



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

Le 09 juillet 2012

Par : David GUILLEMOT

Titre: **Lien entre tarification des événements climatiques en
Multi-Risques Habitation et modèle interne Non-Vie.**

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

Membre du jury de l'Institut des Actuaires

Entreprise :

Groupama SA

Membres du jury I.S.F.A.

M. Jean Claude AUGROS

M. Alexis BIENVENÛE

M. Areski COUSIN

Mme Diana DOROBANTU

Mme Anne EYRAUD-LOISEL

M. Nicolas LEBOISNE

M. Stéphane LOISEL

Mlle Esterina MASIELLO

Mme Véronique MAUME-DESCHAMPS

M. Frédéric PLANCHET

M. François QUITTARD-PINON

Mme Béatrice REY-FOURNIER

M. Pierre RIBEREAU

M. Christian-Yann ROBERT

M. Didier RULLIERE

M. Pierre THEROND

Directeur de mémoire en entreprise :

Hervé PINAUD

Invité :

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

Signature du responsable entreprise

Secrétariat

Mme Marie-Claude MOUCHON

Signature du candidat

Bibliothèque :

Mme Michèle SONNIER

« Tous les modèles sont faux ; certains sont utiles » G. Box¹.

« Un modèle n'est qu'une simplification de la réalité. Il y a des modèles pour comprendre, d'autres pour expliquer, et d'autres pour prédire. » G. Saporta²

¹ George Edward Pelham Box (né le 18 octobre 1919 à Gravesend, Kent, Angleterre) est un statisticien, qui a apporté d'importantes contributions aux domaines du contrôle qualité, des séries temporelles et de l'inférence bayésienne.

² Gilbert Saporta est professeur de Statistiques appliquées au CNAM depuis 1984. Il est responsable de l'équipe MSDMA (Méthodes Statistiques de Data Mining et Apprentissage) au sein du Centre d'Etudes et De Recherche en Informatique.

Résumé

Mots clés : *risque catastrophe, évènements climatiques / aléas climatiques, régime « Cat. Nat. », fréquence, coût moyen, prime pure, tests d'indépendance, tests de significativité, modèles stochastiques, régression (modèles linéaires généralisés), réassurance*

Face au nombre croissant d'aléas climatiques (tempêtes, grêles, ...) et catastrophes naturelles depuis quelques années, une réflexion plus approfondie quant à la survenance de ces événements erratiques doit être accordée, notamment pour des produits d'assurances couvrants ce type de risques. Les retours plus fréquents de ces phénomènes, probablement liés au dérèglement climatique, remettent en cause les concepts actuels (les charges sinistres liées à ces événements prennent un poids de plus en plus important). Les fondements économiques et financiers des différents produits ne sont donc plus nécessairement en adéquation face à ces aléas naturels.

Dans le cadre d'une tarification de telles couvertures d'assurance, il peut alors sembler cohérent de se repositionner sur 3 aspects :

- la période de retour de ces événements (plus rapprochée),
- leurs intensités (Xynthia en 2010 ou Klaus en 2009 en sont de parfaits exemples),
- leurs localisations (comme la répétition des événements du Var en 2009 et 2011).

L'impact des récents événements naturels dans les résultats du Groupe Groupama en France et l'arrivée de la nouvelle réglementation Solvabilité 2 (S2) ont permis de mener une réflexion sur les produits Multi Risques Habitation (MRH) actuellement commercialisés (notamment dans l'approche tarifaire existante).

Avec la mise en place du modèle interne et de par la proposition de la Directive S2, il est précisé qu'« un modèle interne ne pourra être utilisé à des fins réglementaires que si l'organisme montre qu'il est utilisé par ailleurs »³.

L'enjeu de ce rapport est triple :

- Tout d'abord, il offre l'opportunité de créer un lien entre la tarification de la garantie « événements climatiques » d'un produit MRH et le module « risque de souscription non vie » du modèle interne (Sujet principal de ce mémoire) ;
- Deuxièmement, l'objectif est d'apporter aux différentes Entités⁴, une nouvelle approche tarifaire qui soit en lien avec le risque porté par ces dernières ;
- Enfin, il présente de nouveaux aspects, tant dans la considération des capitaux alloués et l'exigence de rentabilité attendue, que dans la perception des risques.

Il est important de mettre à disposition des institutions dirigeantes, des indicateurs synthétisant le profil de risque de leur portefeuille. En effet, les modifications des conditions climatiques peuvent :

- changer le profil de risque d'une région, avec des conséquences directes sur l'exposition de chaque entité assurée⁵ au regard de son implantation géographique ;
- favoriser ou défavoriser la demande de certains produits et jouer ainsi sur le potentiel de marché, affectant la valeur à long terme de la société.

L'analyse et la modélisation des risques et les questions d'assurabilité sont les axes centraux de ce mémoire.

Il sera donc dressé dans une première section, un bref panorama des enjeux économiques que représentent les produits MRH au sein du Groupe Groupama.

Dans une deuxième partie, le cœur même de ce rapport, il sera présenté une méthodologie de tarification, basée sur les modèles de régression et qui sera par la suite, mise en lien avec les travaux de modélisation réalisés au sein de la Direction Actuariat Groupe.

Dans une troisième partie, des aléas et sensibilités des méthodologies précédemment employées seront présentés.

Pour conclure, face au projet de loi déposé au Sénat sur la réforme du système « Cat. Nat. », il sera présenté une réflexion sur les risques naturels et les différentes perspectives qui pourraient y être liées. Ces travaux sont actuellement en cours de réalisation et sont de nature confidentielle.

³ Site Internet de l'Autorité de Contrôle des Assurances et des Mutuelles (ACAM) http://www.acam-france.fr/modeles_internes/

⁴ Vocabulaire spécifique à la compagnie désignant les différentes sociétés constituant le Groupe.

⁵ Une entité assurée peut aussi bien décrire une habitation, une automobile, un homme, une femme, un chien, un chat, ...

Abstract

Keywords: *catastrophe risk, climatic events / climatic hazards, natural catastrophes, frequency, average of claims, pure premium or risk cost, tests of independence, tests of significance, stochastic models, regression (generalized linear models), reinsurance*

Since the last few years, the increasing number of climatic hazards (windstorm, hail,...) and natural catastrophes require a detailed reflection about the cycle of these erratic events, focussing on insurance products linked to those phenomenon. Indeed, the claim expenses, caused by these events, grow larger year on year. Therefore, the economic and financial basements of these various products are not currently relevant anymore.

In addition to that, the more frequent returns of these phenomena, perhaps linked to climate changes, reconsider the current concepts.

For such insurance coverage, it seems to be coherent to re-evaluate three different parts:

- The erratic cycle of these events (higher frequency),
- Their strength (Xynthia in 2010 or Klaus in 2009 are very good examples),
- Their location (as the recurrence of the events in the Var in 2009 and 2011).

The impact of the recent natural events on the Group Groupama's income in France and the Solvency 2 (S2) regulations coming have required a reflection on the household insurance (especially in the current rating approach).

Groupama has decided to set up an internal model. Moreover, the S2 Directive specifies that "an internal model can be used as legal intentions only if the company proves that it's besides applied".

The purpose in this report is triple:

- Firstly, it enables to link the rating of the "climate events" cover of a household insurance to the "Non life CAT risk sub – module" of the internal model (main subject of this report);
- Secondly, the objective is to bring to the various Entities a new rating approach which would be connected to their risk;
- Finally, it presents some perspectives of evolution in the perception of the risk, such as the rating aspect and the preventive aspect.

Representative indicators of the portfolio risk profile have to be provided to institutions leaders (executives officers). Indeed, climatic conditions changes could:

- change the risk profile of a region, with immediate consequences on the exposure of every insured entity (due to its geographical location);
- promote or depreciate products needs and participate on the market potential, impacting the long-term value of the company.

Risk analysis, modelling and insurability matters are the mains subjects of this report.

The first section will introduce a brief panorama of the economical stakes of household insurance in Groupama.

In the second part, the core of this report, a methodology of pricing will be developed, firstly based on regression's models and which will be afterward, linked to the internal model.

In the third part, hazards and sensibilities of methodologies will be exposed.

To conclude, a thought about the new law, submitted to the Senate, dealing with natural catastrophes, will be presented.

Avant-propos

La rédaction d'un rapport est l'occasion de faire un état des lieux de ses connaissances et/ou de découvrir et approfondir de nouvelles notions et concepts.

Depuis mon arrivée dans le Groupe⁶, mes principales missions se sont articulées autour de trois grandes orientations.

La première réside dans la réalisation de diverses études techniques actuarielles de tarification. De nombreuses refontes de structures tarifaires ont été proposées afin de permettre un ajustement des risques et d'assurer l'équilibre financier de différents produits. Ces études ont principalement porté sur les produits d'Assurance de Biens et de Responsabilité (Automobile, Habitation, ...) mais également de Personnes (Santé, Invalidité, ...).

La deuxième orientation est davantage fondée sur le suivi, le pilotage et la détermination des provisions techniques. Dans le cadre des arrêtés de compte, j'ai eu l'opportunité de mettre en place différents modèles qui répondent aux enjeux d'évaluation des provisions et qui permettent ainsi d'estimer correctement le niveau technique des produits.

Enfin, un dernier axe composé d'études transversales qui m'ont permis d'élargir mon périmètre d'action et d'échanger avec les différentes Directions et Entités. Des études de surveillance, notamment sur le risque décennal, ou bien encore des outils d'optimisation des ressources du réseau commercial, ont ainsi vu le jour.

Le sujet développé ci-après, institué par la mise en place du modèle interne et la refonte envisagée des produits MRH, a permis d'approcher le métier sous un angle très large.

Il a permis de revenir, non seulement sur les activités décrites ci-dessus, mais également d'approfondir des notions, telles que les exigences de fonds propres sous Solvabilité 2, et donner une autre perception du métier, avec l'allocation et le coût du capital liés.

Le dérèglement climatique est l'aléa principal des années à venir. Le subir est risqué, il est préférable d'agir en amont en tentant de l'« anticiper ».

Remerciements

Ce rapport est également une occasion de remercier ceux qui vous ont apporté leur savoir faire, leur concours ou leur soutien.

Ma première pensée va bien évidemment à Hervé Pinaud qui a accepté de se lancer dans ces travaux. Grâce à lui, j'ai pu rapprocher la théorie et la pratique sur les notions d'exigence de fonds propres, d'allocation et coût du capital.

Bien évidemment, je remercie mes employeurs (Groupama Grand Est et Groupama SA), qui m'ont permis, par les missions confiées d'avoir accès à une information riche et variée et de travailler sur de nombreux sujets.

Merci aux nombreux collègues de la DADIP mais également mes anciens responsables et collègues de Groupama Grand Est, avec qui j'ai appris mon métier, une méthode de travail et un esprit critique.

Enfin, merci aux différents interlocuteurs des diverses Entités croisés durant ces six années, les échanges sont toujours riches d'enseignements.

Je tiens bien évidemment à remercier Stéphane Loisel, pour sa formation théorique et ses propos sur la rédaction de ce mémoire. J'en profite pour saluer Gérard, compère de formation, avec qui nous avons commencé cette aventure.

Je remercie également mes anciens professeurs de l'ISFA et de la faculté de Beaulieu qui m'ont permis d'approfondir mes connaissances.

Enfin, merci à ma compagne, à ma famille et mes amis pour leur soutien et leur aide.

⁶ Arrivé en 2006 au sein de Groupama Grand Est, puis au sein de Groupama SA depuis 2010

Sommaire

Résumé	3
Abstract.....	4
Avant-propos.....	5
Remerciements.....	5
Sommaire	6
Section I. Réflexion sur les actuels produits MRH commercialisés au sein du Groupe Groupama	9
I.1. Les enjeux de la MRH (au sein de Groupama en France)	9
I.1.1. Part de marché du Groupe en France	9
I.1.2. Les garanties du produit Privatis commercialisé par les Caisses Régionales	10
I.1.3. Sinistralité observée sur les 4 derniers exercices pour le produit Privatis.....	10
I.1.4. Distinction entre « évènement climatique » et « catastrophe naturelle ».....	11
I.1.5. Cartographie des sinistres sur les exercices 2009 et 2010 pour la garantie « évènements climatiques »	12
I.2. Enoncé de la problématique	13
I.3. Définitions	14
Section II. Lien entre la tarification de la garantie « évènements climatiques » et le modèle interne : détermination et analyse de la prime pure.....	15
II.1. Détermination de l'échelle tarifaire de la garantie « évènements climatiques » selon les critères tarifaires du produit Privatis	15
II.1.1. Vérification de l'indépendance des données analysées	18
(a) Stabilité dans la répartition des sinistres d'un exercice à l'autre	18
(b) Tests d'indépendance des variables tarifaires	19
II.1.2. Détermination des échelles de risque de fréquences selon les critères tarifaires.....	23
(a) Les lois possibles pour la modélisation de la fréquence des sinistres	23
(b) Définition du modèle le plus approprié à la modélisation de la fréquence des sinistres	25
(c) Présentation des échelles de risques de la fréquence	28
(d) Analyse sur les résidus pour validation des hypothèses.....	30
II.1.3. Détermination des échelles de risque des coûts moyens selon les critères tarifaires ...	32
(a) Définition d'un seuil d'exclusion entre les sinistres attritionnels et les sinistres non-récurrents dans le cadre d'une modélisation des coûts moyens	33
(b) Définition du modèle le plus approprié à la modélisation des coûts moyens	35
(c) Présentation des échelles de risques des coûts moyens.....	38
(d) Analyse sur les résidus pour validation du modèle.....	39
II.1.4. Déduction des échelles de prime pure	41
II.2. Modélisation des « Forces de la Nature » dans le modèle interne	43
II.2.1. Les raisons d'un modèle interne	43
II.2.2. Processus de modélisation des évènements « Forces de la Nature ».....	45
II.2.3. Modélisation des évènements atypiques et attritionnels pour le risque « Tempête » via l'outil de marché RMS	46
II.2.4. Synthèse des résultats sur l'ensemble des risques couverts par le traité.....	47
II.3. Prise en compte de la réassurance sur les évènements extrêmes modélisés	47
II.3.1. Description du fonctionnement de la réassurance interne pour les Caisses Régionales	48
II.3.2. Description du programme de réassurance externe.....	48
II.3.3. Resimulation des évènements tempêtes sur la plateforme Risk Explorer.....	49
II.3.4. Synthèse des résultats après application des programmes de réassurance.....	51
II.3.5. Intégration de la facturation du coût de la réassurance.....	52
II.4. Clés de répartition des évènements modélisés entre les risques et produits couverts par le programme « Force de la Nature ».....	53

II.5. Interaction des différents éléments dans la détermination de la prime pure et conséquence pour les Caisses Régionales sur le produit Privatis	54
II.5.1. Détermination du niveau de la prime pure	54
II.5.2. Aide à la décision pour les différentes Entités face à ce nouveau calibrage	55
(a) Comparaison entre le nouveau calibrage de la prime pure et les primes pures brutes observées en 2009 et 2010 (« <i>Back-Testing</i> »)	55
(b) Comparaison entre le nouveau calibrage de la prime pure et la prime pure actuellement paramétrée	56
Section III. Sensibilités dans l'approche des modèles proposés et jeux d'hypothèses	59
III.1. Sensibilités liées à la méthodologie employée dans l'analyse de la prime pure ..	59
III.2. Sensibilités du SCR MRH	62
III.2.1. Contexte réglementaire actuel et perspectives	62
(a) Contexte prudentiel actuel en France.....	62
(b) Contexte prudentiel en 2014 avec Solvabilité 2.....	62
(c) Pilier I : Exigences quantitatives	64
III.2.2. Modalités de calcul du SCR selon le modèle standard et interne.....	66
III.2.3. Analyse des sensibilités et comparaisons liées aux formules de calcul des modèles standard et interne	67
(a) Sensibilités liées au risque de primes et de réserves.....	67
(b) Sensibilités liées aux gains de diversification	68
(c) Sensibilités liées au risque de catastrophe.....	69
(d) Comparatif entre le modèle standard et le modèle interne.....	70
III.3. Hypothèses liées à des modifications du solde de réassurance dans l'évaluation du tarif et de la rentabilité pour la garantie « événements climatiques »	71
III.3.1. Evaluation de la rentabilité d'un produit non-vie	71
III.3.2. Modifications du solde de réassurance et sensibilités liées.....	72
III.4. Modèles économiques : Jeux d'hypothèses dans l'allocation du capital et conséquences sur l'exigence de rentabilité	74
Section IV. Les perspectives face aux risques naturels	79
IV.1. Projet de loi portant sur la réforme du régime d'indemnisation des catastrophes naturelles.....	79
IV.2. Les dérivés climatiques	82
IV.3. Un nouvel algorithme dans l'approche tarifaire des risques naturels	83
IV.3.1. Géolocalisations des risques naturels via les données de la MRN et présentées via le logiciel MAP info	84
(a) Géolocalisation du risque inondation.....	84
(b) Géolocalisation du risque sécheresse.....	87
(c) Géolocalisation des mouvements de terrains.....	89
(d) Géolocalisation du risque sismique	90
IV.3.2. Introduction de variables supplémentaires pour mieux appréhender le risque.....	91
IV.4. La prévention pour réduire les effets des aléas naturels	92
IV.4.1. Les différents dispositifs des Pouvoirs publics face aux aléas naturels	93
IV.4.2. Accompagnement des assurés face aux risques naturels	94
Conclusion	96
Bibliographie	97

Annexes.....	100
Annexe [1.] : Les constats environnementaux au niveau mondial	100
Annexe [2.] : Indice FFB de 1980 à fin 2011	104
Annexe [3.] : Les Caisses Régionales Groupama en France métropolitaine.....	105
Annexe [4.] : Compléments sur les méthodes de régression	106
Annexe [4.1.] : Moindres Carrés Ordinaires	106
Annexe [4.2.] : Maximum de Vraisemblance	106
Annexe [4.3.] : Familles exponentielles et fonctions de lien	108
Annexe [4.4.] : Familles de lois conjuguées	110
Annexe [5.] : Descriptif des tests d'indépendance et de significativité	111
Annexe [5.1.] : Test du Khi-Deux	111
Annexe [5.1.bis] : Les limites du test d'indépendance du Khi-Deux.....	112
<i>Programme SAS [5.1.] : Test d'indépendance.....</i>	<i>113</i>
<i>Programme SAS [5.2.] : Extraction d'un échantillon aléatoire</i>	<i>114</i>
<i>Programme SAS [5.3.] : Constitution d'une base agrégée pour un gain de traitement</i>	<i>114</i>
Annexe [5.2.] : Test de Type I contre test de Type III sous SAS	115
Annexe [5.3.] : Test de Fisher-Snedecor et Test de Student	115
Annexe [6.] : Technique de sélection de variables et choix d'un modèle	116
<i>Programme SAS [6] : Calculs des critères de tests et des paramètres des échelles.....</i>	<i>117</i>
Annexe [7.] : Descriptif méthodologique pour l'analyse des résidus.....	118
Annexe [7.1.] : Tests de normalité	118
Annexe [7.1.bis] : Complément sur les tests de normalité	118
<i>Programme SAS [7.1.] : Tests de normalité des résidus.....</i>	<i>120</i>
Annexe [7.2.] : Tests d'indépendance des résidus	121
<i>Programme SAS [7.2.] : Tests d'indépendance des résidus</i>	<i>121</i>
Annexe [7.3.] : Tests d'homogénéité des variances	122
<i>Programme SAS [7.3.] : Tests d'homogénéité des variances</i>	<i>123</i>
Annexe [8.] : Analyses complémentaires permettant la validation de différentes étapes dans l'élaboration de la prime pure proposée	124
Annexe [8.1.] : Validation de la stabilité des échelles d'un exercice sur l'autre	124
Annexe [8.2.] : Analyse complémentaire sur les résidus du modèle de fréquence	125
Annexe [8.3.] : Analyse des échelles et résidus des coûts moyens sans seuil d'écèlement, ni exclusion de données	127
Annexe [9.] : Rappel des différents types de traités de réassurance	130
Annexe [10.] : Mécanisme des primes de reconstitution au sein du Groupe	131
Annexe [11.] : Programmes d'aide à la décision pour les différentes Entités.....	131
<i>Programme SAS [11.1.] : Somme des échelles de primes pures</i>	<i>131</i>
<i>Programme SAS [11.2.] : Elaboration pour la présentation des résultats.....</i>	<i>131</i>
Annexe [12.] : Comparaison entre le nouveau calibrage de la prime pure et les primes pures brutes observées en 2011 (« Back-Testing »)	132
Annexe [13.] : La distribution Tweedie	133
Annexe [14.] : Les provisions pour égalisation, spécificité de la réglementation française	134
Annexe [15.] : Triangle de liquidation	134
Annexe [16.] : Modèles économiques pour 4 lignes de métiers distincts	136

Section I. Réflexion sur les actuels produits MRH commercialisés au sein du Groupe Groupama

Dans un rapport publié à l'occasion du sommet de Durban¹ sur le climat, les Nations Unies ont indiqué que la planète se réchauffait et que 2011 comptera parmi l'une des années les plus chaudes. De même, selon l'Organisation Météorologique Mondiale (OMM), les 13 années les plus chaudes de l'histoire ont également été enregistrées depuis 1997², favorisant des conditions climatiques extrêmes qui accroissent l'intensité des tempêtes, sécheresses et précipitations.

Autre source d'inquiétude, selon l'OMM³, « les concentrations de gaz à effet de serre dans l'atmosphère ont atteint un nouveau niveau record ». Ces concentrations se rapprochent rapidement, d'un niveau qui entraînerait une hausse de 2 à 2,4°C de la température moyenne mondiale. Une telle hausse pourrait déclencher, selon les différents experts, des modifications de grande ampleur et irréversibles sur la planète.

Dans un produit d'assurance couvrant les risques climatiques, telle que l'assurance Habitation par exemple, il est indispensable d'envisager les effets que de telles perturbations pourraient engendrer sur les résultats économiques d'un Groupe comme Groupama.

Dans cette section introductive, l'objectif est de présenter, brièvement, les enjeux financiers que représentent les produits MRH, en France, au sein du Groupe Groupama, notamment au travers de la garantie « événements climatiques ». Après avoir présenté la sinistralité observée des 4 derniers exercices sur le produit Privatis, la problématique de ce rapport sera exposée. Pour la compréhension du document, quelques définitions seront également proposées.

I.1. Les enjeux de la MRH (au sein de Groupama en France)

Dans la suite de ce rapport, l'étude sera portée essentiellement sur le produit Privatis. Ce produit est uniquement commercialisé par les Caisses Régionales (CR). Il représente près de 68% des contrats en portefeuille des CR et plus de 70% de leurs cotisations MRH.

Afin d'engager le sujet, il semble important d'introduire un bref panorama des enjeux économiques que représentent les produits Multi Risque Habitation au sein du Groupe Groupama en France.

I.1.1. Part de marché du Groupe en France

Les chiffres qui suivent sont extraits de l'« Argus de l'assurance » paru le 18 novembre 2011.

