables aléatoires $X_1, \dots, X_n$ indépendantes et identiquement distribuées de loi $\mathcal{L}_\theta$ avec θ le paramètre de la loi. Estimer ce dernier par la méthode du maximum de vraisemblance revient à chercher le paramètre θ maximisant la probabilité d'obtenir l'échantillon $x_1, \dots, x_n$ avec la loi $\mathcal{L}_\theta$: cette probabilité est la vraisemblance de l'échantillon.

Dans le cas où $\mathcal{L}_\theta$ est une loi discrète, la vraisemblance est définie par :

$$L_n(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} \mathbb{P}_\theta(\{X_i = x_i\})$$

Dans le cas où $\mathcal{L}_\theta$ est une loi continue et admet une densité f_θ, la vraisemblance de cet échantillon est définie par :

$$L_n(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} f_\theta(x_i)$$

L'estimateur de θ obtenu par la méthode du maximum de vraisemblance, noté $\hat{\theta}$, est alors défini par :

$$\hat{\theta} = \text{argmax}_\theta \, L_n(x_1, \dots, x_n; \theta)$$

L'estimateur obtenu possède les propriétés suivantes :

- C'est un estimateur convergent. Il converge en probabilité vers le paramètre θ qu'il est sensé estimer.
- Si l'information de Fisher existe, il est asymptotiquement efficace. Autrement dit, la variance de l'estimateur tend vers la borne de Cramer-Rao, c'est-à-dire la variance minimale pour un estimateur non biaisé.
- Il est asymptotiquement distribué selon une loi normale. Autrement dit, plus l'échantillon est large, plus la distribution des valeurs prises par l'estimateur se rapproche d'une distribution normale centrée sur la valeur du paramètre à estimer.

Exemple : cas d'une loi de Poisson de paramètre λ

Supposons que la loi $\mathcal{L}_\theta$ soit une loi de Poisson paramétrée par λ, réel positif. Notons $\hat{\lambda}$ l'estimateur de ce paramètre obtenu avec la méthode du maximum de vraisemblance.

La vraisemblance de l'échantillon est :

$$L_n(x_1, \dots, x_n; \lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \frac{\lambda^{\sum x_i}}{\prod_i x_i!}$$

La log-vraisemblance est donc :

$$\log\big(L_n(x_1, \dots, x_n; \lambda)\big) = -n\lambda + \log(\lambda) \sum_i x_i - \sum_i \log(x_i!)$$

La fonction $\lambda \longmapsto \log\big(L_n(x_1, \dots, x_n; \lambda)\big)$ est deux fois dérivable de dérivée première :

$$\frac{\mathrm{d} \log\big(L_n(x_1, \dots, x_n; \lambda)\big)}{\mathrm{d}\lambda} = -n + \frac{1}{\lambda} \sum_i x_i$$

Calculer la dérivée seconde permet de mettre en évidence la concavité de la vraisemblance. La dérivée première s'annulant en $\frac{1}{n}\sum_i x_i$, nous pouvons en déduire que $\hat{\lambda} = \frac{1}{n}\sum_i X_i$.

L'estimateur du paramètre $\hat{\lambda}$ par la méthode du maximum de vraisemblance est donc $\hat{\lambda} = \frac{1}{n}\sum_i X_i$.