Rang	Assureur	Chiffre d'affaire 2010 (en million d'euros)
1	Groupama**	950,0
2	Axa	873,0
3	Maif	616,0
4	Macif	566,0
5	Allianz	518,0

Tableau 1. 1. Classement des assureurs en 2010 selon leurs chiffres d'affaires

Rang	Assureur	Nombre de contrats fin 2010
1	Macif	4 070 918
2	Axa	4 048 000
3	Groupama**	3 972 170
4	Crédit Agricole Assurances	2 757 802
5	GMF	2 544 607

Tableau 1. 2. Classement des assureurs en 2010 selon le nombre de contrats en portefeuille

Constat :

Avec un montant global de cotisations s'élevant à près d'un milliard d'euros en 2010 (Tableau 1.1.) pour près de 4 millions de contrats en portefeuille (Tableau 1.2.), les produits Multi-Risques Habitation représentent plus de 17% du total des cotisations ABR⁴ de Groupama en France.

Le chapitre qui suit présente succinctement les garanties proposées par le produit Privatis.

¹ Sommet ayant eu lieu en Afrique du Sud entre le 28 novembre et le 11 décembre 2011 réunissant près de 190 pays.

² Année de la conclusion du protocole de Kyoto : 35 pays représentant 30% des émissions à effet de serre avait ratifié un traité conduisant à réduire ces émissions. La Chine et les Etats-Unis, les plus émetteurs, n'avaient pas ratifié ce traité.

³ Cf. Annexe [1.]

** Intégration des périmètres Gan Assurances et Gan Eurocourtage.

⁴ ABR : Assurances de Biens et de Responsabilités (Auto de tourisme 27%, Produits Agricoles 12%, Dommages Entreprises 11%, ACPS 8%, Auto Entreprise 8%, Construction 4%, Autres 13%)

I.1.2. Les garanties du produit Privatis commercialisé par les Caisses Régionales

Le produit Privatis propose diverses garanties aux assurés. Lors de sa souscription, le sociétaire a la possibilité de se couvrir contre :

- les dommages causés aux autres avec la « responsabilité civile vie privée » et/ou la « responsabilité civile propriétaire ou occupant d'immeuble » ;
- la défense des intérêts avec la garantie « Défense - Recours » et/ou la garantie « Protection juridique » ;
- la protection des enfants avec la garantie « Assurance scolaire » ;
- la protection de ses biens avec les garanties :
 - « Incendie et garanties annexes »,
 - « Dommages à l'appareillage électrique »,
 - « Dégâts des eaux »,
 - « Vol »,
 - « Bris de Glaces »,
 - « Bris des vitrages du mobilier usuel »,
 - « Villégiature »,
 - « Objet de valeur »,
 - « Attentats et vandalisme »,
 - « Autres dommages accidentels »,
 - « Catastrophes naturelles »,
 - « Evènements climatiques »,
 - « Frais et perte annexes » ;
- l'« Assistance habitation et aux personnes en déplacement ».

Diverses garanties optionnelles sont également proposées. Elles ne seront pas citées car le poids de ces dernières est relativement faible.

Face à ces garanties de protection des biens (l'essentiel du risque d'un produit MRH), il est présenté dans le point suivant, les charges sinistres qui ont été observées sur les exercices 2007 à 2010.

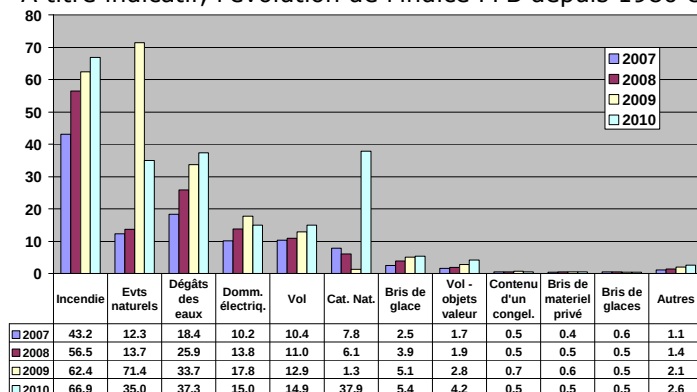
I.1.3. Sinistralité observée sur les 4 derniers exercices pour le produit Privatis

Pour appréhender et percevoir davantage les enjeux économiques et financiers de la garantie « évènements climatiques », la sinistralité observée sur le produit Privatis des 4 derniers exercices, est présentée ci-dessous (uniquement sur les garanties de dommages aux biens).

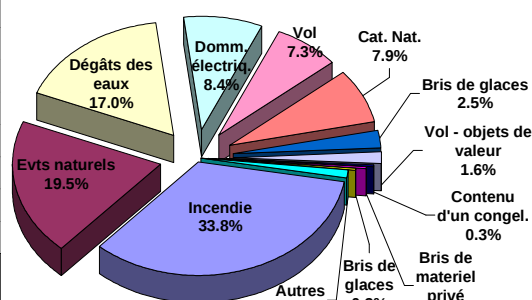
Le périmètre de ces graphiques comprend l'ensemble des Caisses Régionales, excepté Groupama Loire-Bretagne⁵ (GLBR).

Afin de rendre comparable ces différentes années, les charges sinistres sont vues pour chaque exercice *N* en janvier *N*+1 (soit un vieillissement⁶ d'un mois) et actualisées de l'indice de la Fédération Française du Bâtiment (FFB).

A titre indicatif, l'évolution de l'indice FFB depuis 1980 est présentée en *Annexe [2.]*.



Graphe 1. 1. Evolution des charges sinistres par garantie sur les Caisses Régionales (hors GLBR) en million d'euros.



Graphe 1. 2. Répartition de la charge sinistre par garantie des exercices 2007 à 2010.

⁵ Groupama Loire-Bretagne n'étant pas encore passée sous le système SIGMA (système de gestion compagnie), leurs données ne peuvent être extraites.

⁶ Le vieillissement permet d'inclure d'une part, des sinistres qui n'étaient pas déclarés au 31 décembre de l'exercice *N* étudié, et, d'autre part, cela permet de travailler sur une proportion de sinistres plus proche des coûts ultimes. En effet, entre la déclaration d'un sinistre et les différents règlements liés, quelques jours peuvent s'écouler.

Constats :

- De par le *Graphe 1.1.*, il est possible d'observer que les charges sinistres liées aux événements naturels fluctuent dans des proportions relativement importantes entre les exercices.

Si en 2007 et 2008, les charges sinistres étaient partiellement stables, en 2009, l'impact de la tempête Klaus sur la garantie « événements climatiques » se traduit par une charge sinistre de 71,4 millions d'euros (contre 18,7 millions d'euros sur l'exercice précédent). De même en 2010, l'impact de la tempête Xynthia est remarquable sur la garantie « catastrophes naturelles » et la garantie « événements naturels » pour un coût additif, là aussi, de plus de 75 millions d'euros.

- En termes de répartition, sur ces 4 exercices, la charge liée aux aléas naturels représente plus de 27,4% des charges sinistres (*Graphe 1.2.*). Soit le deuxième poste de charge après l'incendie.

Face à ces deux notions « événement climatique » et « catastrophe naturelle », il peut sembler opportun de préciser la méthodologie (ou « règle ») de distinction.

I.1.4. Distinction entre « événement climatique » et « catastrophe naturelle »

En France, l'assurance des risques naturels (hors risques agricoles) est régie par deux systèmes complémentaires. Il y a, d'une part, un dispositif assurantiel contractuel classique pour les risques considérés comme « assurables », et, d'autre part, un second système appelé « régime d'indemnisation des catastrophes naturelles », mis en place par les Pouvoirs publics.

En effet, compte tenu du fort risque d'anti-sélection que présentent les événements cités ci-après et en *Annexe [1.]* (inondations, tremblement de terre, ...), la couverture de ces risques n'a trouvé de réponse assurantielle qu'en rendant la garantie obligatoire à tous les assurés.

Les Pouvoirs publics ont donc légiféré une loi en ce sens le 13 juillet 1982. Ainsi, tous les contrats d'assurance de dommages aux biens (notamment Automobile et Habitation) doivent inclure une garantie « catastrophes naturelles ». Sa tarification est fixée par voie réglementaire. En MRH, il s'agit d'une surprime (identique pour tous les assurés) définie telle que :

Le taux annuel de la cotisation pour la garantie est fixé à 12 % de l'ensemble des primes ou cotisations afférentes au contrat, à l'exception des primes ou cotisations afférentes aux garanties des dommages aux biens résultant d'actes de terrorisme ou d'attentats, lorsque ces primes ou cotisations sont individualisées dans l'avis d'échéance du contrat, des primes ou cotisations afférentes aux garanties de responsabilité civile générale, de protection juridique, d'assistance et de dommages corporels. (Article A.125-2 du Code des Assurances)

Dans un contexte de réforme du système français d'assurance des catastrophes naturelles (système dit « Cat. Nat. »), il semble important d'apporter quelques précisions :

- Tout d'abord, les aléas naturels relèvent de l'état de catastrophe naturelle suivant leur gravité ou leur intensité. Cet état de catastrophe naturelle est constaté par arrêté ministériel. L'indemnisation est prise en charge par l'Etat, via la Caisse Centrale de Réassurance (CCR).

Ces catastrophes naturelles sont définies légalement en France comme étant 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 » (Article L.125-1 du Code des Assurances).

- Deuxièmement, la garantie fait l'objet depuis sa création, d'une obligation d'isolation dans les états comptables réglementaires.

- Enfin, seuls les mouvements de terrain (également appelés « subsidences »), les inondations, les séismes, les raz de marée, les avalanches et, dans les DOM, les cyclones de forces supérieurs à 3 sur l'échelle de Simpson depuis 1999⁷ sont explicitement cités dans le Code des assurances. Depuis 2002, les affaissements de terrains dus à des cavités souterraines et à des marnières (Article L.125-1) ont été ajoutés.

Sur la période 1982-2010, les inondations constituent en France l'événement ayant donné le plus lieu à indemnisations par le régime « Cat. Nat. » (en terme de répartition : 55% d'inondations, 41% de sécheresses, 4% d'autres aléas)⁸.

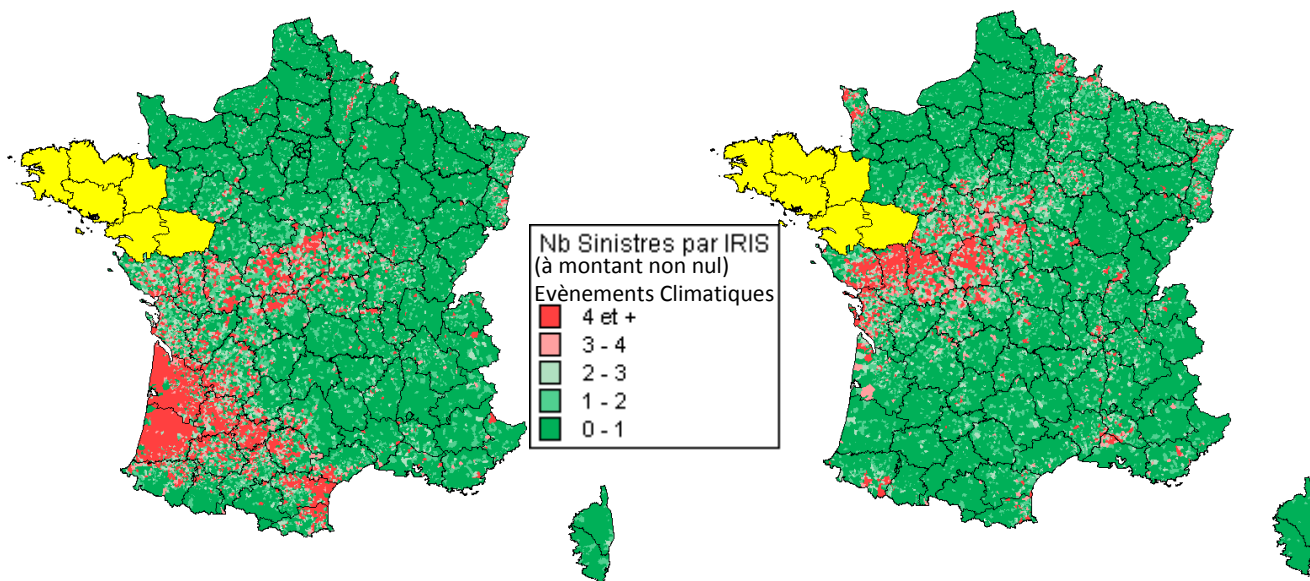
Ce cadre réglementaire posé, la sinistralité géolocalisée, observée sur les exercices 2009 et 2010 pour la garantie « événements climatiques » est brièvement présenté dans le chapitre suivant.

⁷ PARTRAT Christian, BESSON Jean-Luc (2005) *Assurance Non-Vie Modélisation, Simulation*. Economica

⁸ RASPILLER Sébastien (28 juin 2012), *Présentation du projet de réforme du régime d'indemnisation des Catastrophes Naturelles*, Direction Générale du Trésor

I.1.5. Cartographie des sinistres sur les exercices 2009 et 2010 pour la garantie « événements climatiques »

Pour conclure cette introduction, une cartographie des sinistres liés à la couverture de la garantie « événements climatiques » sur les exercices 2009 et 2010 est présentée ci-dessous :



Graphe 1. 3. Nombre de sinistres recensés par IRIS* pour la garantie "Evénements Climatiques" sur l'exercice 2009.

Graphe 1. 4. Nombre de sinistres recensés par IRIS pour la garantie "Evénements Climatiques" sur l'exercice 2010.

Les 5 départements de la Caisse Régionale Groupama Loire-Bretagne (GLBR) ressortent en jaune sur les 2 cartes ci-dessus. Leur sinistralité n'a pu être extraite car GLBR n'est pas encore passée sous le système SIGMA (système de gestion compagnie). Pour plus de précision, les abréviations, libellés et périmètres des Caisses Régionales sont référencés en *Annexe [3.]*.

Constats :

- En 2009, le sud de la France (périmètre de la Caisse Régionale GOC) avait été fortement impacté par le passage de la tempête Klaus. Cela peut s'observer sur le *Graphe 1.3.* par le nombre important d'IRIS à fond rouge.
- De même, sur le *Graphe 1.4.*, il est possible d'observer le passage de la tempête Xynthia en 2010. La Caisse Régionale GCA avait alors été particulièrement impactée. Toutefois, le nombre d'IRIS à fond rouge peut sembler moins important qu'en 2009. En effet, un arrêté ministériel de catastrophe naturelle avait été publié pour certaines communes du bassin atlantique.

Point de transition :

Comme il a été vu, la charge sinistre des événements climatiques et catastrophes naturelles est erratique et impacte considérablement les résultats techniques d'un produit MRH en cas de survenances majeures.

Dans la suite de ce rapport, seule la garantie « événements climatiques » sera étudiée. En effet, le régime des « Cat. Nat. » est d'une part, légiféré, comme il a été vu précédemment, et, d'autre part, les engagements sont limités par la convention existante avec la CCR⁹.

Pour clore définitivement cette introduction, il est précisé ci-après la portée de couverture de la garantie « événements climatiques ».

A ce titre et selon les conditions générales du produit, cette garantie couvre les détériorations accidentelles subies par le bâtiment ou le mobilier usuel résultant de l'action directe :

- du vent ou du choc d'un corps renversé ou projeté par le vent,
- de la grêle sur les toitures,
- du poids de la neige ou de la glace sur les toitures,
- d'une avalanche,
- de glissement ou affaissement de terrain (lorsqu'un arrêté interministériel reconnaît l'état de catastrophes naturelles, c'est la garantie « catastrophes naturelles » qui intervient).

⁹ CCR : Caisse Centrale de Réassurance

* La définition de l'IRIS est renseignée au chapitre IV.3.

Ce rapide diagnostic permet d'introduire le sujet principal de ce rapport. Celui-ci est décrit dans ses différentes étapes, et est complété de quelques définitions, nécessaires à une meilleure compréhension, par les deux chapitres suivants.

I.2. Enoncé de la problématique

L'approche actuarielle classique se décompose généralement en deux temps :

- Tout d'abord, à partir de données historiques, une estimation des primes pures espérées est réalisée.
- Ensuite, ces primes pures sont actualisées à la fois d'une « projection » à l'ultime (grâce à l'utilisation de triangle de liquidation¹⁰) et d'un indice d'inflation. En effet, les données historiques ne fournissent une estimation des paiements espérés futurs que si les règlements (sommes des primes pures espérées et actualisées) sont les mêmes d'un exercice à l'autre.

Les incertitudes liées au dérèglement climatique de la planète et les conséquences sur la fréquence et l'ampleur des événements futurs rendent délicate cette approche tarifaire pour des garanties type « événements climatiques ».

En effet, même avec une connaissance approfondie du passé, le dérèglement climatique ne laisse pas supposé un comportement identique dans les années à venir.

« La volatilité de ce type de risque en fait l'essentiel même du coût »¹¹.

Pour répondre à cette problématique de volatilité, l'approche tarifaire décrite dans ce rapport s'articulera essentiellement en 4 étapes :

- La première étape consistera à définir une échelle de coût du risque (ou échelle de prime pure, selon la littérature), basée sur les exercices 2009-2010 et selon les critères tarifaires de l'actuel produit « Privatis ». En effet, pour refléter au mieux un « comportement », il est préférable de se baser sur les exercices les plus récents. Pour ce faire, les méthodes de régression (type *GLM* : Modèle Linéaire Généralisé) seront utilisées. Au final, sur chaque contrat, et selon les modalités tarifaires qui le composent, une échelle de risque globale sera calculée.

Les trois étapes qui suivent consisteront à déterminer le « niveau » de cette prime pure. Ce niveau sera calculé à partir des modélisations de charges sinistres espérées, via les travaux réalisés par la Direction Actuariat Groupe, dans le module risque « CAT » du modèle interne.

- Des événements attritionnels et atypiques (ayant une probabilité de retour d'une fois tous les 200 ans, comme l'exige la Directive Solvabilité 2) seront modélisés. Des lois de fréquences et d'intensités pour ces événements seront établies sur des bases relativement théoriques, en utilisant le logiciel de modélisation de catastrophes (RMS¹²). Compte tenu des probabilités liées à chaque événement, une espérance de charge est ainsi obtenue.

- Le programme de réassurance sera appliqué sur chaque événement précédemment modélisé afin d'en réduire l'impact et la volatilité. Le coût de la facture de réassurance devant également être financé, il sera réinjecté via un « taux de chargement ».

- Pour terminer, l'espérance du coût des sinistres (réassurance incluse) sera divisée par la somme des échelles de risque global définies à la première étape. Cette valeur calculée sera assimilée à un coefficient multiplicatif et permettra de déterminer la prime pure de chaque profil de risque.

Chaque Entité aura ainsi :

- une mutualisation robuste des échelles de risque au niveau national,
- une prime pure reflétant le risque porté par son portefeuille.

Avant de rentrer définitivement dans le sujet (Section II.) et comme il l'a été noté auparavant, quelques définitions sont précisées dans le chapitre qui suit.

¹⁰ Une illustration des triangles de liquidation est proposée en Annexe [15].

¹¹ PARTRAT Christian, BESSON Jean-Luc (2005), *Assurance Non-Vie Modélisation, Simulation*. Economica

¹² RMS : Logiciel permettant d'appréhender de manière globale l'exposition d'un portefeuille par rapport à un péril climatique. Ce logiciel a été « validé » par l'ensemble des réassureurs.

I.3. Définitions

Avant de poursuivre, il est important de rappeler ou préciser quelques termes qui seront employés dans la suite de ce document :

- Entité assurée : Une entité assurée peut aussi bien décrire une habitation, une automobile, un homme, une femme, un chien, un chat, ... L'entité assurée est explicitement rédigée dans les conditions générales et/ou particulières d'un contrat d'assurance.

- Contrat imputé (police imputée ou exposition) : Les différents contrats en portefeuille sont proratisés à leur temps de présence sur l'année. Par exemple, l'exposition pour un contrat souscrit au 1^{er} juillet d'une année *N* et résilié au 1^{er} octobre de la même année, vaut 0,25 (3 mois de présence sur 12).

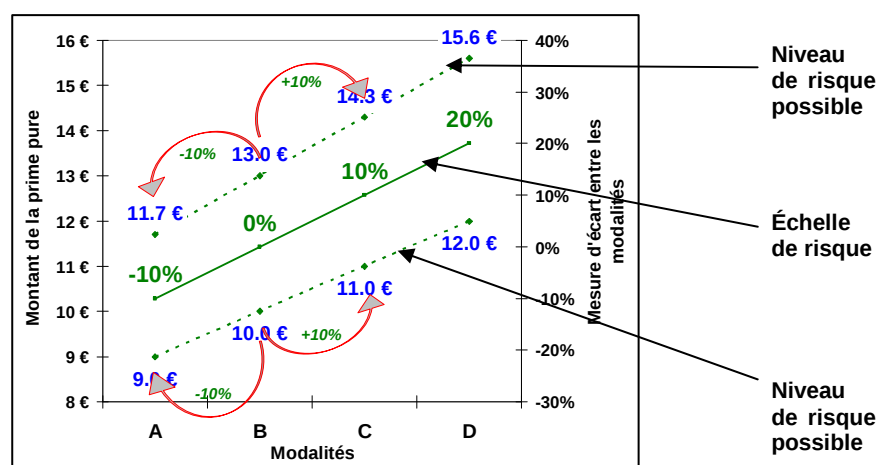
- Fréquence : La fréquence de sinistres d'un exercice correspond au ratio « Nombre de sinistres » divisé par le « Nombre de contrats imputés ». Plus précisément, cette fréquence correspond à l'ensemble des déclarations de sinistres qu'effectuent les sociétaires au cours d'un exercice. Il est important d'insister sur le mot « déclaration » qui reflète un aspect comportemental des assurés. En effet, en exemple, un sociétaire peut être amené ou non à déclarer un sinistre selon la franchise appliquée à son contrat.¹³

- Coût moyen : Le coût moyen correspond à la somme des charges sinistres de l'année divisée par le nombre de sinistres total.

- Prime pure : La prime pure, ou coût du risque selon la littérature, correspond à la multiplication de la fréquence et du coût moyen (par simplification, c'est également la somme des charges sinistres de l'année divisée par la somme des contrats imputés). La prime pure se décrit comme étant le montant des sinistres moyens auxquels devra faire face l'assureur pour une entité assurée donnée. En sommant les primes pures de chaque entité assurée, l'assureur doit être en mesure de payer l'ensemble des charges sinistres cumulées tout au long de l'année.

- Échelle de risque : Il est entendu par échelle de risque, une mesure d'écart entre les différentes modalités d'un critère tarifaire et une modalité de référence (généralement la modalité la plus représentée du critère tarifaire étudié). Sur le *Graphe 1.5.* ci-dessous, en considérant la modalité *B* comme la modalité de référence, il est possible de conclure que la modalité *C* a un coût du risque 10% plus important que la modalité *B*.

- Niveau (ou point référent) et coefficient multiplicateur : Les échelles de risque donnent un comportement à la prime pure, mais ne donnent nullement les montants en euros associés. Sur ce *Graphe 1.5.*, il est donné en exemple deux niveaux de risque possible (courbes en pointillés). Suivant le coût du risque de la modalité *B* (le point référent) et la valeur associée (le coefficient multiplicateur), ici 10€ ou 13€, le risque *C* a respectivement une prime pure de 11€ ou 14,3€.



Graphe 1. 5. Graphe de synthèse des différentes définitions

¹³ De plus en plus, l'assurance est considérée comme un bien de consommation par les différents assurés. En effet, l'assuré qui cotise depuis plusieurs années, va demander à bénéficier d'avantages ou de remboursements par rapport à des sinistres, même si ces derniers ne figurent pas dans les Conditions Générales (CG) ou Conditions Particulières (CP).

Section II. Lien entre la tarification de la garantie « évènements climatiques » et le modèle interne : détermination et analyse de la prime pure

Bien souvent, à partir des données observées et actualisées d'un portefeuille (pour un historique de 2 à 3 exercices selon la quantité d'informations), une tarification peut être élaborée.

En MRH, cette pratique est relativement commune pour des garanties à fréquence élevée telles que les garanties « Dégâts des eaux », « Bris des glaces », « Dommages à l'appareillage électrique ». De plus, les charges globales sont relativement stables d'un exercice à l'autre, y compris pour la garantie « Incendie » (voir *Graphique 1.1.*). La garantie « évènements climatiques » est bien la seule garantie qui possède de la volatilité (tant en termes de fréquence qu'en coût).

Or, comme évoqué précédemment, l'incertitude liée au dérèglement climatique de la planète et ses effets sur la fréquence et l'ampleur des évènements futurs (ainsi que leur localisation) n'en rend que plus complexe l'approche tarifaire. La seule connaissance des statistiques sur les derniers exercices s'avérerait d'une utilité limitée pour l'évaluation du coût du risque.

Afin de répondre à la problématique, la Section II. se découpe en 5 chapitres :

- Dans le chapitre II.1., il est présenté la méthodologie de détermination des échelles de primes pures selon l'algorithme de l'actuel produit ;
- Les modélisations effectuées, dans le modèle interne, sur le risque tempête seront présentées dans le chapitre II.2.. Ces modélisations sont réalisées sur l'ensemble des risques couverts par le traité de réassurance « Force de la Nature » ;
- Dans le chapitre II.3., le programme de réassurance sera appliqué sur les évènements extrêmes modélisés dans le chapitre précédent ;
- Les calculs ayant été effectués sur l'ensemble des risques couverts par le traité, des clés de répartition seront appliquées et exposées dans le chapitre II.4. pour ne cibler que le produit Privatis ;
- Enfin, pour clore cette section, le chapitre II.5. liera les différents travaux réalisés et permettra ainsi de fixer la prime pure (objectif de l'étude).

II.1. Détermination de l'échelle tarifaire de la garantie « évènements climatiques » selon les critères tarifaires du produit Privatis

Sur le produit Privatis, la tarification de la garantie « évènements climatiques » est définie selon l'interaction de trois variables principales :

- le nombre de pièces,
- le type de bâtiment,
- la limite de capital mobilier garantie.

Des coefficients de « modulations » sont ensuite appliqués sur ce tarif de base. En l'occurrence, ces coefficients dépendent des critères tarifaires suivants :

- la superficie de dépendance,
- la formule,
- la franchise,
- l'absence de mobilier (ou mobilier assuré par ailleurs),
- si le bâtiment est assuré par ailleurs,
- la présence de domotique,
- la présence de piscine,
- la qualité de l'assuré,
- le rééquipement à neuf (abrégé RAN).

A ces facteurs explicatifs, deux autres variables seront ajoutées :

- la Caisse Régionale,
- l'exercice de survenance.

Ces deux variables supplémentaires n'ont aucune influence sur les échelles tarifaires des autres variables (complète indépendance avec les autres critères). Néanmoins, elles permettent avant tout de capter (détecter), pour la première variable, une possible influence géographique et pour la seconde, d'être cohérente avec une fréquence de sinistre annuelle.

N.B. : La tarification actuelle du produit impose un « comportement » identique quelle que soit la géolocalisation du risque. A noter, tout de même, que chaque Caisse Régionale a la possibilité d'élaborer son niveau tarifaire selon la constitution de son portefeuille.

Avant toute chose, il est bien évidemment nécessaire de se consolider une base de données. Le logiciel SAS a été utilisé à cet effet. Cette base de données est construite à partir des exercices 2009 et 2010 avec un vieillissement¹ de 2 mois et comprend 10 des 11 Caisses Régionales². Au total, se sont 2,7 millions de contrats qui seront analysés.

Dans une approche actuarielle classique, un calibrage de prime pure se réalise généralement en deux temps :

- en premier lieu, une analyse des fréquences de déclarations des sinistres,
- suivi par une étude des coûts moyens.

La prime pure est alors calibrée par le produit des deux analyses précédemment réalisées (cf. chapitre I.3.).

Pour ce faire, les fréquences de déclarations des sinistres et les coûts moyens seront analysés et modélisés grâce aux techniques de régression (dites également *GLM* : Modèle Linéaire Généralisé) définies ci-dessous.

La régression (ou modèle) linéaire généralisée³

Le but général de la régression est d'expliquer au mieux une grandeur Y (la réponse), en fonction d'autres grandeurs X (matrice de variables explicatives, appelées également « régresseurs », ou bien encore « facteurs »).

Si y_i désigne une des observations et $X_i=(x_{i1},...,x_{in})$ représente n données (variables explicatives) liées à cette observation, alors il est possible d'expliquer y_i par diverses équations (modèles).

Par exemple, avec β_i désignant des coefficients spécifiques appliqués à chaque modalité des variables explicatives, y_i peut être exprimée telle que :

- $y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_n x_{in}$ ou bien,
- $y_i = \beta_0 + \beta_1 \log(x_{i1}) + \dots + \beta_n \log(x_{in})$ ou bien encore, un modèle mixte tel que
- $y_i = \beta_0 + \beta_1 \log(x_{i1}) + \dots + \beta_j x_{ij} x_{ij+1} + \dots + \beta_{n-1} \exp(x_{in})$.

Dans le cadre de cette étude, les variables à expliquer Y sont :

- des nombres naturels en ce qui concerne les nombres de sinistres,
- des nombres réels positifs en ce qui concerne les coûts moyens.

Le modèle linéaire généralisé permet, via une famille de distributions de probabilités spécifiques ainsi qu'une fonction réelle $\eta \rightarrow g(\eta)$, dite fonction de lien, de modéliser la grandeur Y .

Cette fonction de lien permet de définir la relation entre les variables réponses (fréquences, coûts moyens) et la combinaison linéaire des variables de tarification.

Dans le cadre de ce projet, ce sont les familles exponentielles (avec un paramètre de nuisance) qui seront sollicitées en raison des propriétés très avantageuses qu'elles possèdent.

Par définition, une distribution de probabilités appartient à la famille exponentielle si, pour une mesure de Lebesgue ou de comptage désigné par $\nu(dy)$ sur $\mathbb{R}$, sa loi f peut s'exprimer sous la

forme $f(y|\theta, \phi) = \exp\left\{\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)\right\}$ telle que $f(y|\theta, \phi)\nu(dy)$ forme une distribution de

probabilité pour tout ϕ variant dans un certain domaine de $\mathbb{R}_+$, avec θ désignant le profil de risque et ϕ le paramètre de nuisance.

Le terme $c(y, \phi)$ contient toute la partie non-linéaire en y non déjà présente dans $\nu(dy)$ permettant, dans la pratique, de définir la variable ϕ de la paramétrisation, puis ensuite θ .

ϕ fixé, il vient $E[y] = b'(\theta)$ et $Var(y) = \phi.b''(\theta)$.

¹ Pour chaque exercice N , les sinistres sont vus en février de l'exercice $N+1$

² Groupama Loire-Bretagne n'étant pas encore passée sous le système SIGMA (système de gestion compagnie), leurs données ne peuvent être extraites.

³ Extrait du cours de *Régression*, Bernard Delyon, Université de Rennes 1. Toute la théorie ne sera pas développée ici.

N.B. : b' désigne la dérivée première de la fonction b en θ et b'' désigne la dérivée seconde en θ .

De même, par définition, un modèle linéaire généralisé pour un jeu de données $(x_i, y_i)_{i=1, \dots, n}$ est une distribution pour la suite (y_i) déterminée par :

- une famille exponentielle à un paramètre de nuisance ϕ
- une fonction r (dont l'inverse g est appelé fonction de lien), g définie telle que $g(r(x)) = x$,
- une valeur ϕ et un vecteur de régression⁴ β ,

avec les propriétés suivantes (les régresseurs x_i sont déterministes) :

- indépendance des y_i ,
- y_i suit la loi $f(y|\theta, \phi)v(dy)$
- $b'(\theta_i) = r(x_i \cdot \beta)$

La dernière relation permet ainsi de déterminer θ_i en fonction de $x_i \beta$.

De plus, la fonction b' est inversible, en raison de la stricte convexité de b . Il vient en particulier :

$$E[y_i] = r(x_i \beta)$$

$$\text{Var}(y_i) = \phi \cdot \text{Var}(r(x_i \beta))$$

$$y_i \in \text{Supp}(v)$$

Ce point théorique étant à présent clos, les différentes étapes et méthodologies employées pour déterminer les échelles de primes pures sont décrites dans le tableau ci-dessous :

	Objectif	Méthodologie	Résultats / et hypothèses à valider	Validation des hypothèses
Détermination des échelles de primes pures	<u>Étape 1 :</u> Vérification de l'indépendance des données analysées	- Comparer les distributions empiriques selon le nombre de sinistres déclarés par contrat ; - Comparer les valeurs des coûts moyens observés selon le nombre de sinistres déclarés par contrat ; - Analyser l'indépendance des variables tarifaire.	<u>Résultat :</u> Si aucun biais n'est observé, alors l'indépendance des données est établie. <u>Hypothèse émise :</u> Les échelles sont stables d'un exercice à l'autre.	
	<u>Étape 2 :</u> Détermination des échelles de risque de fréquences selon les critères tarifaires	- Vérifier la significativité des facteurs de l'actuel algorithme tarifaire ; - Choisir le modèle le plus adapté à la modélisation de la fréquence.	<u>Résultat :</u> Échelles de fréquences déterminées <u>Hypothèse du modèle à valider :</u> Étude sur les résidus	- Vérifier l'indépendance des résidus ; - Vérifier l'homoscédasticité des résidus ; - Vérifier la normalité des résidus.
	<u>Étape 3 :</u> Détermination des échelles de risque des coûts moyens selon les critères tarifaires	- Définir un seuil d'exclusion entre les sinistres attritionnels et les sinistres non-récurrents ; - Vérifier la significativité des facteurs ; - Choisir le modèle le plus adapté à la modélisation des coûts moyens.	<u>Résultat :</u> Échelles des coûts moyens déterminées <u>Hypothèse du modèle à valider :</u> Étude sur les résidus	- Vérifier l'indépendance des résidus ; - Vérifier l'homoscédasticité des résidus ; - Vérifier la normalité des résidus.
	<u>Étape 4 :</u> Dédution des échelles de primes pures	- Multiplication des échelles précédemment modélisées.		- Confirmer la stabilité des échelles de risques (hypothèse validée en Annexe [8.1.] par l'utilisation de l'exercice 2011).

Tableau 2. 1. Méthodologie de détermination des échelles de primes pures

Avant de réaliser la modélisation des échelles de fréquences et coûts moyens, il est donc nécessaire d'établir l'indépendance des observations entre exercices. Ainsi, l'hypothèse de stabilité des échelles de risques d'une année à l'autre pourra être émise.

⁴ β est l'estimateur du maximum de vraisemblance des moindres carrés ordinaires : Voir Annexe [4.1.]

II.1.1. Vérification de l'indépendance des données analysées

Pour rappel, la fréquence de déclarations des sinistres d'un exercice correspond au ratio « Nombre de sinistres » divisé par la somme des contrats imputés. Le coût moyen quant à lui correspond à la somme des charges sinistres de l'année divisée par le nombre de sinistres total.

Sur la base étudiée, la fréquence de sinistres de la garantie « événements climatiques » est de 1,75% en 2009 et de 1,60% en 2010. Sur la base agrégée, il ressort une fréquence de 1,668%.

Pour ces deux exercices, le nombre de sinistres par contrat est réparti comme suit :

Nombre de sinistres par contrat :	0	1	2	3
Effectifs recensés (données 2009)	1 270 806	21 995	514	12
Effectifs recensés (données 2010)	1 374 976	21 739	502	18

Tableau 2. 2. Répartition du nombre de sinistres référencés par contrat sur la garantie "événements climatiques"

Interprétation :

Pour la colonne « 0 » du Tableau 2.2., il s'agit de l'exposition des contrats. Pour les colonnes « 1 », « 2 » et « 3 », il s'agit du nombre de sinistres recensés. En lecture, à titre d'exemple, en 2009, il n'y a eu que 4 contrats qui ont eu 3 sinistres ($4 \times 3 = 12$).

Le but de ce chapitre est :

- D'une part, de vérifier que les contrats recensés avec plus de 1 sinistre ont des montants de règlements homogènes aux contrats ayant un unique sinistre et ce, pour chaque exercice. Il s'agit ainsi de confirmer la stabilité dans la répartition et le coût des sinistres ;
- D'autre part, de vérifier l'indépendance des critères tarifaires.

(a) Stabilité dans la répartition des sinistres d'un exercice à l'autre

Afin de répondre au premier point, la méthodologie employée va se dérouler en deux temps :

- Premièrement, une comparaison des distributions empiriques selon le nombre de sinistres déclarés par contrat doit être réalisée ;
- Deuxièmement, il est nécessaire de comparer les coûts moyens observés selon le nombre de sinistres déclarés au sein d'une même « colonne » conformément au Tableau 2.2..

Avant de poursuivre et afin d'avoir une homogénéité dans les résultats, il est important de préciser que les charges sinistres des différents exercices ont été actualisées⁵.

	Comparaison des distributions	Données empiriques et commentaires
Contrat ayant eu un sinistre		Coût Moyen 2009 des sinistres < 5700€ avec exclusion des sinistres à montant nul : 1473,23€ N.B. : 17,0% des sinistres ont un montant nul. Coût Moyen 2010 des sinistres < 5700€ avec exclusion des sinistres à montant nul : 1451,26€ N.B. : 29,8% des sinistres ont un montant nul. L'écart observé n'est que de 5€, soit un écart de moins de 0,4%. L'inclusion de ces sinistres à montant non nul aurait biaisé l'estimation des coûts moyens.
Contrat ayant eu 2 sinistres		Coût Moyen 2009 des sinistres < 5700€ avec exclusion des sinistres à montant nul : 1544,61€ 29,5% des sinistres ont un montant nul. Coût Moyen 2010 des sinistres < 5700€ avec exclusion des sinistres à montant nul : 1447,39€ 35,1% des sinistres ont un montant nul. L'écart de coût moyen n'est plus de 5€ mais de 50€, ce qui représente un écart de près de 4%.

⁵ cf Annexe [2.] : Indice FFB de 1980 à fin 2011

Contrat ayant eu 3 sinistres	Sachant qu'il n'y a eu que 12 sinistres en 2009 et 18 sinistres en 2010, la comparaison des données n'est pas possible. A noter tout de même lors de l'étude de ces sinistres, que les montants affichés étaient relativement importants. Les coûts (actualisés) sur ces 2 exercices étaient compris dans une fourchette de 1 274€ à 20 497€.
------------------------------	--

Tableau 2. 3. Analyse de la stabilité dans la répartition des sinistres

Constats :

- Les deux histogrammes ci-dessus ont une allure relativement similaire.
 - Les coûts moyens 2009 et 2010 sont très proches, aussi bien sur les contrats ayant eu 1 sinistre (moins de 0,4%) que les contrats ayant eu 2 sinistres (moins de 4%).
- En conclusion, quels que soient le nombre de sinistres et l'amplitude des événements sur un exercice, la distribution des sinistres reste stable.

N.B. : Ce constat n'aurait pu être vu sans l'exclusion des sinistres à montant nul et sans l'exclusion des sinistres ayant un coût supérieur à 5700€ (réf. paragraphe b) du chapitre II.1.3.).

Précision : Sur les deux histogrammes précédents, le nombre de sinistres à montant nul est plus important en 2010 qu'en 2009. Ceci est dû à une pratique de gestion propre à Groupama. En effet, lors de la tempête Xynthia, la volonté du Groupe a été de satisfaire et d'indemniser rapidement les assurés. De ce fait, tous les sinistres déclarés ont d'abord été renseignés dans la garantie « événements climatiques ». Par la suite, un décret « Cat. Nat. » a été appliqué sur certaines communes. Les déclarations de sinistres saisis au préalable sur la garantie « événements climatiques » ont alors été remis sur la garantie « catastrophes naturelles ». D'où cette proportion supérieure de sinistres à montant nul sur la garantie « événements climatiques » en 2010 (aspects remarquables sur les graphes 1.3. et 1.4 et tableaux 2.3. et 2.4.).

Taux de sinistres sans suite	2009	2010	Agrégat des 2 exercices
1 sinistre par contrat	17,0%	29,8%	23,5%
2 sinistres par contrat	29,5%	35,1%	28,3%
Total	17,3%	30,0%	23,6%

Tableau 2. 4. Taux de sinistres à montant nul sur les exercices 2009 et 2010

Rappel objectif	Méthodologie employée	Point retenu	Étape suivante
Stabilité dans la répartition des sinistres d'un exercice à l'autre	- Comparer les distributions empiriques selon le nombre de sinistres déclarés par contrat - Comparer les valeurs des coûts moyens observés selon le nombre de sinistres déclarés par contrat	Quels que soient les sinistres et l'amplitude des événements sur un exercice, la distribution des sinistres reste stable.	Tests d'indépendance des variables tarifaires

Avant de poursuivre et de déterminer le modèle le plus approprié à la modélisation de la fréquence des sinistres, il est également nécessaire de tester⁶ l'indépendance des variables tarifaires.

(b) Tests d'indépendance des variables tarifaires

La qualité des données est une démarche inhérente à la gestion du risque d'entreprise et qui a été renforcée par la Directive Solvabilité 2.

Même si ce n'est pas le but de ce mémoire, il peut être bon de préciser que des procédures sur la qualité des données ont été réalisées. Ce type d'analyse permet de remarquer rapidement des anomalies, notamment la présence de valeurs manquantes ou la création de nouvelles modalités. Une simple analyse descriptive univariée permet également d'identifier des tendances de fond.

Sur la base de tables de contingence des expositions, des tests du Khi-Deux (ou Khi-Carrés ou χ^2) ont été utilisés⁷ afin de déterminer l'hypothèse d'indépendance entre les variables.

Ces tests consistent à réaliser une comparaison entre la distribution statistique observée dans l'échantillon et une distribution théorique. Cette distribution théorique s'obtient en supposant l'indépendance des deux variables considérées (H_0 : hypothèse d'indépendance des variables). Les tests permettent ainsi de comparer si les écarts entre ces deux distributions sont assimilables aux fluctuations d'échantillonnage, ou si au contraire, les écarts sont trop importants pour que l'hypothèse H_0 soit acceptée.

⁶ cf. Annexe [5.] : Descriptif des tests d'indépendance et de significativité

⁷ Annexe [5.1.] : Test du Khi-Deux

Dans un premier temps, l'étude des tables de contingence a été réalisée sur la base agrégée des exercices 2009 et 2010 (selon les critères tarifaires présentés dans le *Tableau 2.11.*).

Face à des résultats relativement surprenants, une attention particulière a été portée sur la statistique du Khi-Deux. En effet, avec un échantillon important, il est toujours possible d'obtenir un test du Khi-Deux significatif (c'est à dire dépassant les valeurs critiques de la table à 5% ou 1%).

Par exemple, si tous les effectifs des cellules d'un tableau de contingence sont multipliés par 100, alors la statistique du Khi-Deux est également multipliée par 100 (sans pour autant que les liaisons n'aient été changées). Sachant que la base de données comprend 2,7 millions de contrats, le test du Khi-Deux est vraisemblablement biaisé.

Pour palier cette spécificité⁸, d'autres statistiques peuvent être lues avec le logiciel SAS⁹ :

- le Test du rapport de vraisemblance,
- le coefficient ϕ^2 de Pearson (noté « Coefficient Phi » dans la sortie SAS),
- le coefficient de contingence C , ou bien encore,
- le V de Cramer.

Malgré tout, leur lecture est délicate, car la loi de ces tests n'est pas connue de façon théorique.

A titre d'exemple, sur la base totale, il est présenté ci-dessous, le tableau des statistiques lié à l'analyse d'indépendance entre la variable « Type de bâtiment » (CTNBI) et « Nombre de pièces » (QNNBP).

Statistiques pour la table de QNNBP par CTNBI

Statistique	DDL	Valeur	Prob
Khi-2	11	747385	<.0001
Test du rapport de vraisemblance	11	764693	<.0001
Khi-2 de Mantel-Haenszel	1	568393	<.0001
Coefficient Phi		0.53054	
Coefficient de contingence		0.46866	
V de Cramer		0.53054	

Tableau 2. 5. Statistiques d'indépendance sur la base

Constats :

- Sur ce *Tableau 2.5.*, selon le test du Khi-Deux et le Test du rapport de vraisemblance, les deux variables « Type de bâtiment » et « Nombre de pièces » sont *a priori* significativement indépendantes.

- Dans le même temps, le coefficient ϕ^2 de Pearson, le coefficient de contingence C et le V de Cramer ont des valeurs aux alentours des 0,5 (Pour ces statistiques, plus les valeurs sont proches de 0, plus les critères sont, *a priori*, indépendants).

Avec ces résultats opposés, il est relativement difficile de conclure à l'indépendance de tels critères.

Afin de combler les biais de ces tests et faire « converger » les statistiques, des échantillons, extraits de la base de données, ont été créés.

Pour obtenir un test du Khi-Deux fiable et valide, trois conditions sont nécessaires :

- les échantillons doivent être représentatifs de la population totale,
- la taille de l'échantillon ne doit pas dépasser les 32 000 observations (contrainte SAS),
- le nombre de contrats à extraire est déterminé de telle sorte que dans les futurs tableaux de contingence, les effectifs de toutes les cases soient au moins égaux à 5.

Une fois ces diverses extractions réalisées¹⁰, les procédures de tests ont de nouveau été effectuées.

Cette méthodologie a été relativement longue, car l'ensemble des croisements possibles génère 66 combinaisons ($= C_{12}^2$, 2 variables parmi 12 critères tarifaires possibles. Les critères « Caisse Régionale » et « Exercice » ayant été exclus de l'analyse).

Pour optimiser au mieux cette analyse, il est semble opportun de préciser qu'une table de données agrégée a été créée, via une « *proc means* » sous SAS¹¹. Cette agrégation permet de réduire le temps de traitement sans pour autant nuire à la qualité de l'information.

⁸ Annexe [5.1.bis] : Les limites du test d'indépendance du Khi-Deux

⁹ Annexe : Programme SAS [5.1.] : Test d'indépendance

¹⁰ Annexe : Programme SAS [5.2.] : Extraction d'un échantillon aléatoire

¹¹ Annexe : Programme SAS [5.3.] : Constitution d'une base agrégée pour un gain de traitement

Le nombre de modalités étant relativement important pour certaines variables (comme le critère « Nombre de pièces »), il n'a pas semblé nécessaire de rapporter l'ensemble des tableaux.

Sur la base totale, à titre d'exemple et pour sa lisibilité, seuls les résultats du croisement « Type de bâtiment » (CTNBI) et « Domotique » (CINBF) sont décrits dans le tableau ci-dessous :

Table de CINBF par CTNBI

CINBF(CINBF)	CTNBI(CTNBI)		
Fréquence Attendu Pourcentage Pctage en ligne Pctage en col.	C	P	Total
N	674450 671100 25.40 25.62 99.65	1958450 1.96E6 73.76 74.38 98.99	2632900 99.16
O	2361 5710.6 0.09 10.54 0.35	20043 16693 0.75 89.46 1.01	22404 0.84
Total	676811 25.49	1978493 74.51	2655304 100.00

Tableau 2. 7. Comparaison effectif théorique / effectif empirique sur la base totale

Statistiques pour la table de CINBF par CTNBI

Statistique	DDL	Valeur	Prob
Khi-2	1	2659.2337	<.0001
Test du rapport de vraisemblance	1	3182.0843	<.0001
Khi-2 continuité ajustée	1	2658.4398	<.0001
Khi-2 de Mantel-Haenszel	1	2659.2327	<.0001
Coefficient Phi		0.0316	
Coefficient de contingence		0.0316	
V de Cramer		0.0316	

Tableau 2. 6. Statistiques d'indépendance sur la base totale

Constats :

- Sur le Tableau 2.7., il est possible de remarquer qu'il y a près de 75% d'habitations individuelles (modalité « P ») et plus de 25% d'habitations collectives (modalité « C ») dans le portefeuille du produit Privatis.

Pour ce qui est de la « Domotique », plus de 99% des habitations n'en sont pas équipées.

- Sur le Tableau 2.6., les tests ϕ^2 de Person, le coefficient de Contingence et le V de Cramer sont proches de zéro. Il en est de même pour les statistiques du Khi-Deux. Mais ces tests peuvent être biaisés par rapport à l'effectif du portefeuille étudié.

Sur un des échantillons testé, il ressort pour ces deux mêmes variables les statistiques suivantes :

Table de CINBF par CTNBI

CINBF(CINBF)	CTNBI(CTNBI)		
Fréquence Attendu Pourcentage Pctage en ligne Pctage en col.	P	C	Total
N	16760 16783 74.03 74.61 99.08	5702 5679.2 25.19 25.39 99.62	22462 99.22
O	155 132.25 0.68 87.57 0.92	22 44.752 0.10 12.43 0.38	177 0.78
Total	16915 74.72	5724 25.28	22639 100.00

Tableau 2. 9. Comparaison effectif théorique / effectif empirique sur un des échantillons test

Statistiques pour la table de CINBF par CTNBI

Statistique	DDL	Valeur	Prob
Khi-2	1	15.6038	<.0001
Test du rapport de vraisemblance	1	18.0895	<.0001
Khi-2 continuité ajustée	1	14.9255	0.0001
Khi-2 de Mantel-Haenszel	1	15.6031	<.0001
Coefficient Phi		-0.0263	
Coefficient de contingence		0.0262	
V de Cramer		-0.0263	

Tableau 2. 8. Statistiques d'indépendance sur un des échantillons test

Constats :

- Hormis un ordonnancement différent des modalités, dans le Tableau 2.9., la répartition de l'échantillon est relativement proche de celle de la population totale.

- Les tests d'indépendance énumérés dans le Tableau 2.8. sont tous significatifs. L'hypothèse d'indépendance est donc validée (les variables « Type de bâtiment » et « Domotique » sont significativement indépendantes).

Cette démarche a été réalisée de la même façon pour l'ensemble des critères.

De cette analyse complète, deux points en sont ressortis :

- Une dépendance a été constatée entre le type de bâtiment et le nombre de pièces (test du Khi-Deux significatif). Ce qui n'est pas « curieux ». En effet, en termes de répartition, il est assez commun de retrouver une quantité importante d'appartements de 1 à 2 pièces, et de ne pas trouver un effectif important avec des maisons de 1 à 2 pièces.
De la même façon, des maisons de 4 à 6 pièces sont relativement fréquentes, ce qui n'est pas le cas pour les appartements.
Pour ces raisons, une nouvelle variable, croisement du type d'habitation et du nombre de pièces, a été créée.
Cette variable a été étudiée dans la modélisation des fréquences de sinistres et coûts moyens. Les résultats étaient assez intéressants et « surprenants » en ce sens, que les échelles de risques étaient relativement similaires aux résultats des non croisements de ces variables. Par conséquent, pour plus de stabilité, il a été décidé de poursuivre les travaux en laissant dissociées les variables « type d'habitation » et « nombre de pièces ».
- Pour le reste, il y a complète indépendance avec/et entre les autres critères tarifaires.

Rappel objectif	Méthodologie employée	Point retenu
Tests d'indépendance des variables tarifaires	Comparaison entre la distribution statistique observée dans l'échantillon et la distribution théorique	Chaque variable sera considérée comme indépendante pour la suite de l'étude, y compris les deux critères « nombre de pièces » et « type de bâtiment »

Point de transition :

Quels que soient le nombre de sinistres et l'amplitude des événements sur un exercice, la distribution des sinistres semble régulière. C'est pourquoi, toute chose étant égale par ailleurs, il est posé l'hypothèse suivante : les échelles de risque restent stables d'un exercice à l'autre.

Lors de la rédaction de ce rapport, seuls les exercices 2009 et 2010 étaient à disposition. En mai, une nouvelle base de données exercice 2011, vieillie à 2 mois, était à disposition. A partir de cette base, l'hypothèse de stabilité des échelles de risques a été renforcée. Ce constat est illustré en *Annexe [8.1.]*.

A présent que l'étude d'indépendance des données est achevée, la détermination des échelles de risques (fréquence et coût moyen) va pouvoir être réalisée.

A la fin de chaque chapitre, un tableau synthétique dressera ce qui a été effectué, tout comme le *Tableau 2.10.* ci-dessous. Il y sera également décrit les points qui seront traités par la suite.

Etat des lieux	Objectif	Méthodologie	Résultats et/ou Validation des hypothèses	Validation des hypothèses
Etape effectuée	<u>Étape 1 :</u> Vérification de l'indépendance des données analysées	- Comparer les distributions empiriques selon le nombre de sinistres déclarés par contrat ; - Comparer les valeurs des coûts moyens observés selon le nombre de sinistres déclarés par contrat ; - Analyser l'indépendance des variables tarifaire.	<u>Résultat :</u> Si aucun biais n'est observé, alors l'indépendance des données est établie. <u>Hypothèse émise :</u> Les échelles sont stables d'un exercice à l'autre.	
Prochaine étape	<u>Étape 2 :</u> Détermination des échelles de risque de fréquences selon les critères tarifaires	- Vérifier la significativité des facteurs de l'actuel algorithme tarifaire ; - Choisir le modèle le plus adapté à la modélisation de la fréquence.	<u>Résultat :</u> Échelles de fréquences déterminées <u>Hypothèse du modèle à valider :</u> Étude sur les résidus	- Vérifier l'indépendance des résidus ; - Vérifier l'homoscédasticité des résidus ; - Vérifier la normalité des résidus.

Tableau 2. 10. Point d'étape dans l'analyse et la détermination de la prime pure

A présent que l'étude d'indépendance des données est validée et que l'hypothèse des échelles de risques est définie comme stable d'un exercice à l'autre, le point suivant s'attache à déterminer l'échelle de risque liée à la fréquence de déclaration des sinistres.

II.1.2. Détermination des échelles de risque de fréquences selon les critères tarifaires

L'objectif de ce chapitre est de modéliser le nombre de sinistres à partir d'un modèle de régression défini tel que :

y_{ij} = le nombre de sinistres déclarés par un assuré i durant l'exercice j .

x_{ij} = une matrice de régresseurs (facteurs explicatifs).

Cette matrice de variables explicatives est décrite dans le tableau ci-dessous :

Facteurs	Modalités		
Nombre de pièces	12 modalités : de 1 à 12 pièces		
Type Bâtiment	2 modalités : Collectif (C) ou Individuel (P)		
Limite Capital Mobilier	5 modalités : sans capital, 7.7 fois l'indice FFB, 11 FFB, illimité, autres (clos en souscription)		
Dépendance	5 modalités : [0 - 50 m²],]50 - 100 m²],]100 - 150 m²],]150 - 300 m²], au-delà		
Franchises	12 modalités : 12 valeurs indexées sur l'indice FFB		
Présence de mobilier	(ou mobilier assuré par ailleurs) 2 modalités : Oui, Non		
Bâtiment est assuré par ailleurs	2 modalités : Oui, Non		
Domotique	2 modalités : Oui, Non		
Rééquipement à Neuf	2 modalités : Sans RAN, Avec RAN		
Présence de piscine	5 modalités :		
	I Piscine intérieure		
	A Piscine extérieure non-couverte		
	N Pas de piscine		
	O Piscine extérieure avec couverture automatisée		
	U Piscine extérieure avec couverture manuelle		
Formule	6 modalités : Economique, Confort, Sécurité, Tranquillité, Libre, PNO		
Qualité de l'assuré	15 modalités agrégées en 3 :		
	AL	LOCATAIRE D ALSACE - MOSELLE	Locataire
	AU	AUTRES	Propriétaire
	BA	BAILLEUR	Propriétaire
	CL	CO-LOCATAIRE	Locataire
	CN	COPROPRIETAIRE NON OCCUPANT	Copropriétaire
	CP	COPROPRIETAIRE	Copropriétaire
	FE	FERMIER	Propriétaire
	LF	LOGEMENT DE FONCTION	Locataire
	LO	LOCATAIRE	Locataire
	NU	NU-PROPRIETAIRE	Propriétaire
	OG	OCCUPANT A TITRE GRATUIT	Locataire
	PN	PROPRIETAIRE NON OCCUPANT	Propriétaire
	PR	PROPRIETAIRE	Propriétaire
	SO	SOUS LOCATAIRE	Locataire
US	USUFRUITIER	Propriétaire	

Tableau 2. 11. Matrice des critères tarifaires du modèle actuel

Dans le cadre d'une tarification, la distribution du nombre de sinistres est communément ajustée à une loi discrète. En l'occurrence, dans ce rapport, ce sont essentiellement deux lois qui ont été étudiées de par leur appartenance aux familles des lois exponentielles. Elles sont décrites ci-après.

(a) Les lois possibles pour la modélisation de la fréquence des sinistres

Trois lois discrètes sont souvent citées pour modéliser le nombre de sinistres :

- la loi de Poisson : $P(X = n) = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$, pour $\lambda > 0$ et $k = 0, 1, \dots$

Cette distribution permet de décrire parfaitement un nombre d'événements k (par exemple des sinistres tempêtes) avec une probabilité associée, indépendante du temps écoulé depuis l'événement précédent.

Le paramètre λ représente à la fois la moyenne d'apparition des événements et la variance associée ($Espérance[X] = Variance(X) = \lambda$).

- la loi Binomiale Négative : $P(X = k) = C_{k-1}^{n-1} \cdot p^n \cdot (1-p)^{k-n}$, pour $k \geq n \geq 1$ et $0 < p < 1$

Plusieurs formes d'équation existent pour cette distribution. Pour cette expression, la loi décrit la probabilité d'obtenir n succès parmi k réalisations indépendantes (p désigne la probabilité d'un succès).

Cette distribution est fréquemment utilisée en assurance pour décrire un nombre de sinistres par police, sur un temps fixé et des portefeuilles à risques hétérogènes. Le côté « déclaratif » des assurés est ainsi davantage pris en compte.

Le nombre de réalisation moyen vaut n/p . La variance est égale à $n(1-p)/p^2$.

- et la loi Binomiale : $P(X = k) = C_n^k \cdot p^k \cdot (1-p)^{n-k}$, pour $k = 0, \dots, n$ et $0 < p < 1$

Cette loi est moins utilisée que les deux précédemment citées, et ce pour deux raisons. Tout d'abord, le nombre de sinistres est limité (réduisant la flexibilité de la distribution) et sa variance est faible (inférieure à son espérance).

Pour ces raisons, la loi Binomiale sera exclue de l'étude.

D'autres types de lois peuvent également leur être préférés, telles que les familles de distributions mélanges. Ces lois ont été définies afin d'accroître la surdispersion due à l'hétérogénéité d'un portefeuille. Les distributions Poisson-mélange sont notamment une de ces familles de loi¹².

Malgré l'attrait que pourrait présenter l'étude de ce type de lois, il ne faut pas omettre le côté interprétation d'un modèle. En effet, la compréhension des méthodes utilisées doit être assimilée par tous les acteurs de l'entreprise (notamment les instances dirigeantes).

Pour conclure ce paragraphe, il est présenté ci-dessous une comparaison des lois de Poisson et Binomiale Négative, aux données empiriques.

Afin de comparer ces distributions, les propriétés ou les estimateurs du maximum de vraisemblance¹³ de ces lois ont été utilisées.

Sur les données agrégées, la fréquence est de 1,668% et la variance vaut 0,01762.

Estimation du paramètre de la loi Poisson.

La loi de Poisson de paramètre $\lambda > 0$ se définit telle que $P(Y = n) = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$. Pour un échantillon de

n observations, l'estimateur du maximum de vraisemblance $\hat{\lambda}$ est défini tel que $\hat{\lambda} = \frac{\sum_{i=1}^n k_i}{n}$ où k_i désigne la valeur de la $i^{\text{ème}}$ observation.

De par les données observées sur la base, le paramètre λ de cette loi doit être égal à 1,668%.

Estimation du paramètre de la loi Binomiale Négative.

Étant donné qu'il n'existe pas d'expression analytique simple pour les paramètres de la loi Binomiale Négative (cf. *Annexe [4.2.]*), il a été décidé de les choisir de manière empirique.

Ainsi, pour cette loi, si p désigne la probabilité de non déclaration de sinistres, de par ses propriétés, l'espérance est égale à $\frac{nq}{p}$ et la variance est égale à $\frac{nq}{p^2}$ avec $q=1-p$.

Algébriquement, il vient $p = \frac{\text{Espérance}}{\text{Variance}} \cong 94,63\%$ et $n = \frac{p \times \text{Espérance}}{q} \cong 0,294$.

Or, n doit être un entier naturel non nul. Par conséquent, n sera « forcé » à 1. La probabilité p est alors être déterminée à l'aide de l'estimateur du maximum de vraisemblance $\hat{p} = \frac{m}{m + \sum_{i=1}^m \left(\frac{k_i}{n} - 1 \right)}$ (m

désigne le nombre d'observations et k_i désigne la valeur de la $i^{\text{ème}}$ observation).

Numériquement, $\hat{p} = 98,360\%$.

¹² cf. Polycopier de cours ISFA 2009, « *Gestion des risques d'assurances et de théorie de la ruine* », Stéphane Loisel.

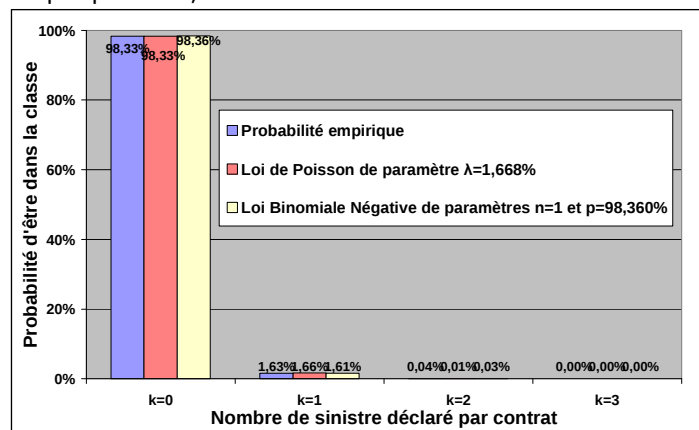
¹³ La définition ainsi que les détails des calculs sont présentés en *Annexe [4.]* et *Annexe [4.1.]*

Soit le tableau de comparaison des distributions de fréquence des sinistres :

$P(X=k)$	k=0	k=1	k=2	k=3
Probabilité empirique	98,332%	1,629%	0,038%	0,001%
Loi de Poisson de paramètre $\lambda=1,668\%$	98,326%	1,660%	0,014%	0,000%
Loi Binomiale Négative de paramètres $n=1$ et $p=98,360\%$	98,360%	1,613%	0,026%	0,000%

Tableau 2. 12. Répartition des sinistres selon les probabilités empiriques, de Poisson et Binomiale Négative

Graphiquement, il vient :



Graphe 2. 1. Représentation graphique des sinistres selon les probabilités empiriques, de Poisson et Binomiale Négative

Constats :

A la lecture du *Tableau 2.12.*, du *Graphe 2.1.* et du coefficient de corrélation entre les données empiriques et les valeurs des lois théoriques, il semblerait que la loi la plus appropriée à la fréquence de déclaration des sinistres, soit la loi Binomiale Négative.

Ces premiers résultats vont à présent être confrontés à une démarche plus analytique. En effet, comme il l'a été noté dans le chapitre I.3., la fréquence correspond à l'ensemble des déclarations qu'effectuent les sociétaires au cours d'un exercice. Il est important d'insister sur le mot « déclaration », qui reflète un aspect comportemental des assurés. Cet aspect va être capté par les facteurs explicatifs de l'algorithme tarifaire, via les techniques de régression.

Avant de poursuivre, un complément a été réalisé en *Annexe [4.3.]*, afin de montrer l'appartenance de la loi de Poisson et de la loi Binomiale Négative à la famille exponentielle. Il conviendra dès lors de choisir le modèle *GLM* le plus approprié à la déclaration des sinistres. Les travaux qui seront présentés ci-dessous ont été réalisés à l'aide du logiciel SAS.

(b) Définition du modèle le plus approprié à la modélisation de la fréquence des sinistres

Deux notions sont à comprendre lorsqu'il s'agit de définir un modèle adéquat (*GLM* approprié) :

- établir une équation avec des facteurs significatifs,
- choisir le « meilleur » modèle explicatif aux données observées.

Tests de significativité des variables

Les tests du Khi-Deux¹⁴ ou de Fisher¹⁵ sont souvent cités dans ce type d'étude. Ces statistiques permettent d'exclure, ou non, des variables qui ne seraient pas pertinentes dans la description des risques. Ils sont communément appelés tests de significativité. Dans ce chapitre, la significativité des facteurs est regardée par rapport à la déclaration des sinistres recensés.

Seuls le modèle poissonien et le modèle Binomiale Négatif sont développés ici. La fonction de lien présentée pour ces deux modèles est la fonction logarithmique (plusieurs fonctions de lien ont été testées, c'est la seule où les résultats ont été les plus probants pour ces deux régressions).

Il est rappelé qu'aux 10 variables précédemment citées, 2 autres critères de modélisation que sont l'« exercice » de survenance et les « Caisses Régionales » ont été ajoutés. Leurs influences sont significatives, mais ne changent en rien l'échelle tarifaire des autres variables (complète indépendance).

Seul le test de Type III est étudié de par sa sévérité par rapport au test de Type I¹⁶.

¹⁴ cf. Annexe [5.1.] : Test du Khi-Deux

¹⁵ cf. Annexe [5.3.] : Test de Fisher-Snedecor et Test de Student

¹⁶ cf. Annexe [5.2.] : Test de Type I contre test de Type III sous SAS

Pour les 2 modèles étudiés (Poisson et Binomiale Négative), les tableaux suivants sont ressortis :

Facteur testé	Selon la loi de Poisson					
	DDL Num.	DDL Res.	Valeur F	Pr > F	Khi-2	Pr > Khi-2
Caisse	9	11696	244,35	<,0001	2199,12	<,0001
Exercice	1	11696	24,18	<,0001	24,18	<,0001
Nombre de pièces	11	11696	322,55	<,0001	3548	<,0001
Type de bâtiment	1	11696	447,37	<,0001	447,37	<,0001
Limite capital mobilier	4	11696	35,54	<,0001	142,14	<,0001
Superficie de dépendance	4	11696	464,93	<,0001	1859,71	<,0001
Franchise	9	11696	75,27	<,0001	677,43	<,0001
Formule	10	11696	55,17	<,0001	551,72	<,0001
Absence de mobilier	1	11696	87,69	<,0001	87,69	<,0001
Bâtiment assuré par ailleurs	1	11696	78,45	<,0001	78,45	<,0001
Présence de domotique	1	11696	283,52	<,0001	283,52	<,0001
Présence de piscine	2	11696	323,82	<,0001	647,64	<,0001
Qualité de l'assuré	14	11696	122,81	<,0001	1719,31	<,0001
RAN (rééquipement à neuf)	1	11696	12,85	0,0003	12,85	0,0003

Tableau 2. 13. Test de significativité des variables pour le modèle poissonien

N.B. : Si le modèle naturel est poissonien, alors y suit une loi de Poisson de paramètre $x\beta$, ce qui peut également s'écrire en espérance $E[y] = x\beta$.

La fonction de lien canonique étant la fonction de lien logarithmique, le modèle est donc de la forme $\ln(y_i) = \sum_i x_i \beta_i$. C'est le modèle présenté dans ce Tableau 2.13.

Facteur testé	Selon la loi Binomiale Négatif					
	DDL Num.	DDL Res.	Valeur F	Pr > F	Khi-2	Pr > Khi-2
Caisse	9	11696	102,81	<,0001	925,29	<,0001
Exercice	1	11696	36,92	<,0001	36,92	<,0001
Nombre de pièces	11	11696	166,2	<,0001	1828,15	<,0001
Type de bâtiment	1	11696	90,05	<,0001	90,05	<,0001
Limite capital mobilier	4	11696	29,88	<,0001	119,51	<,0001
Superficie de dépendance	4	11696	419,16	<,0001	1676,65	<,0001
Franchise	9	11696	67,25	<,0001	605,21	<,0001
Formule	10	11696	7,9	<,0001	79,02	<,0001
Absence de mobilier	1	11696	111,87	<,0001	111,87	<,0001
Bâtiment assuré par ailleurs	1	11696	146,93	<,0001	146,93	<,0001
Présence de domotique	1	11696	463,79	<,0001	463,79	<,0001
Présence de piscine	2	11696	407,83	<,0001	815,67	<,0001
Qualité de l'assuré	14	11696	131,16	<,0001	1836,18	<,0001
RAN (rééquipement à neuf)	1	11696	11,53	0,0007	11,53	0,0007

Tableau 2. 14. Test de significativité des variables pour le modèle Binomiale Négatif

Constats :

- Quel que soit le modèle étudié, tous les critères tarifaires sont significatifs.
- Le critère RAN (rééquipement à neuf) est la variable qui a la plus faible significativité (que ce soit pour le test de Fisher ou le test du Khi-Deux, la probabilité avoisine les 0,07%).

Au regard de la garantie « événements climatiques », même si ce critère est le moins significatif, il dénote tout de même un comportement. Les sociétaires souscrivant ce type d'option sont plus enclins à déclarer leur sinistre (le critère RAN a une influence très forte pour ce qui est de la garantie « Dommages à l'appareillage électrique » ce qui dénote un réel comportement).

N.B.1 : A noter de par les propriétés de la loi de Fisher, que la valeur de Fisher (« Valeur F ») est égale à la valeur du test du Khi-Deux divisée par le degré de liberté de la variable étudiée ($Valeur F = Valeur Khi-Deux / Degré de Liberté$).

N.B.2 : Pour ces modèles, le nombre de sinistres a été pondéré par l'exposition. En effet, un contrat présent toute l'année n'a pas la même probabilité de sinistres qu'un contrat présent 8 mois. Cette pondération a été possible grâce à l'option « offset » de SAS.

Rappel objectif	Méthodologie employée	Points retenus	Étape suivante
Tests de significativité des variables	Utilisation des tests de Type III sur les modèles GLM poissonien et Binomiale Négatif	- Chaque variable est considérée comme significative - Ajout des variables « Exercice » et « Caisse Régionale »	Sélection du modèle le plus approprié à la déclaration des sinistres

Sélection du modèle le plus approprié à la déclaration des sinistres

Afin de déterminer le modèle le plus ajusté à la distribution des déclarations des sinistres, les critères d'information *AIC* et *BIC* ont été utilisés. Un descriptif de ces critères est détaillé en Annexe [6.].

N.B. : Le coefficient de détermination R^2 aurait également pu être utilisé puisqu'il mesure l'adéquation géométrique du modèle aux données. Cependant, ce critère n'est pas à l'abri de critiques. En effet, lorsqu'un prédicteur peu corrélé avec la réponse y est ajouté au modèle, le coefficient R^2 s'approche de 1 (R^2 croît de façon monotone avec le nombre de variables).

Les tableaux des tests de sélection pour les deux modèles étudiés sont les suivants :

Criterion	DF**	Value	Value/DF
Deviance	12000	25273,0976	2.1608
Scaled Deviance	12000	11696	1
Pearson Chi-Square	12000	49582,7723	4.2393
Scaled Pearson X2	12000	22946,1427	1.9619
Full Log Likelihood	—	-12695,8421	—
AIC (smaller is better)	—	25531,6842	—
AICC (smaller is better)	—	25532,5341	—
BIC (smaller is better)	—	26047,7921	—

Tableau 2. 15. Test d'ajustement pour le modèle de régression Poisson

Criterion	DF**	Value	Value/DF
Deviance	12000	12214,8427	1.0444
Scaled Deviance	12000	11696	1
Pearson Chi-Square	12000	25526,5422	2.1825
Scaled Pearson X2	12000	24442,2662	2.0898
Full Log Likelihood	—	-23638,1629	—
AIC (smaller is better)	—	47418,3258	—
AICC (smaller is better)	—	47419,2001	—
BIC (smaller is better)	—	47941,8067	—

Tableau 2. 16. Test d'ajustement pour le modèle de régression Binomiale Négatif

** DF : Degré de liberté

Constats :

- Le modèle le plus approprié est le modèle dont les critères *AIC* et *BIC* ont les valeurs les plus faibles. Des deux tableaux ci-dessus, le modèle poissonien ressort avec des statistiques *AIC* et *BIC* inférieures au modèle Binomiale Négatif.

- Etant donné que les deux modèles ont un nombre identique de variables, le critère R^2 peut toutefois être regardé. Le modèle poissonien apparait avec la valeur la plus proche de 1.

De ces remarques, le modèle poissonien semble donc être le modèle le plus adéquat à la distribution des déclarations des sinistres. Ce dernier sera donc choisi pour modéliser la fréquence.

Pour rappel, ces résultats ont été obtenus sur la base agrégée résultant de la « *proc means* » de SAS. A titre informel, et uniquement pour ces deux modèles, une comparaison avec la table d'origine a été réalisée. Les degrés de liberté sont bien évidemment supérieurs (3 675 266), la log-vraisemblance avoisine les - 219 000 et les critères *AIC* et *BIC* ressortent à des valeurs supérieures aux 438 000. Néanmoins, les résultats restent similaires : le modèle poissonien avec une fonction de lien logarithmique semble être le modèle le plus probant.

Pour la suite de l'étude et pour la lisibilité des résultats, le logiciel Pretium a été utilisé.

La table de données initiale (non agrégée) a été reprise comme base de travail pour ce logiciel.

Rappel objectif	Méthodologies employées	Point retenu	Étape suivante
Sélection du modèle le plus approprié à la déclaration des sinistres	- Utilisation des critères <i>AIC</i> et <i>BIC</i> - Validation confirmée en utilisant la base initiale	Le modèle poissonien est retenu	Présentation des échelles de risques de la fréquence

Le modèle poissonien étant à présent choisi comme le modèle le plus approprié à la déclaration des sinistres, il convient de présenter les résultats des échelles de risques selon les critères tarifaires.

N.B. : Il est précisé que les deux modèles précédemment décrits, ici, ont été testés avec la variable déduite du croisement « nombre de pièces » et « Type de bâtiment ». L'intervalle de confiance était moins stable, mais les mesures de risques semblaient converger. C'est pourquoi, ces 2 critères tarifaires sont restés dissociés et ont simplifié l'algorithme du modèle et sa lecture.

(c) Présentation des échelles de risques de la fréquence

Dès lors que le modèle poissonien a été défini comme le modèle le plus approprié, il est décrit ci-après les échelles de fréquences selon les critères tarifaires actuels. Seule une partie des résultats est présentée ci-dessous :

	Description de l'axe des abscisses	Exposition au risque	Estimateur non lissé loi de Poisson	P-value du critère étudié	Rang (pertinence) selon la loi de Poisson
Nombre de pièces	1	192 592	-0,99958	0,0%	12/14
	2	321 330	-0,49846		
	3	489 055	-0,18079		
	4	560 964	0,00000		
	5	475 223	0,16264		
	6	308 203	0,25270		
	7	167 189	0,36158		
	8	89 100	0,39625		
	9	42 558	0,50895		
	10	22 356	0,58699		
	11	9 248	0,68414		
	12	7 917	0,75944		
Plafond de mobilier garanti par pièce	Sans capital mobilier	302 422	-0,07964	0,0%	7/14
	7,7 FFB	557 561	0,00157		
	11 FFB	311 430	-0,17085		
	illimité	1 315 069	0,00000		
	Autres	199 252	0,00000		
Type de logement	Collectif (C)	687 405	-0,66775	0,0%	10/14
	Individuel (P)	1 998 329	0,00000		
Dépendance	[0 - 50 m²]	2 332 741	0,00000	0,0%	11/14
	]50 - 100 m²]	193 585	0,29039		
	]100 - 150 m²]	66 470	0,39837		
	]150 - 300 m²]	80 156	0,57752		
	Au-delà	12 783	0,73485		
Présence de domotique	N	2 663 316	0,00000	0,0%	2/14
	O	22 418	0,13991		
...	...	...	...	...	...
Bâtiment assuré par ailleurs	N	2 660 342	0,00000	0,0%	6/14
	O	25 392	-0,03851		

Tableau 2. 17. Échelles de risques de la fréquence

Lecture du Tableau 2.17. :

Les variables étudiées sont présentes dans la première colonne. Dans la seconde, ce sont les différentes modalités composant le critère tarifaire qui sont présentées.

Viennent ensuite l'exposition, les estimateurs non lissés (calculés en fonction de la modalité de référence), la « *p-value* » (test de significativité des variables étudiées) et le classement de la pertinence de chaque variable.

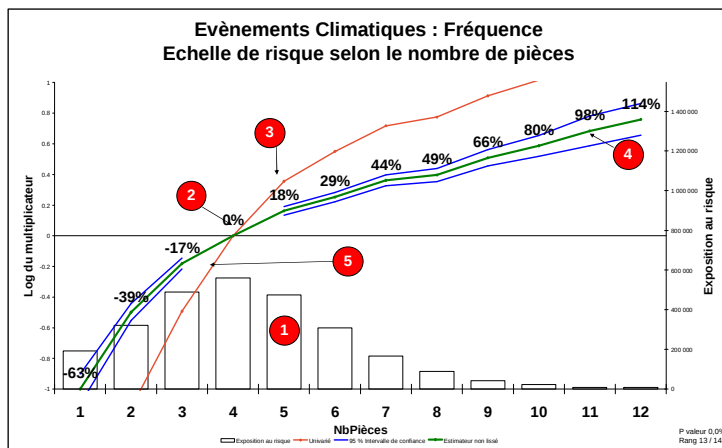
Les variables ayant une « note » de « 12/14 » sont plus pertinentes et donc plus discriminantes que celles qui ont une note de « 2/14 ». Certaines variables ont un pouvoir prédictif identique, leur notation est donc dans ce cas la même.

Il est assez difficile d'interpréter ces coefficients sous cette forme. C'est pourquoi, il est proposé des représentations graphiques de ces paramètres. A partir de ces représentations graphiques, il sera possible d'identifier plus aisément l'impact des différentes modalités des critères étudiés et donc d'en déduire un comportement (dans les déclarations des sinistres).

Dans la suite de ce rapport, les graphiques proposés seront portés essentiellement sur les 3 variables suivantes :

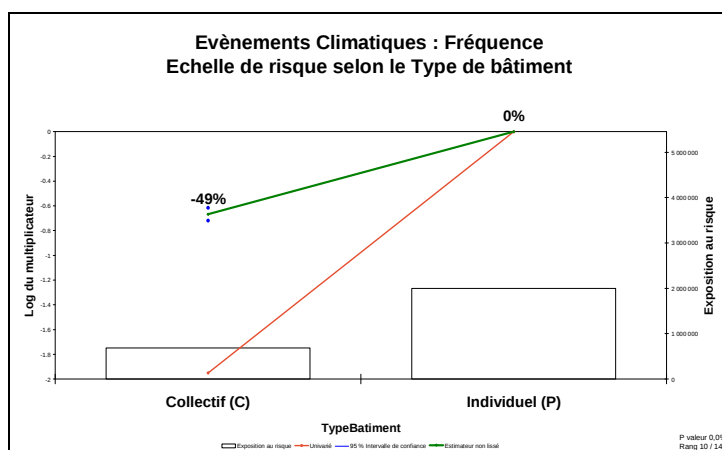
- nombre de pièces dans l'habitation,
- type de bâtiment,
- et surface de la dépendance.

Ces facteurs tarifaires ressortent comme étant les 3 plus pertinents au niveau du modèle.



Graphe 2. 2. Echelle de risque de la fréquence selon le nombre de pièces

3. De ce fait, toutes choses étant égales par ailleurs, les habitations possédant 5 pièces ont 18% de chances supplémentaires de déclarer un sinistre « événements climatiques » par rapport aux habitations possédant 4 pièces.
4. L'intervalle de confiance est représenté par les courbes bleues. Plus les courbes s'éloignent de la courbe de référence en vert, plus l'incertitude est importante sur le pouvoir prédictif de la modalité concernée.
5. La courbe orange représente, quant à elle, la fréquence univariée observée sur le portefeuille.



Graphe 2. 3. Echelle de risque de la fréquence selon le type de bâtiment

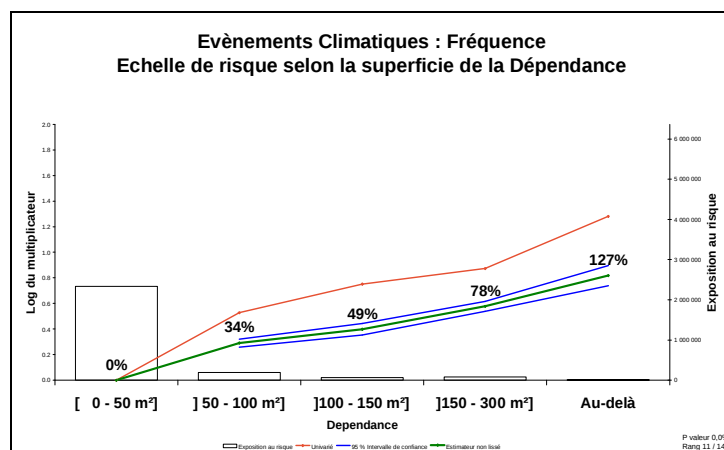
Le graphique ci-contre peut s'interpréter de la façon suivante :

1. Le rectangle blanc représente ici le poids dans le portefeuille des habitations ayant 5 pièces.
2. Le nombre d'habitation possédant 4 pièces est la classe de référence (l'exposition sur cette modalité est la plus importante du portefeuille). Ainsi, l'échelle de risque, plus précisément, les écarts de sinistralité sont mesurés par rapport à cette modalité.

Constats :

Sur ce graphe, il est aisé de voir :

- que les habitations individuelles ont une fréquence de sinistres deux fois plus élevée que les habitations collectives (appartements).
- l'intervalle de confiance est ici représenté par deux points bleus relativement proche de la valeur -49% de la modalité Collective (C).



Graphe 2. 4. Echelle de risque de la fréquence selon la superficie de la Dépendance

Constats :

- Graphiquement, il est possible de conclure que la déclaration de sinistres est croissante par rapport à la surface de dépendance. Ce qui, somme toute, est assez logique : Plus une surface est étendue, plus il y a de probabilité qu'un événement climatique impacte la surface.
- L'échelle est robuste au regard de l'intervalle de confiance.

N.B. : Les autres critères ont fait l'objet d'une analyse similaire.

À présent que les échelles de risques de fréquences sont établies, il convient de regarder si toutes les hypothèses émises dans la modélisation proposée sont validées. Afin de confirmer ces hypothèses, une analyse des résidus doit être réalisée.

(d) Analyse sur les résidus pour validation des hypothèses

L'étude des résidus est bien souvent la seule façon de vérifier empiriquement le bien-fondé des hypothèses posées pour un modèle. Cette analyse vise plusieurs objectifs, à savoir :

- vérifier l'hypothèse de normalité, d'indépendance et d'homoscédasticité des résidus,
- détecter des données aberrantes qui s'écarteraient considérablement du modèle et qui joueraient un rôle important dans la détermination de la régression,
- détecter des tendances particulières et des relations sur les résidus qui pourraient permettre d'affiner le modèle.

Pour ce faire, les méthodes employées sont décrites dans l'*Annexe [7.]* de ce rapport.

En régression, les résidus sont définis comme étant la différence entre la valeur observée y_i et la valeur prédite $\hat{y}_i$ déterminée par le modèle.

Or, les valeurs prédites $\hat{y}_i$ sont des nombres réels positifs, tandis que les nombres de sinistres y_i observés dans la base de données sont des nombres naturels.

De ce fait, selon la valeur de y_i , les résidus bruts $\hat{u}_i = y_i - \hat{y}_i$ auront une distribution asymétrique et n'auront pas la même variance. L'interprétation des résidus de Pearson ou de déviance n'en rend pas une interprétation plus aisée.

Toutefois, les hypothèses d'indépendance et d'homogénéité des variances (ou homoscédasticité selon la littérature) peuvent être vérifiées.

Pour cette vérification, les résidus ont été ordonnés selon la catégorie socio-professionnel (« CSP ») du souscripteur (codé « CDGEDL »). Il ressort les résultats suivants :

	Statistiques sur résidus Bruts	Constats																																												
Hypothèse d'indépendance	Coefficients de corrélation de Pearson Proba > r sous H0: Rho=0 Nombre d'observations <table><tr><td></td><td>ResBrut</td><td>var_new</td></tr><tr><td>ResBrut</td><td>1.00000</td><td>0.03872 <.0001 3685253 3682469</td></tr><tr><td>var_new</td><td>0.03872 <.0001 3682469</td><td>1.00000 3685252</td></tr></table>		ResBrut	var_new	ResBrut	1.00000	0.03872 <.0001 3685253 3682469	var_new	0.03872 <.0001 3682469	1.00000 3685252	Afin de valider l'hypothèse d'indépendance des résidus, la méthodologie employée est référencée en <i>Annexe [7.2.]</i>. <u>Lecture des résultats :</u> Étant donné que le coefficient de corrélation de Pearson¹⁷ est relativement proche de 0, il est possible de conclure que les résidus sont indépendants. Le <i>Programme SAS [7.2.]</i> de cette sortie est également référencé en <i>Annexe</i>.																																			
	ResBrut	var_new																																												
ResBrut	1.00000	0.03872 <.0001 3685253 3682469																																												
var_new	0.03872 <.0001 3682469	1.00000 3685252																																												
Hypothèse d'homoscédasticité	The ANOVA Procedure Bartlett's Test for Homogeneity of ResBrut Variance <table><tr><td>Source</td><td>DDL</td><td>Khi-2</td><td>Pr > Khi-2</td></tr><tr><td>CDGEDL</td><td>182</td><td>408175</td><td><.0001</td></tr></table> The ANOVA Procedure Brown and Forsythe's Test for Homogeneity of ResBrut Variance ANOVA of Absolute Deviations from Group Medians <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>182</td><td>144.5</td><td>0.7939</td><td>71.49</td><td><.0001</td></tr><tr><td>Error</td><td>3.68E6</td><td>40910.0</td><td>0.0111</td><td></td><td></td></tr></table> The ANOVA Procedure Levene's Test for Homogeneity of ResBrut Variance ANOVA of Squared Deviations from Group Means <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>182</td><td>73.9819</td><td>0.4065</td><td>34.25</td><td><.0001</td></tr><tr><td>Error</td><td>3.68E6</td><td>43725.1</td><td>0.0119</td><td></td><td></td></tr></table>	Source	DDL	Khi-2	Pr > Khi-2	CDGEDL	182	408175	<.0001	Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	182	144.5	0.7939	71.49	<.0001	Error	3.68E6	40910.0	0.0111			Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	182	73.9819	0.4065	34.25	<.0001	Error	3.68E6	43725.1	0.0119			Pour vérifier l'hypothèse d'homogénéité des variances, plusieurs tests ont été réalisés et s'appuient sur les méthodologies proposées en <i>Annexe [7.3.]</i>. Ils sont tous basés sur la « <i>proc ANOVA</i> » de SAS. (Voir également le <i>Programme SAS [7.3.]</i> référencé en <i>Annexe</i>). <u>Lecture des résultats :</u> Des sorties SAS, l'hypothèse d'égalité des variances est validée. En effet, les tests de Bartlett, Brown & Foresthe et Levene sont significatifs.
Source	DDL	Khi-2	Pr > Khi-2																																											
CDGEDL	182	408175	<.0001																																											
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																									
CDGEDL	182	144.5	0.7939	71.49	<.0001																																									
Error	3.68E6	40910.0	0.0111																																											
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																									
CDGEDL	182	73.9819	0.4065	34.25	<.0001																																									
Error	3.68E6	43725.1	0.0119																																											

Tableau 2. 18. Analyse des résidus pour le modèle de fréquence

¹⁷ Ce coefficient correspond au coefficient de corrélation multiple R , spécifié en *Annexe [6.]*

Pour ce qui est de la vérification du test de normalité, comme il l'a été noté, l'interprétation est plus délicate. Toutefois, il va être tenté d'en vérifier l'hypothèse.

Plusieurs méthodes existent. Dans ce rapport, ce sont essentiellement 3 tests qui seront exposés. La méthodologie et l'interprétation de ces tests sont référencées en *Annexe [7.1.]* et *[7.1.bis]*.

Afin d'obtenir des graphiques lisibles, l'histogramme des résidus et la Droite de Henry ont été produits sur le logiciel Excel. Sachant le nombre de lignes conséquent, il a été décidé de tirer aléatoirement¹⁸ un nombre de 30 000 lignes (la sortie graphique de SAS, peu compréhensible, est référencée en *Annexe sous Programme SAS [7.1.]*).

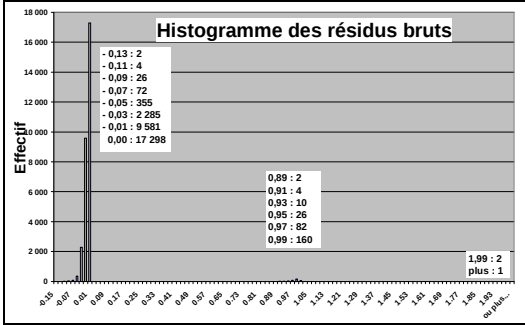
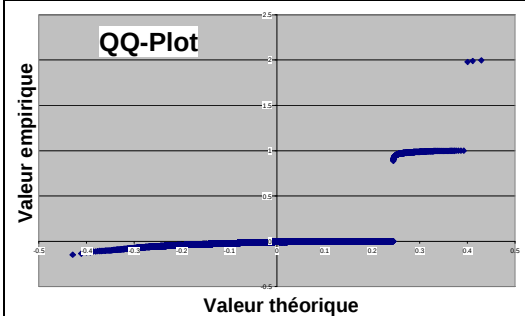
		Statistiques sur résidus bruts	Constats																								
Hypothèse de normalité	Test de normalité	Tests de tendance centrale : $\mu_0=0$ <table><thead><tr><th>Test</th><th>-Statistique-</th><th>-----p-Value-----</th></tr></thead><tbody><tr><td>t de Student</td><td>t 0</td><td>Pr > t 1.0000</td></tr><tr><td>Signe</td><td>M -1800614</td><td>Pr >= M <.0001</td></tr><tr><td>Rang signé</td><td>S -3.24E12</td><td>Pr >= S <.0001</td></tr></tbody></table> Tests de normalité <table><thead><tr><th>Test</th><th>-Statistique--</th><th>-----p-Value-----</th></tr></thead><tbody><tr><td>Kolmogorov-Smirnov</td><td>D 0.4886</td><td>Pr > D <0.0100</td></tr><tr><td>Cramer-von Mises</td><td>W-Sq 223178.2</td><td>Pr > W-Sq <0.0050</td></tr><tr><td>Anderson-Darling</td><td>A-Sq 1084797</td><td>Pr > A-Sq <0.0050</td></tr></tbody></table>	Test	-Statistique-	-----p-Value-----	t de Student	t 0	Pr > t 1.0000	Signe	M -1800614	Pr >= M <.0001	Rang signé	S -3.24E12	Pr >= S <.0001	Test	-Statistique--	-----p-Value-----	Kolmogorov-Smirnov	D 0.4886	Pr > D <0.0100	Cramer-von Mises	W-Sq 223178.2	Pr > W-Sq <0.0050	Anderson-Darling	A-Sq 1084797	Pr > A-Sq <0.0050	Même si les tests de Kolmogorov-Smirnov, Cramer-von Mises et Anderson-Darling sont significatifs, il est difficile de conclure à une quelconque normalité. En effet, des sauts sont observables sur les résidus (résultat de la soustraction, nombre de sinistres – valeur prédite) : Voir les deux graphiques suivants.
	Test	-Statistique-	-----p-Value-----																								
t de Student	t 0	Pr > t 1.0000																									
Signe	M -1800614	Pr >= M <.0001																									
Rang signé	S -3.24E12	Pr >= S <.0001																									
Test	-Statistique--	-----p-Value-----																									
Kolmogorov-Smirnov	D 0.4886	Pr > D <0.0100																									
Cramer-von Mises	W-Sq 223178.2	Pr > W-Sq <0.0050																									
Anderson-Darling	A-Sq 1084797	Pr > A-Sq <0.0050																									
	Histogramme des résidus	 Les différents histogrammes de répartition étudiés ont permis de constater que les valeurs des résidus bruts étaient comprises entre :]-0,25 ; 0], [0,85 ; 1], [1,85 ; 2] et [2,90 ; 3].																									
	Droite d'Henry	 - L'introduction du graphique <i>QQ-Plot</i> est intéressante, en ce sens qu'une quasi-linéarité est observable sur chaque « escalier ». - Toutefois, cette linéarité ne semble pas complète sur les parties gauches de ces droites.																									

Tableau 2. 19. Analyse des résidus pour le modèle de fréquence (suite)

Il est relativement difficile d'interpréter une distribution gaussienne avec ce type de représentation (y compris avec les résidus de Pearson, Student ou de déviance).

L'idée d'étudier une loi composée a donc été soulevée. Le principe consiste à modifier la loi de Poisson en zéro (cf. le « *Projet Etienne Marceau, Méthodes statistiques en assurance non vie* » (2007) de Sauser Carine, Groisne Mélanie et Milhaud Xavier).

L'étude de ce type de résidus (pour un modèle de régression basé sur une loi discrète) peut faire l'objet de tout un dossier. En effet, il n'existe pas à ce jour de « vraies méthodes » pour valider une quelconque normalité à ces valeurs.

Une analyse complémentaire sur ces résidus est toutefois proposée en *Annexe [8.2.]*.

¹⁸ Voir le *Programme SAS [7.1.] : Test de normalité des résidus de l'Annexe [7.] : Descriptif méthodologique pour l'analyse des résidus*

Rappel objectif	Méthodologie employée	Points retenus	Étape suivante
Analyse des résidus pour validation du modèle	Vérification des hypothèses : - d'indépendance, - d'homoscédasticité, - de normalité.	- Hypothèses <i>a priori</i> validées sur les résidus. - Pas de données aberrantes. - Le modèle choisi est donc <i>a priori</i> adapté à la modélisation de la fréquence.	Détermination de l'échelle de risques des coûts moyens selon les critères tarifaires

Point de transition :

Pour certains contrats présents en portefeuille, le taux de fréquences sinistres est de 22,52% (soit plus d'un sinistre tous les 5 ans). Ce taux est particulièrement élevé et donc « inquiétant ». En effet, la fréquence globale de la base étudiée n'est que de 1,67%.

Face à une fréquence aussi élevée, une analyse plus poussée sur ces contrats doit être réalisée, car seule la garantie « Evénements Climatiques » a été étudiée. Si, en parallèle, ce type de profil génère des sinistres sur d'autres garanties, l'entreprise doit mettre en place et adapter une certaine politique pour faire face à ce niveau de risque et le réduire.

Le tableau qui suit, présente le point d'avancement dans la modélisation de la prime pure.

Etat des lieux	Objectif	Méthodologie	Résultats et/ou Validation des hypothèses	Validation des hypothèses
Etape effectuée	<u>Étape 2 :</u> Détermination des échelles de risques de fréquences selon les critères tarifaires	- Vérifier la significativité des facteurs de l'actuel algorithme tarifaire - Choisir le modèle le plus adapté à la modélisation de la fréquence	<u>Résultat :</u> Échelles de fréquences déterminées <u>Hypothèse du modèle à valider :</u> Étude sur les résidus	- Vérifier l'indépendance des résidus - Vérifier l'homoscédasticité des résidus - Vérifier la normalité des résidus
Prochaine étape	<u>Étape 3 :</u> Détermination des échelles de risques des coûts moyens selon les critères tarifaires	- Définir un seuil d'exclusion entre les sinistres attritionnels et les sinistres non-récurrents - Vérifier la significativité des facteurs - Choisir le modèle le plus adapté à la modélisation des coûts moyens	<u>Résultat :</u> Échelles des coûts moyens déterminées <u>Hypothèse du modèle à valider :</u> Étude sur les résidus	- Vérifier l'indépendance des résidus - Vérifier l'homoscédasticité des résidus - Vérifier la normalité des résidus

Tableau 2. 20. Point d'étape dans l'analyse et la détermination de la prime pure.

A présent que les échelles de fréquences ont été déterminées, il convient de s'intéresser à la modélisation des échelles de coûts moyens.

II.1.3. Détermination des échelles de risque des coûts moyens selon les critères tarifaires

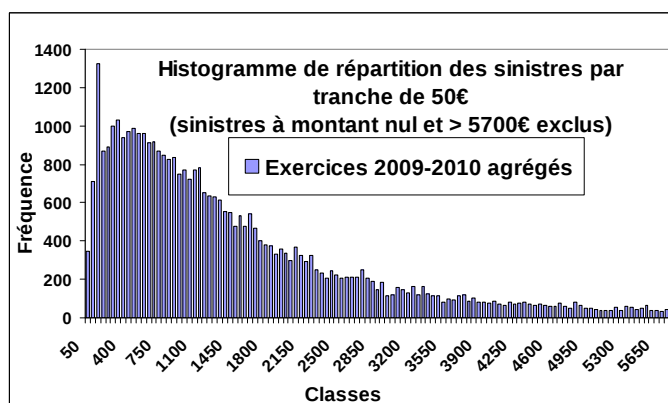
Pour rappel, le coût moyen correspond à la somme des charges sinistres de l'année divisée par le nombre de sinistres total.

Avant de poursuivre et afin d'avoir une homogénéité dans les résultats, il est important de rappeler et préciser que les montant des sinistres des différents exercices ont été actualisées¹⁹.

¹⁹ cf Annexe [2.] : Indice FFB de 1980 à fin 2011

Sur ces 2 exercices, les charges sinistres sont réparties comme suit :

Tranche de coût	% effectif cumulé	% charge sinistre cumulé
[0€]	23,6%	0,0%
]0;100]	29,8%	0,4%
]100;200]	34,4%	1,0%
]200;300]	37,9%	1,7%
]300;400]	41,6%	2,6%
]400;500]	45,2%	3,7%
]500;600]	48,5%	5,0%
]600;700]	51,5%	6,3%
]1700;1800]	75,3%	24,3%
]1800;1900]	76,6%	25,8%
]3400;3500]	89,9%	47,0%
]3500;3600]	90,3%	48,0%
]5600;5700]	94,9%	61,2%
]5700;5800]	95,1%	61,7%
]9900;10000]	97,9%	75,1%
]10000;10100]	98,0%	75,5%
]14400;14500]	98,9%	82,9%
]14500;14600]	99,0%	83,1%
]20800;20900]	99,4%	88,5%
]20900;21000]	99,5%	88,6%



Graphe 2. 5. Histogramme de répartition des sinistres de la garantie « événements climatiques » par tranche de 50 € (sinistres à montant nul et > 5700€ exclus)

Tableau 2. 21. Répartition des sinistres « événements climatiques » sur les exercices 2009 et 2010.

Constats :

Plusieurs remarques sont à formuler suite à ces figures :

- Il est possible d'observer une forte proportion de sinistres dont le montant des règlements est nul. La majeure partie de ces contrats sinistrés possède une franchise. Les franchises ont donc un enjeu non négligeable sur les risques de cumuls.
- Près d'un sinistre sur deux a un coût inférieur à 700 euros. Même si la proportion est moins importante, une partie de ces contrats sinistrés a également une franchise. En revanche, la charge cumulée de l'ensemble de ces sinistres inférieurs à 700€ ne représente que 6,3% de la charge totale.
- Il y a au total 243 sinistres qui ont un coût supérieur à 21 000€, dont 5 qui ont un coût de plus de 100 000€. Le sinistre maximum a un coût actualisé de 139 295€.
- Un pic est observable sur la tranche des coûts de sinistres compris entre 150€ et 200€. Cette tranche de coût correspond à l'envoi d'un expert.

Afin de pouvoir élaborer une échelle de risque des coûts moyens, il est impératif de déterminer un seuil (montant) permettant de distinguer les sinistres récurrents, des sinistres plus exceptionnels. En effet, ces sinistres non-récurrents pourraient impacter de façon disproportionnée une ou des modalités sur l'échelle de risques des coûts moyens, comme par exemple les sinistres dont les montants sont supérieurs aux 100 000€.

(a) Définition d'un seuil d'exclusion entre les sinistres attritionnels et les sinistres non-récurrents dans le cadre d'une modélisation des coûts moyens

À ce jour, il n'existe pas de réelle démarche pour définir un seuil entre les sinistres attritionnels et les sinistres à montant moins courant. Néanmoins, cette analyse est importante, si ce n'est fondamentale. En effet, deux notions sont à comprendre à travers ce seuil :

- Le récurrent (ou selon la littérature l'« attritionnel ») porte le coût du risque,
- L'exceptionnel doit être mutualisé (de par son caractère) sur l'ensemble du portefeuille.

N.B. : Une façon de mutualiser ces sinistres non-récurrents est de répartir la charge en appliquant un taux proportionnel à chaque profil de risque. En effet, des profils ayant un coût du risque attritionnel important ont, proportionnellement, plus de probabilité d'engendrer des sinistres à caractères exceptionnels.

Ainsi, si f désigne la fonction de densité des montants de sinistre $y > 0$, il est possible d'exprimer cette distribution sous une forme de loi mélange telle que :

$f(y) = P(y < \text{seuil})f_{\text{récurrent}}(y) + P(y \geq \text{seuil})f_{\text{non-récurrent}}(y)$ où $f_{\text{récurrent}}$ et $f_{\text{non-récurrent}}$ désignent des fonctions de densité et où P désigne la probabilité que y dépasse ou non un certain seuil.

Pour un ensemble de points décrit ci-après et avec des études réalisées en parallèle, la décision a été prise de fixer ce seuil à hauteur de 5700€.

Cette décision s'appuie sur les points suivants :

- Robustesse statistique dans la description du risque : 95% des sinistres sont pris en compte (voir *Tableau 2.21.*) dont 71,4% des sinistres à montant non nul ;
- Pratique de gestion et du calcul des provisions : Dans le cadre d'une gestion des sinistres aux coûts moyens, un seuil de règlement trop élevé entraînerait une gestion systématique au coût moyen. Or, hormis le fait que les coûts moyens fluctueraient dans des proportions importantes, l'entreprise perdrait la qualité d'appréciation des gestionnaires sinistres pour l'estimation des événements non récurrents.
Une fois les sinistres supérieurs à 5700€ exclus, les coûts moyens 2009 et 2010 sont respectivement de 1475€ et 1451€ (hors sinistre à montant nul), soit une fluctuation de 1,7% uniquement) ;
- Robustesse prédictive : Stabilité des échelles d'un exercice à l'autre.

N.B.1 : Les sinistres ont été actualisés par rapport à l'indice FFB (indice d'indexation des franchises). L'écart de 1,7% observé sur les coûts moyens 2009 et 2010 peut être expliqué par le fait que l'indice FFB soit calculé à partir du prix de revient d'un immeuble de rapport de type courant à Paris. Cet indice n'est peut être donc pas suffisamment représentatif, puisque ce n'est pas un indicateur de niveau national (voir *Annexe [2.]*).

N.B.2 : Il est également important de mettre en avant les notions de « règlement », « provision en règlement », « recours » et « prévision de recours ». Au moment de la déclaration du sinistre, une première évaluation est effectuée. Cette estimation tient essentiellement compte du coût global d'indemnisation (ou selon la littérature, du montant de règlement brut). Ce coût n'est alors constitué que de provisions.

En effet, diverses étapes peuvent avoir lieu avant la clôture du dossier. De ce fait, le coût total se compose de divers règlements successifs (indemnisation de l'entité assurée, de l'expert, d'un possible relogement, ...) qui viennent diminuer, au fur et à mesure, la provision constituée lors de la déclaration. La notion de recours intervient au regard de la responsabilité engagée. En effet, la compagnie peut se retourner contre un tiers (ou objet de ce tiers, tel un arbre) qui serait à l'origine du sinistre. Dans ce cas, une demande est faite à ce tiers (via sa compagnie d'assurance) dont la responsabilité est engagée. Pour la garantie « événements climatiques », il n'y a pas eu de recours sur les exercices étudiés.

N.B.3 : Dans une gestion des sinistres au coût moyen, le nombre de sinistres « sans suite » (i.e. à montant nul) doit toujours être inclus. Hors, un taux de « sans suite » fluctuant d'un exercice à l'autre, engendre de fortes perturbations sur l'estimation des coûts moyens.

Sur l'exercice 2009, le taux de « sans suite » était à près de 17,3%, contre un taux proche des 30% pour l'exercice 2010, soit mécaniquement une baisse du coût moyen de plus de 12%. Comme il l'a été vu, en ayant au préalable exclus ces sinistres à montant nul et supérieurs à un montant de 5700€, il n'est apparu que 1,7% d'écart sur les coûts moyens 2009 et 2010.

N.B.4 : Dans ce genre d'analyse, il est également important d'exclure les sinistres dont le montant de règlement est supérieur à un certain seuil (et non les écrêter). En effet, dans le cadre d'un écrêtement, les coûts moyens calculés seraient biaisés de par le point d'accumulation créé à ce seuil. A titre informel, il est renseigné en *Annexe [8.3.]*, une analyse complète sans seuil d'exclusion, ni d'écrêtement. Dans cette analyse, il est aisé de remarquer que le modèle ne donne pas de résultats satisfaisants.

À noter en plus, que l'ordre de pertinences des critères était relativement discordant (par exemple, le « type de bâtiment » arrive dans les dernières positions pour la description du risque).

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Définition d'un seuil d'exclusion entre les sinistres attritionnels et les sinistres non-récurrents dans le cadre d'une modélisation des coûts moyens	- Robustesse statistique dans la description du risque - Pratique de gestion et du calcul des provisions - Robustesse prédictive	- Les règlements de sinistres dépassant les 5700€ sont exclus afin d'éviter un point d'accumulation - Stabilité des échelles d'un exercice à l'autre	Définition du modèle le plus approprié à la modélisation des coûts moyens

A présent que ce seuil est déterminé, il convient de définir le modèle le plus adapté à la modélisation des échelles de risques des coûts moyens.

(b) Définition du modèle le plus approprié à la modélisation des coûts moyens

Les lois Exponentielle, Gamma, LogNormale, Bêta, Inverse Gaussienne, Weibull, LogGamma ou bien encore Pareto sont les lois les plus utilisées pour modéliser une distribution des montants de sinistres.

Cependant, afin de continuer à utiliser les techniques de régression décrites au début de ce chapitre, il est nécessaire que la loi choisie appartienne à la famille des lois exponentielles. Ainsi, même si les lois LogNormale, Weibull, LogGamma, Pareto modélisent parfaitement une distribution de montant de sinistres, elles doivent être exclues, car elles ne font pas parties des familles de lois exponentielles²⁰ ou parce qu'elles introduisent un biais dans l'estimation du coût (le logarithme de l'espérance des coûts n'est pas égal à l'espérance des logarithmes des coûts²¹).

Dans ce modèle de régression (au même titre que la modélisation de la fréquence de déclarations des sinistres), l'objectif est de modéliser les coûts moyens tel que :

y_{ij} = la somme des charges sinistres déclarés par un assuré i durant l'exercice j divisé par le nombre déclaré.

x_{ij} = [la même matrice qu'au chapitre II.1.1.].

N.B. : Si un sociétaire i a déclaré 2 sinistres durant l'exercice j , alors y_{ij} est le résultat de la somme des règlements fait au titre de ces 2 sinistres. Dans l'analyse, le modèle en tient compte via l'option « offset » de SAS.

Il peut sembler important de préciser que les valeurs des coûts moyens sont à valeurs dans $\mathbb{R}^+$, donc la loi recherchée doit être définie sur ce périmètre (une loi gaussienne doit par exemple être rejetée).

Les lois de la famille exponentielle, évoquées dans le chapitre précédent II.1., sont pratiques de par les facilités de calculs qu'elles peuvent induire, en particulier lors des calculs de loi *a posteriori*. Le second avantage réside dans le fait qu'il existe des lois *a priori* conjuguées²² naturelles pour quelques familles exponentielles usuelles.

Dans le tableau suivant²³, quelques familles exponentielles usuelles sont recensées avec leurs lois *a priori* conjuguées.

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
Loi de Poisson $P(\theta)$	Loi Gamma $G(\alpha, \beta)$	$G(\alpha + x, \beta + 1)$
Loi Binomiale Négative $BN(n, \theta)$	Loi Bêta $Be(\alpha, \beta)$	$Be(\alpha + n, \beta + x)$
Loi Binomiale $B(n, \theta)$	Loi Bêta $Be(\alpha, \beta)$	$Be(\alpha + x, \beta + n - x)$
Loi Normale $N(\theta, \sigma^2)$	Loi Normale $N(\mu, \tau^2)$	$N(\frac{\sigma^2\mu + \tau^2x}{\sigma^2 + \tau^2}, \frac{\sigma^2\tau^2}{\sigma^2 + \tau^2})$
Loi Gamma $G(v, \theta)$	Loi Gamma $G(\alpha, \beta)$	$G(\alpha + v, \beta + x)$

Tableau 2. 22. Table de correspondance de lois *a priori* conjuguées

Etant donné que la loi de Poisson est la loi qui s'apparente le mieux à la modélisation des déclarations de sinistres, il semblerait donc naturel, de par le tableau ci-dessus, que les montants des sinistres soient modélisés par la loi Gamma.

Afin d'étayer ce choix, une approche empirique est proposée. Trois courbes de distribution vont être ajoutées à l'histogramme de répartition des sinistres par tranche de 50 € (Graphe 2.5.) :

- la courbe de distribution de la loi Exponentielle,
- la courbe de distribution de la loi Gamma,
- la courbe de distribution de la loi Inverse Gaussienne.

N.B. : L'appartenance de ces lois aux familles de lois exponentielles est montrée en Annexe [4.3.]

²⁰ Lawrence Brown *Fundamentals of statistical exponential families with applications in statistical decision theory*

²¹ PARTRAT Christian, BESSON Jean-Luc (2005) *Assurance Non-Vie Modélisation, Simulation*. Economica

²² Voir définition dans l'Annexe [4.4.] : Familles de lois conjuguées

²³ Extrait du livre *Le choix bayésien: principes et pratique* de Christian Robert (2006)

Pour comparer ces courbes de distribution théorique à la distribution empirique, les propriétés des lois ci-dessus et/ou l'estimateur du maximum de vraisemblance²⁴ ont été utilisés, afin de générer 26 000 sinistres à chaque fois (nombre de sinistres compris dans l'intervalle] 0 ; 5700€[dans la base de données).

Estimation du paramètre de la loi exponentielle

La loi exponentielle a pour fonction de densité $f(x) = \theta e^{-\theta x}$. Or, pour un échantillon de n observations, l'estimateur du maximum de vraisemblance de la loi exponentielle est $\hat{\theta} = \frac{n}{\sum_{i=1}^n x_i}$ où x_i désigne la valeur de la $i^{\text{ème}}$ observation (voir l'Annexe [4.1.] pour le détail des calculs).

Sur les sinistres compris dans l'intervalle] 0 ; 5700€[, la moyenne empirique est de 1464€. Ainsi, le paramètre de la loi exponentielle est défini et 26 000 sinistres ont pu être générés.

Sous Excel, il est également possible de tracer rapidement des courbes de tendance, notamment une courbe de tendance exponentielle. Les paramètres de cette courbe sont optimisés selon la méthode des moindres carrés²⁵.

Estimation des paramètres de la loi Gamma

La loi Gamma a pour fonction de densité $f_{(\alpha, \beta)}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ ou selon la littérature $f_{(\alpha, \lambda)}(x) = \frac{1}{\lambda^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\lambda}}$ avec $\alpha > 0$ et $\lambda = \frac{1}{\beta} > 0$. En restant sur cette dernière expression,

l'espérance de la loi Gamma vaut $\alpha\lambda$, quant à la variance, elle est égale à $\alpha\lambda^2$.

Ainsi, grâce aux propriétés de la loi Gamma, il est possible d'estimer les paramètres α et λ . En effet, il vient $\lambda = \frac{\text{Variance}}{\text{Espérance}}$ et $\alpha = \frac{\text{Espérance}^2}{\text{Variance}}$.

Numériquement, de par la moyenne et la variance empirique des sinistres observés, α et λ valent respectivement 1,4031 et 1028,59. Sous le logiciel Excel, grâce à ces valeurs, il a été possible de générer, là aussi, 26 000 sinistres.

N.B.1 : Pour la loi Gamma, il n'existe pas de solutions simples quant à la détermination des paramètres selon la méthode du maximum de vraisemblance (voir Annexe [4.1.] pour les calculs).

N.B.2 : La loi Gamma est sensible aux valeurs extrêmes. Mais, de par le paragraphe (a), les sinistres considérés ici appartiennent uniquement à l'intervalle] 0 ; 5700€[.

Estimation des paramètres de la loi Inverse Gaussienne

La loi Inverse Gaussienne a pour fonction de densité $f_{(a,b)}(x) = \frac{\sqrt{a}}{\sqrt{2\pi} x^3} e^{-\frac{a(x-b)^2}{2b^2x}}$, son espérance vaut b , quant à la variance, elle est égale à $\frac{b^3}{a}$. De par les données empiriques et grâce aux

propriétés de la loi Inverse Gaussienne, il est possible d'estimer les paramètres a et b . En effet, il vient $a = \frac{\text{Espérance}^3}{\text{Variance}}$ et $b = \text{Espérance}$.

Sur les outils en possession, il n'existe pas de générateur pour cette loi. Par conséquent, il convient de le créer. Pour ce faire, le calibrage de ce générateur s'appuie sur les cours de *Simulation* dispensés par Bernard Delyon, à l'Université de Rennes 1. La méthodologie est la suivante :

1 - Soit Y une variable aléatoire de loi normale centrée réduite.

2 - Il est nécessaire de calculer les deux solutions de l'équation suivante : $y^2 = \frac{a(x-b)^2}{b^2x}$. Si Δ

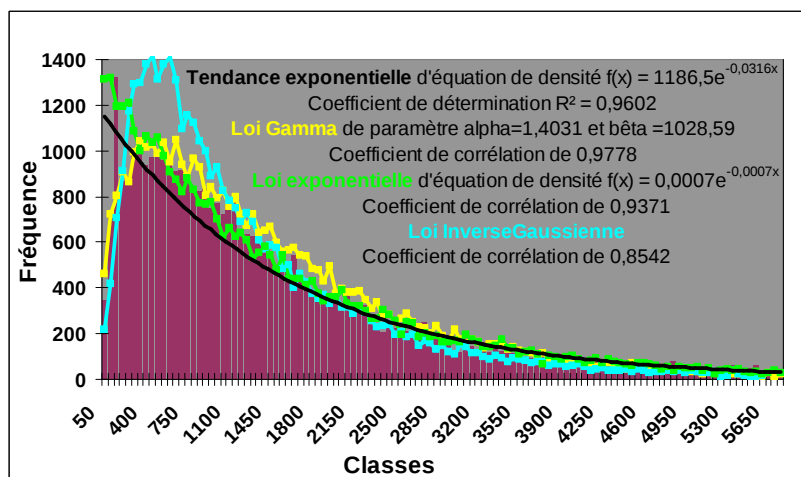
est le discriminant, alors $\Delta = 4ab^3y^2 + b^4y^4$ et $X_{\pm} = \frac{-2(b^2y^2 + 2ab) \pm \sqrt{\Delta}}{-2a}$.

3 - Soit U une variable aléatoire uniforme sur (0,1). La solution X_- est prise lorsque la probabilité $U < p$ avec $p = (1 + X_-/b)^{-1}$. Le cas contraire, il faut prendre X_+ .

²⁴ Voir Annexe [4.2.] : Maximum de Vraisemblance

²⁵ Voir Annexe [4.1.] : Moindres Carrés Ordinaires

A présent, il est représenté sur le graphique ci-dessous les résultats des différentes simulations.



Graph 2. 6. Comparaison des distributions théoriques et de la distribution empirique sur les sinistres appartenant à l'intervalle]0 ; 5700€[

Ainsi, de par le *Graph 2.6.*, la répartition des coûts moyens empiriques et la distribution théorique Gamma semblent avoir la meilleure compatibilité (le coefficient de corrélation de la loi Gamma avec les données empiriques atteint la valeur de 0,9778). Même s'il semble qu'il y ait une surreprésentation des montants compris entre 1800 et 2500€, opter pour une modélisation des coûts moyens par le biais d'un modèle Gamma reste cohérent et prudent.

Ces premiers résultats ont également été confrontés à la démarche analytique proposée dans le chapitre II.1.1. :

- Dans un premier temps, des tests de significativité des variables ont donc été réalisés. Tous les critères étaient significatifs, mais les rangs étaient différents (assez paradoxalement, ici aussi, la variable RAN avait la significativité la plus faible. Alors qu'il n'est pas appliqué de vétusté sur les biens assurés, il aurait pu être logique d'observer un rang plus important).
- Puis, les critères AIC et BIC ont été utilisés pour déterminer le modèle le plus adapté à la modélisation des coûts moyens.

Suite à cette démarche, le modèle Gamma est ressorti comme étant le modèle le plus adapté à la modélisation des échelles des coûts moyens des sinistres.

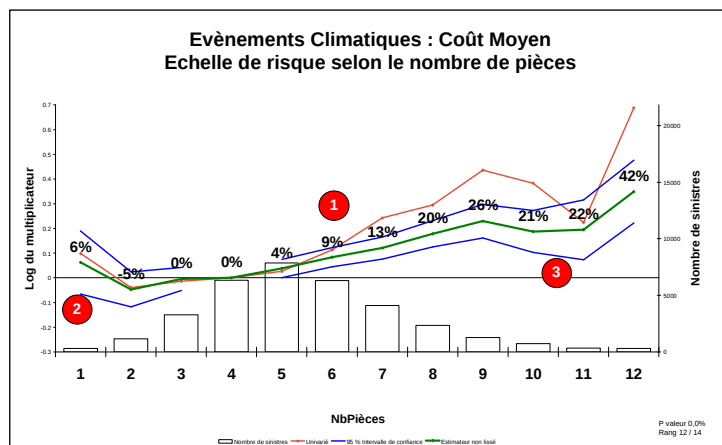
N.B. : Pour information, la loi Bêta qui appartient à la famille de lois exponentielles et dont la fonction de densité s'écrit $f_{(\alpha,\beta)}(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1}$ permet de modéliser un taux de dommages. Si la loi Binomiale Négative avait été choisie pour la modélisation des fréquences de sinistres, il aurait donc été nécessaire de regarder le rapport montant de sinistres sur capital assuré. Dans la base d'analyse, cette information n'était pas présente. C'est pourquoi, cette loi n'a pas été sélectionnée.

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Définition du modèle le plus approprié à la modélisation des coûts moyens	- Lien avec la loi conjuguée - Appartenance à la famille des lois exponentielles - Comparer la distribution empirique et les distributions théoriques - Vérifier de la significativité des facteurs (Utilisation des tests de Type III sur les modèles GLM) - Choisir le modèle le plus adapté à la modélisation des coûts moyens (Utilisation des critères AIC et BIC)	- Le modèle Gamma est ressorti comme le modèle le plus adapté - Chaque variable est considérée comme significative	Présentation des échelles de risques des coûts moyens

Le modèle Gamma étant à présent choisi comme le modèle le plus approprié à la modélisation des montants des sinistres, il convient de présenter les résultats des échelles de risques selon les critères tarifaires. Pour une question de lisibilité, il ne sera proposé que les graphiques.

Comme pour la fréquence de déclaration des sinistres, seuls les facteurs « nombre de pièces », « type de bâtiment » et « dépendance » seront présentés.

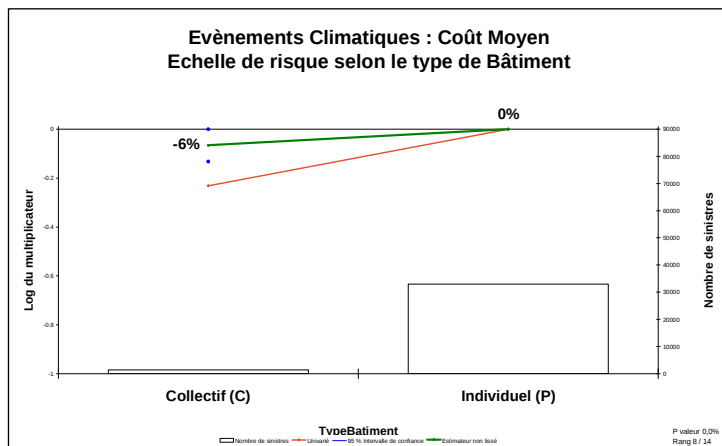
(c) Présentation des échelles de risques des coûts moyens



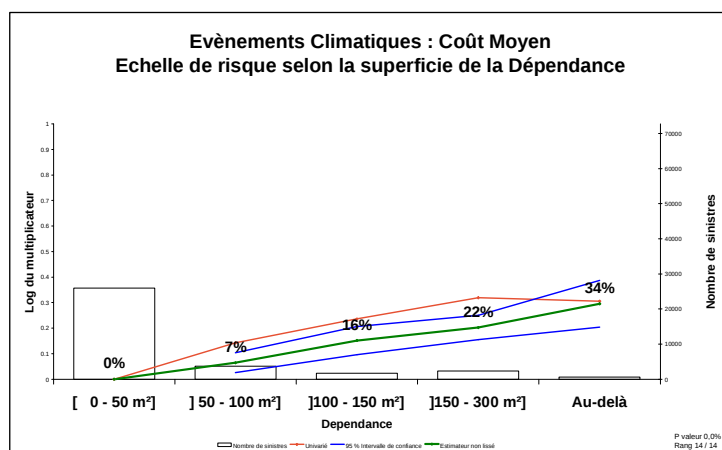
Graphe 2. 7. Echelle de risque du coût moyen selon le nombre de pièces

Sur ce Graphe 2.7., l'analyse du modèle met de plus en évidence :

- 1.** Un coût moyen croissant entre les habitations ayant un nombre de pièces compris entre 2 et 12 pièces (hors habitations à 10 et 11 pièces).
- 2.** Les habitations ayant une pièce ont un coût moyen plus élevé que les habitations possédant 2 pièces. L'interprétation est délicate, si ce n'est que l'intervalle de confiance est assez éloigné de la courbe verte (GLM). Pour les 2 exercices, il n'y a eu que 305 montants de sinistres non nuls recensés sur les habitations d'une pièce.
- 3.** Le fait que les habitations ayant 10 et 11 pièces aient un coût moyen en dehors de l'allure croissante est dû à l'exercice 2010 (Voir Annexe [8.1.]). En effet, sur cet exercice, les montants sur ces habitations étaient relativement faibles. De plus, tout comme le constat précédent **2.**, l'intervalle de confiance est relativement éloigné de la courbe du GLM. Il y a eu respectivement 739 et 338 sinistres recensés sur les habitations de 10 et 11 pièces.



Graphe 2. 8. Echelle de risque du coût moyen selon le type de bâtiment



Graphe 2. 9. Echelle de risque du coût moyen selon la superficie de la Dépendance

Constats :

Sur ce graphique, il est possible de constater que la courbe du GLM reste relativement « écartée » de celle de l'univarié. Sans l'exclusion des sinistres supérieurs à 5700€, ces deux courbes seraient relativement proches (cf. Annexe [8.2.])

Constats :

Deux points peuvent être mis en évidence avec ce graphique :

- les habitations collectives (appartements) ont un coût moyen 6% plus faible que les habitations individuelles (maisons).
- l'intervalle de confiance est fiable.
- à noter tout de même que le pouvoir prédictif de cette variable est moins important que la fréquence (note de « 8/14 »).

Constats :

- Graphiquement, il est aisé de remarquer un coût moyen croissant selon la superficie des dépendances.

- Point intéressant : le critère « superficie de la dépendance » ressort comme la variable la plus significative du modèle (note de « 14/14 »).

Rappel objectif	Points retenus	Étape suivante
Présentation des échelles de risques des coûts moyens	- Les échelles sont moins robustes sur certaines modalités (l'intervalle de confiance est plus large). - L'intégration d'un exercice supplémentaire permettrait de stabiliser davantage les échelles	Analyse sur les résidus pour validation du modèle

Les échelles de risques de coûts moyens ayant été réalisées, il convient de vérifier à présent que les hypothèses émises dans la modélisation proposée soient validées. Tout comme pour la validation des échelles de fréquences, une analyse des résidus a donc été réalisée et est présentée dans le point suivant.

(d) Analyse sur les résidus pour validation du modèle

Les valeurs prédites $\hat{y}_i$ et les nombres de sinistres y_i observés sont cette fois-ci des nombres réels positifs.

L'analyse des résidus est donc « hypothétiquement » plus simple, dans ce sens où il est plus aisé de les « normaliser ».

Il n'est présenté ici que l'étude des résidus bruts (les résultats étant suffisamment probants). Les méthodologies employées sont référencées à l'Annexe [7.].

	Statistiques sur résidus bruts	Constats																																												
Hypothèse d'indépendance	Coefficients de corrélation de Pearson Proba > r sous H0: Rho=0 Nombre d'observations <table><tr><td></td><td>ResBrut</td><td>var_new ResBrut</td></tr><tr><td>ResBrut</td><td>1.00000</td><td>0.02102 0.0001 33863</td></tr><tr><td>var_newResBrut</td><td>0.02102 0.0001 33863</td><td>1.00000 33863</td></tr></table>		ResBrut	var_new ResBrut	ResBrut	1.00000	0.02102 0.0001 33863	var_newResBrut	0.02102 0.0001 33863	1.00000 33863	Le test d'indépendance de Pearson est clairement significatif (pour rappel, ce test expliqué en <i>Annexe [7.2.]</i>). L'hypothèse d'indépendance des résidus est donc validée. N.B. : Le test de corrélation s'effectue entre les résidus ordonnées (noté ResBrut) et eux même décalés d'un pas dans la séquence (noté var_newResBrut).																																			
	ResBrut	var_new ResBrut																																												
ResBrut	1.00000	0.02102 0.0001 33863																																												
var_newResBrut	0.02102 0.0001 33863	1.00000 33863																																												
Hypothèse d'homoscédasticité	Test d'homogénéité des variances The ANOVA Procedure Bartlett's Test for Homogeneity of Pred Variance <table><tr><td>Source</td><td>DDL</td><td>Chi-2</td><td>Pr > Chi-2</td></tr><tr><td>CDGEDL</td><td>129</td><td>45292.5</td><td><.0001</td></tr></table> Brown and Forsythe's Test for Homogeneity of Pred Variance ANOVA of Absolute Deviations from Group Medians <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>124</td><td>1.005E12</td><td>8.1033E9</td><td>14.12</td><td><.0001</td></tr><tr><td>Error</td><td>25524</td><td>1.464E13</td><td>5.7371E8</td><td></td><td></td></tr></table> Levene's Test for Homogeneity of Pred Variance ANOVA of Squared Deviations from Group Means <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>124</td><td>6.529E22</td><td>5.265E20</td><td>3.47</td><td><.0001</td></tr><tr><td>Error</td><td>25524</td><td>3.877E24</td><td>1.519E20</td><td></td><td></td></tr></table>	Source	DDL	Chi-2	Pr > Chi-2	CDGEDL	129	45292.5	<.0001	Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	124	1.005E12	8.1033E9	14.12	<.0001	Error	25524	1.464E13	5.7371E8			Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	124	6.529E22	5.265E20	3.47	<.0001	Error	25524	3.877E24	1.519E20			Les tests de Bartlett, de Brown & Forsythe et Levene sont tous significatifs (ces tests sont expliqués en <i>Annexe [7.3.]</i>). L'hypothèse d'homogénéité des variances est donc validée.
Source	DDL	Chi-2	Pr > Chi-2																																											
CDGEDL	129	45292.5	<.0001																																											
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																									
CDGEDL	124	1.005E12	8.1033E9	14.12	<.0001																																									
Error	25524	1.464E13	5.7371E8																																											
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																									
CDGEDL	124	6.529E22	5.265E20	3.47	<.0001																																									
Error	25524	3.877E24	1.519E20																																											

Tableau 2. 23. Analyse des résidus pour le modèle de coût moyen

N.B. : Comme pour la fréquence, pour ces deux tests, l'analyse a été réalisée sur le critère « CSP »²⁶ du souscripteur (codé « CDGEDL »).

²⁶ CSP : Critère socio-professionnel

Il convient, à présent, d'étudier la normalité des résidus (il est rappelé que les 3 tests utilisés dans le cadre de cette analyse sont référencés dans l'Annexe [7.1.] et [7.1.bis]).

Afin d'obtenir des graphiques lisibles, les histogrammes des résidus et les Droites de Henry ont été produits sur le logiciel Excel. Sachant le nombre de lignes conséquent, il a été décidé de tirer aléatoirement²⁷ un nombre de 20 000 lignes.

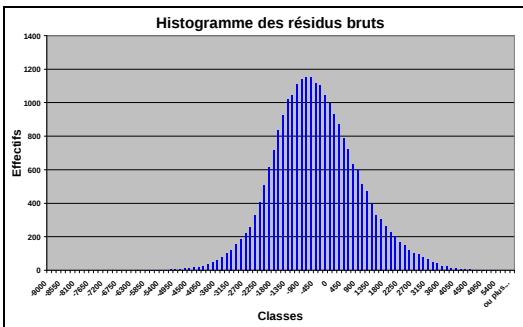
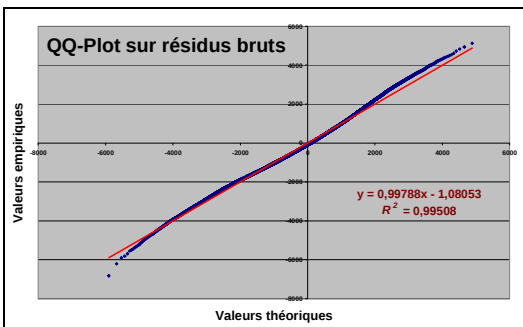
		Sur résidus bruts	Constats												
Hypothèse de normalité	Test de normalité	Tests de normalité <table><thead><tr><th>Test</th><th>-Statistique-</th><th>-p-Value-</th></tr></thead><tbody><tr><td>Kolmogorov-Smirnov</td><td>D 0.061733</td><td>Pr > D <0.0100</td></tr><tr><td>Cramer-von Mises</td><td>W-Sq 52.02995</td><td>Pr > W-Sq <0.0050</td></tr><tr><td>Anderson-Darling</td><td>A-Sq 301.6346</td><td>Pr > A-Sq <0.0050</td></tr></tbody></table>	Test	-Statistique-	-p-Value-	Kolmogorov-Smirnov	D 0.061733	Pr > D <0.0100	Cramer-von Mises	W-Sq 52.02995	Pr > W-Sq <0.0050	Anderson-Darling	A-Sq 301.6346	Pr > A-Sq <0.0050	- Les tests de Kolmogorov-Smirnov (K-S), Cramer-von Mises et Anderson-Darling sont significatifs. - A noter tout de même, que le test de K-S a une probabilité « moins » significative. Comme il l'est notifié en Annexe [7.1.], ce test est sensible à l'existence de points aberrants.
	Test	-Statistique-	-p-Value-												
	Kolmogorov-Smirnov	D 0.061733	Pr > D <0.0100												
Cramer-von Mises	W-Sq 52.02995	Pr > W-Sq <0.0050													
Anderson-Darling	A-Sq 301.6346	Pr > A-Sq <0.0050													
Histogramme des résidus		La représentation graphique a une allure relativement proche de celle d'une loi normale.													
Droite de Henry		- Sur ce <i>QQ-Plot</i>, il est possible d'observer sur la gauche, des valeurs négatives supérieures au 5700 €. Ce constat est somme toute logique car même si les sinistres supérieurs à 5700€ ont été exclus, le modèle peut prédire des valeurs supérieures à ce seuil. - A l'inverse, sur la partie haute de la droite, les valeurs ne dépassent pas les 5700. - Une droite de régression linéaire a été tracée sur cette courbe. Le coefficient de corrélation avoisine les 0,995 pour les résidus bruts. Ce qui démontre une très remarquable normalité des résidus.													

Tableau 2. 24. Analyse des résidus pour le modèle de coût moyen (suite)

L'ensemble des hypothèses ayant été validé, le modèle Gamma semble donc être un modèle adapté pour la modélisation des échelles de coûts moyens.

N.B. : L'analyse des résidus de Pearson n'a pas été montrée, puisqu'il ne s'agit que d'une réduction des valeurs des résidus par la variance.

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Analyse sur les résidus pour validation du modèle	Vérification des hypothèses : - d'indépendance, - d'homoscédasticité, - de normalité.	- Les hypothèses sur les résidus sont validées. - Le modèle choisi est donc adapté à la modélisation des coûts moyens	Déduction de l'échelle de risque de la prime pure

²⁷ Voir le Programme SAS [7.1.] : Test de normalité des résidus de l'Annexe [7.] : Descriptif méthodologique pour l'analyse des résidus

Point de transition :

Le modèle Gamma semble donc être un modèle adapté pour la modélisation des échelles de coûts moyens.

Il est toutefois à noter un intervalle de confiance élevé pour certaines modalités, notamment pour les habitations de plus de 10 pièces (l'intervalle de confiance y est relativement important).

	Etat des lieux	Objectif	Méthodologie	Résultats et/ou Validation des hypothèses	Validation des hypothèses
Détermination des échelles de primes pures	Etape effectuée	<u>Étape 1 :</u> Vérification de l'indépendance des données analysées	- Comparer les distributions empiriques selon le nombre de sinistres déclarés par contrat ; - Comparer les valeurs des coûts moyens observés selon le nombre de sinistres déclarés par contrat ; - Analyser l'indépendance des variables tarifaire.	<u>Résultat :</u> Si aucun biais n'est observé, alors l'indépendance des données est établie. <u>Hypothèse émise :</u> Les échelles sont stables d'un exercice à l'autre.	
	Etape effectuée	<u>Étape 2 :</u> Détermination des échelles de risque de fréquences selon les critères tarifaires	- Vérifier la significativité des facteurs de l'actuel algorithme tarifaire ; - Choisir le modèle le plus adapté à la modélisation de la fréquence.	<u>Résultat :</u> Échelles de fréquences déterminées <u>Hypothèse du modèle à valider :</u> Étude sur les résidus	- Vérifier l'indépendance des résidus ; - Vérifier l'homoscédasticité des résidus ; - Vérifier la normalité des résidus.
	Etape effectuée	<u>Étape 3 :</u> Détermination des échelles de risque des coûts moyens selon les critères tarifaires	- Définir un seuil d'exclusion entre les sinistres attritionnels et les sinistres non-récurrents ; - Vérifier la significativité des facteurs ; - Choisir le modèle le plus adapté à la modélisation des coûts moyens.	<u>Résultat :</u> Échelles des coûts moyens déterminées <u>Hypothèse du modèle à valider :</u> Étude sur les résidus	- Vérifier l'indépendance des résidus ; - Vérifier l'homoscédasticité des résidus ; - Vérifier la normalité des résidus.
	Prochaine étape	<u>Étape 4 :</u> Dédution des échelles de primes pures	- Multiplication des échelles précédemment modélisées.		- Confirmer la stabilité des échelles de risques (hypothèse validée en <i>Annexe [8.1.]</i> par l'utilisation de l'exercice 2011).

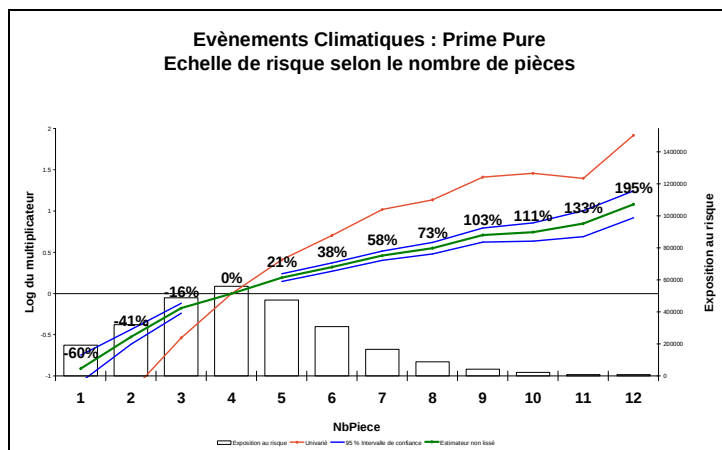
Tableau 2. 25. Récapitulatif de la méthodologie utilisée dans la détermination de l'échelle de primes pures

Ces échelles de fréquences et coûts moyens étant à présent déterminées, il est possible de définir les échelles de primes pures.

II.1.4 Dédution des échelles de prime pure

Par définition et comme il l'a été énoncé au chapitre I.3., la prime pure correspond à la multiplication de la fréquence et du coût moyen. Ainsi, les échelles de primes pures présentées ci-dessous correspondent au produit des paramètres des fréquences et coûts moyens déterminés dans les chapitres précédents.

Il ne sera présenté ici que les échelles de primes pures suivantes des variables tarifaires « nombre de pièces », « type de bâtiment » et « dépendance » :

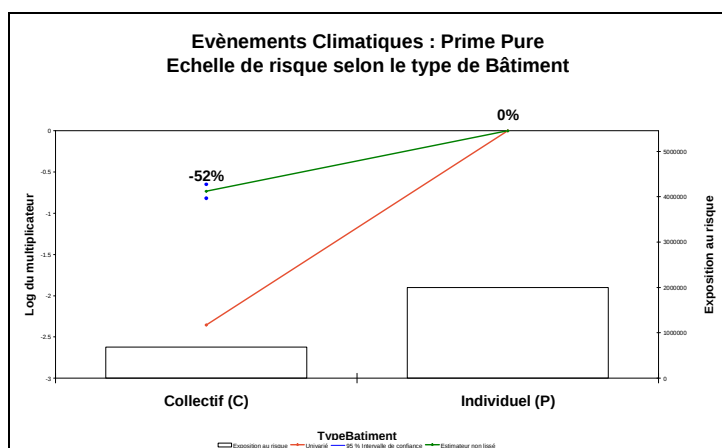


Graph 2. 10. Echelle de risque de la prime pure selon le nombre de pièces

Constats :

Le graphique met en évidence :

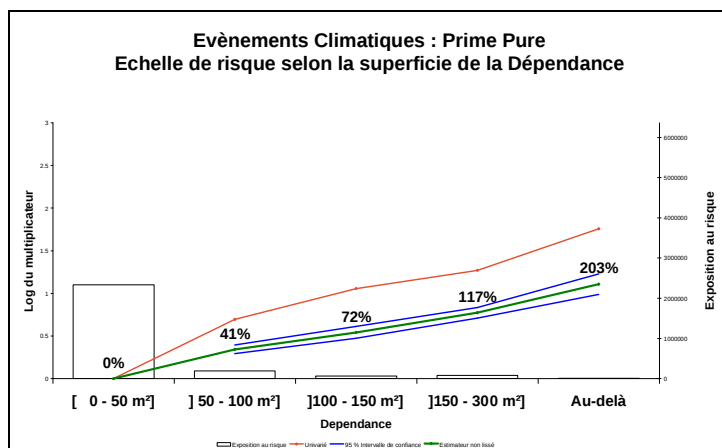
- un coût du risque croissant en fonction du nombre de pièces dans l'habitation ;
- une prédiction du critère relativement robuste du fait du faible intervalle de confiance (hormis peut être sur les habitations de 10 à 12 pièces et 1 pièce où l'intervalle de confiance est un peu plus important. Effets liés à l'échelle des coûts moyens où la volatilité était plus grande).



Graph 2. 11. Echelle de risque de la prime pure selon le type de bâtiment

Constats :

- Il est possible de remarquer que les habitations collectives (appartements) ont un coût du risque 52% plus faible que celui des habitations individuelles ;
- Le pouvoir prédictif semble robuste (faible intervalle de confiance)



Graph 2. 12. Echelle de risque de la prime pure selon la superficie de la Dépendance

Constats :

Le graphe suivant met en évidence :

- un coût du risque croissant suivant la superficie de la dépendance ;
- une robustesse de l'échelle par rapport à l'intervalle de confiance.

Point de transition :

Les allures de courbes semblent relativement stables et pertinentes. Il est donc possible de clore ce chapitre sur l'évaluation des échelles du coût du risque.
A présent, il convient de fixer le niveau de cette prime pure.

N.B. : Les échelles ci-dessus ont été retraduites (transformées) en paramétrage de saisie pour le système de gestion SIGMA afin que les Entités puissent tester cette nouvelle grille tarifaire.

II.2. Modélisation des « Forces de la Nature » dans le modèle interne

Les travaux présentés ci-dessous ont été réalisés par la Direction Actuariat Groupe. Plus précisément, dans le pôle Non-Vie de la Direction Actuariat Groupe, dont Hervé Pinaud, tuteur de ce mémoire, est le responsable.

Avant de poursuivre, il convient de réaliser un bref chapitre sur les raisons de la mise en place d'un modèle interne au sein de Groupama.

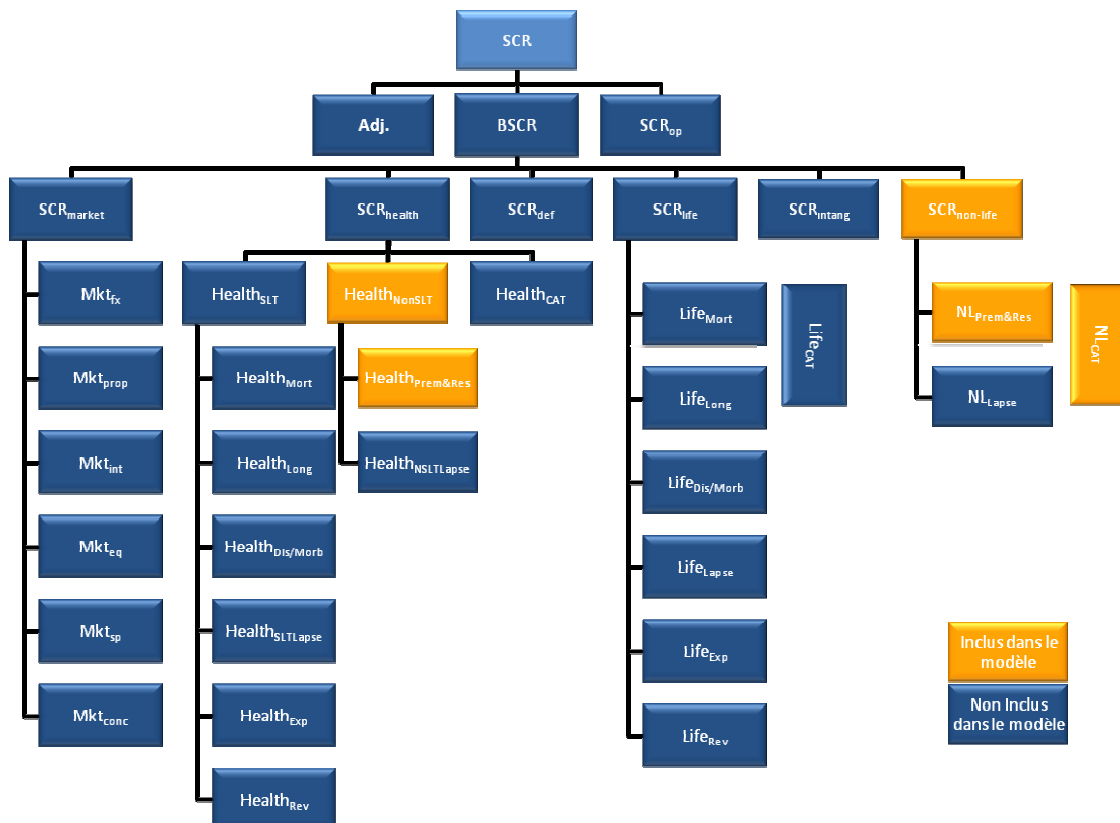
II.2.1. Les raisons d'un modèle interne

A ce jour, Groupama dispose d'un modèle interne non vie et a pour objectif d'être utilisé dès l'entrée en vigueur de Solvabilité 2 (S2). Ce modèle interne sera notamment utilisé pour évaluer l'exigence de capital²⁸ (appelée « SCR ») du Groupe et de deux de ses Entités constituantes.

Le développement de ce modèle interne répond à une volonté du Groupe, axée sur 3 aspects :

- refléter au mieux le profil des principaux risques portés par Groupama et prendre en compte la structure spécifique du Groupe, afin d'intégrer, pour l'ensemble des décisions stratégiques, leurs impacts sur l'évolution du profil de risques et les besoins en fonds propres ;
- capter au mieux les effets de diversification liés à la structure du Groupe et prendre en compte les particularités de la réassurance externe pour le calcul d'exigence de capital réglementaire ;
- répondre aux besoins de l'« ERM/ORSA »²⁹ en intégrant le modèle de gestion des risques dans les processus de pilotage du Groupe, tels que la Planification Stratégique Opérationnelle³⁰ (PSO), les études de rentabilité, les processus de souscription, la politique de réassurance,...

Le *Graphe 2.13.* ci-dessous se réfère à la structure de calcul du SCR décrite dans les projets de mesures d'implémentation consolidés par la Commission Européenne en octobre 2010.



Graphe 2. 13. Structure de calcul selon l'implémentation de la Commission Européenne

²⁸ Un point plus précis sur cette exigence est réalisé dans la Section III.

²⁹ ERM : Enterprise Risk Management / ORSA : Own Risk Solvency Asset

³⁰ La PSO est une vision projetée à horizon 3 ans (un « Business Plan »)

La détermination du *SCR* (capital requis) repose donc sur une structure modulaire de risques. En l'occurrence, ce sont six modules de risques composés chacun de sous-modules qui conduisent à la détermination de ce capital.

Pour chaque sous-module, une approche basée sous forme de « choc » permet de leur déterminer un capital requis. Selon des matrices de corrélations qui traduisent la dépendance des risques et leur diversification, l'agrégation de ces sous-modules permet d'aboutir au capital requis final réglementaire.

Le *Graphe 2.13*. ci-dessus présente :

- en jaune, les modules et sous-modules de risques inclus dans le modèle interne :
 - risques de primes et de réserves en non vie et santé similaire à la non-vie
 - risques de catastrophe en non vie
- en bleu, les modules et sous-modules de risques non modélisés et qui feront l'objet d'un calcul de *SCR* en formule standard :
 - risques de marché
 - risques de souscription vie et santé similaire à la vie
 - risques de catastrophe en santé
 - risques de rachat en non vie et santé similaire à la non vie
 - risques de défaut de contrepartie
 - risques opérationnels.

Dans le cadre de ce rapport, une précision est à apporter concernant le risque « CAT » :

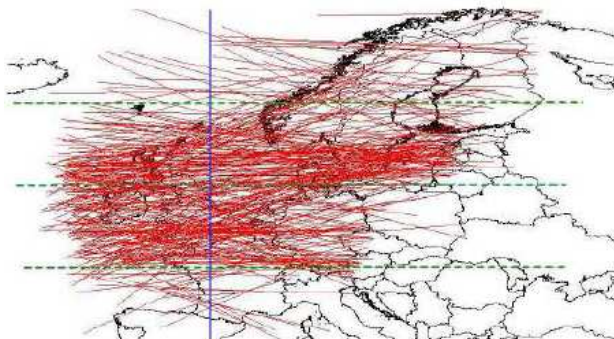
- Tout d'abord, par rapport au *QIS4*, l'architecture du *QIS5*³¹ a été revue avec une décomposition en 4 sous-modules dans le but d'assurer une meilleure prise en compte de la réassurance. Le chapitre II.2.3. traite spécifiquement de la couverture de réassurance de Groupama sur le risque Tempête.

- Deuxièmement, aujourd'hui, chacun des assureurs européens évalue individuellement l'exposition de son portefeuille de couvertures d'assurances tempêtes. Pour cela, les acteurs utilisent soit des modèles internes, soit des logiciels du marché. Mais ces approches donnent des réponses divergentes étant donné les incertitudes sur les champs de vent, les temps de retour et la relation vent-dommages propre à chaque outil.

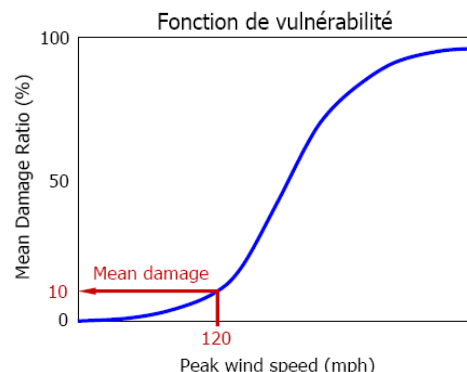
A ce jour, le Groupe a fait le choix d'utiliser le logiciel RMS³² pour modéliser ses événements atypiques. Ce logiciel fonctionne par module et permet de :

- modéliser l'évènement : les trajectoires des tempêtes sont créées en fonction de la vitesse et de la direction historique des tempêtes passées. Le cycle de vie complet d'une tempête, de son intensification à son extinction est pris en compte.

- modéliser « l'aléa » : l'aléa intègre à la fois l'évènement lui-même ainsi que tous les facteurs qui influencent l'évènement (par exemple, la topographie du sol)



- modéliser la vulnérabilité : ce module prévoit des dommages en fonction de l'intensité de l'aléa (comme la vitesse maximale du vent). La vulnérabilité d'un bâtiment ou son contenu est caractérisée par une courbe de destruction. Pour les tempêtes, le risque est quantifié par la vitesse de pointe de vent. Les fonctions de vulnérabilité sont définies par des caractéristiques de construction clés (Type de couverture - bâtiments, contenus, pertes d'exploitation -, Type d'occupation, Classe de construction, Hauteur du bâtiment, Année de construction).



³¹ *QIS5* : « *Quantitative Impact Study* », Etude Quantitative d'Impact. Depuis un certain nombre d'années, des travaux ont été effectués à travers l'Europe pour une révision et une amélioration substantiellement du régime d'assurance et de sa supervision : Solvabilité 2. Le développement de Solvabilité 2 est testé régulièrement à travers ces *QIS*. 2010 a été l'année de la cinquième édition, dite « *QIS5* ».

³² RMS : Logiciel permettant d'appréhender de manière globale l'exposition d'un portefeuille pour un péril climatique.

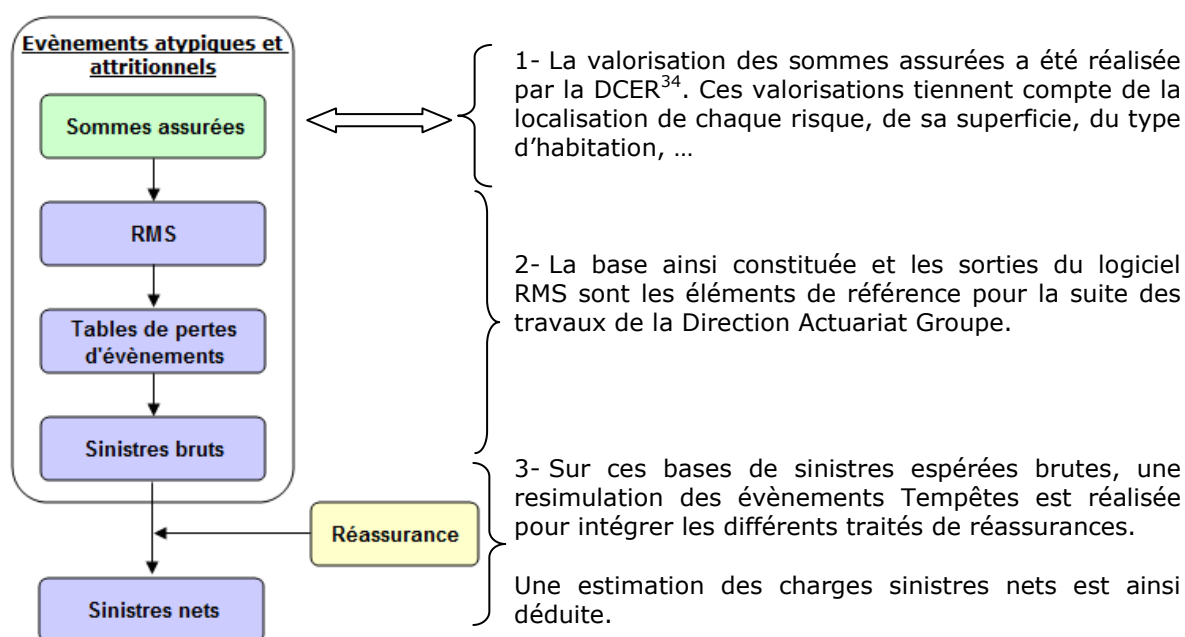
Rappel objectif	Points retenus	Étape suivante
Les raisons d'un modèle interne	- Refléter au mieux le profil des principaux risques portés par Groupama et prendre en compte la structure spécifique du Groupe - Le modèle interne couvre les modules de souscription non vie et Santé similaire à la non vie (l'ensemble des périls est inclus dans ces modules). - La structure du modèle interne couvre toutes les lignes d'activité (nommées « LoB » (<i>Line of Business</i>)) du QIS5 de Solvabilité 2, notamment les risques de primes et de réserves.	Processus de modélisation des événements « Forces de la Nature »

Face à la structure spécifique du Groupe et aux caractéristiques liées au programme de réassurance « Force de la Nature », il est exposé ci-après, le processus de modélisation qui a été élaboré et qui a permis de quantifier l'exposition au risque « Tempête » des diverses Entités concernées.

II.2.2. Processus de modélisation des événements « Forces de la Nature »

Les contraintes de réassurance ont imposé une modélisation du périmètre des Caisses Régionales, mais également des périmètres de Gan Assurances, Gan Eurocourtage et Groupama Insurances. En effet, le traité « Tempête » Groupe (appelé également « Force de la Nature ») fonctionne en excédent de sinistres³³ (à l'évènement) et couvre l'ensemble de ces Entités (y compris l'ensemble des périls Agricole, Collectivité, ... et Serres pour les Caisses Régionales).

Pour chaque Entité, le processus de modélisation du risque Tempête peut être décrit graphiquement, selon la démarche suivante :



Graph 2. 14. Mécanisme de modélisation des forces de la nature

En quatrième étape, afin d'être en adéquation avec la tarification des événements climatiques du produit Privatis, des clés de répartition seront élaborées pour répartir les charges sinistres modélisées. Ces clés de répartition seront retranscrites dans le chapitre II.4..

Rappel objectif	Points retenus	Étape suivante
Processus de modélisation des événements « Forces de la Nature »	- Forte interaction entre les Directions - Non maîtrise des valorisations (spécificité d'indépendance entre les parties prenantes)	Présentation des sorties de modélisation des événements atypiques et attritionnels via l'outil de marché RMS

Les modélisations réalisées à partir du logiciel RMS et les sorties fournies par ce logiciel sont traitées dans le chapitre II.2.3. et II.2.4.

³³ Les 4 principaux mécanismes de réassurance sont rappelés en Annexe [9.]

³⁴ DCER : Direction Cessions Externes de Réassurance

II.2.3. Modélisation des événements atypiques et attritionnels pour le risque « Tempête » via l'outil de marché RMS

Comme il l'a été signalé, le logiciel RMS a permis de modéliser des événements atypiques et des événements attritionnels.

Etant donné que le traité « Force de la Nature » fonctionne en excédent de sinistre et couvre un ensemble d'Entités et de périls, il est fourni, en « *input* », une base avec tous les contrats qui ont été présents lors de la clôture de l'exercice 2010 (dans cette base, la DCER a valorisé la somme couverte pour chaque entité assurée).

En « *output* », le logiciel RMS produit, pour chaque Entité et chaque péril, une liste d'événements catastrophes susceptibles de se produire dans une année, avec notamment une fréquence et un montant de sinistre associés liés à chacun de ces événements. Le tout est présenté dans une « *Event Loss Table* ».

Ces tables de sortie contiennent entre autres :

- l'identifiant de chaque tempête (« *EVENT ID* »),
- sa fréquence annuelle (« *RATE* »),
- le coût moyen qu'elle engendre (« *PERSPVALUE* »),
- l'écart-type de ces coûts indépendants et corrélés (respectivement « *STDDEVI* » et « *STDDEVC* »),
- le coût maximal, c'est-à-dire l'exposition (« *EXPVALUE* ») (ou bien encore selon la littérature le SMP : Sinistre Maximum Probable).

Les deux tables suivantes sont des exemples d'événements simulés par RMS, pour deux Entités couvertes par le programme « Force de la Nature » :

	EVENTID	RATE	PERSPVALUE	STDDEVI	STDDEVC	EXPVALUE
Entité A	859011	0,0000000001	486 624	141 959	148 298	847 678 845
	859020	0,0000000001	13 637 078	468 844	6 800 354	48 907 562 763
	859030	0,0001413000	573 640	86 824	531 401	18 947 627 586
	859031	0,0000353300	16 930 137	744 638	8 824 320	49 050 351 072
	859032	0,0000353300	74 291 096	2 025 181	24 520 185	49 166 778 316
	EVENTID	RATE	PERSPVALUE	STDDEVI	STDDEVC	EXPVALUE
Entité B	859010	0,0000000001	122 501 483	2 607 540	49 493 389	143 082 506 260
	859011	0,0000000001	16 889 676	1 040 958	8 749 342	76 455 554 877
	859030	0,0001413000	2 085 569	180 523	1 783 066	53 195 122 563
	859031	0,0000353300	104 595 608	2 281 345	44 791 348	143 006 383 227
	859032	0,0000353300	76 150 233	2 182 600	30 995 838	143 053 601 015

Tableau 2. 26. Exemple de sortie du logiciel RMS pour deux Entités

Constats :

- Un événement donné peut impacter une seule ou plusieurs Entités simultanément. C'est le cas de l'événement 859031 dans l'exemple ci-dessus. Cet événement impacte l'Entité A et B, avec une fréquence annuelle commune de 0,003533% et une loi de sévérité spécifique à l'exposition de chaque Entité.
- L'écart-type se divise en une composante indépendante et une composante corrélée. La composante corrélée permet de tenir compte de la concentration géographique des événements au sein du portefeuille : plus les risques assurés sont géographiquement proches, plus l'écart-type du coût est important. L'écart-type total correspond à la somme de ces deux composantes.
- Sur chaque police (contrat), les montants qui ressortent de ces tables sont bruts de réassurance, nets de franchises et de limites de couverture.

Les lois de fréquence et le taux de destruction dommage sont fixés (plus précisément imposés) par RMS : la fréquence suit une loi de Poisson et le taux de destruction dommage suit une loi Bêta (taux qui est appliqué sur les valorisations effectuées par la DCER).

Rappel objectif	Points retenus	Étape suivante
Modélisation des événements atypiques et attritionnels pour le risque Tempête via l'outil de marché RMS	- RMS fournit des tables de sorties de différents événements susceptibles de se produire au cours d'un exercice avec notamment des événements extrêmes dont la survenance n'est que d'une fois tous les 200 ans. - Les lois de fréquence et le taux de destruction dommage sont fixés (imposés) par RMS : la fréquence suit une loi Poisson et le taux de destruction dommage suit une loi Bêta.	Synthèse des résultats sur l'ensemble des risques couverts par le traité

Le fonctionnement du logiciel RMS étant à présent expliqué, une synthèse des résultats est proposée dans le chapitre suivant.

II.2.4. Synthèse des résultats sur l'ensemble des risques couverts par le traité

Il est présenté une vision globale des Caisses Régionales (Hors GLBR) et les résultats de deux Entités (GCA et GSUD), notamment pour leur opposition géographique.

Perte GCA	Sinistres Bruts
Espérance	16,2
Ecart-type	35,6
Coefficient de variation	219,7%
Période de retour 1/10	25,6
Période de retour 1/20	47,5
Période de retour 1/50	101,1
Période de retour 1/100	152,5
Période de retour 1/200	230,3

Tableau 2. 27. Résultat des modélisations RMS pour GCA

Perte GSUD	Sinistres Bruts
Espérance	7,0
Ecart-type	15,0
Coefficient de variation	212,7%
Période de retour 1/10	9,4
Période de retour 1/20	12,6
Période de retour 1/50	26,2
Période de retour 1/100	48,2
Période de retour 1/200	83,6

Tableau 2. 28. Résultat des modélisations RMS pour GSUD

Perte Toutes CR (hors GLBR)	Sinistres Bruts
Espérance	126,8
Ecart-type	189,5
Coefficient de variation	149,4%
Période de retour 1/10	189,2
Période de retour 1/20	317,8
Période de retour 1/50	617,2
Période de retour 1/100	927,4
Période de retour 1/200	1 335,6

Tableau 2. 29. Résultat des modélisations RMS pour l'ensemble des Caisses Régionales (hors GLBR)

Constats :

Sur ces 3 tableaux, il est possible de remarquer de forts coefficients de variations, notamment au niveau Caisse Régionale. Au niveau national, ce constat est moins prononcé.

Cette observation permet de mettre en exergue les différents attraits de la mutualisation (tant au niveau des périls que des portefeuilles).

Les tableaux présentés ici sont les résultats de la somme de tous les périls couverts par les différentes Caisses Régionales.

N.B. : à noter que RMS fournit des tableaux spécifiques pour chaque péril et Entité. C'est un aspect qui est précisé car il permettra de constituer par la suite les clés de répartition.

Rappel objectif	Points retenus	Étape suivante
Synthèse des résultats sur l'ensemble des risques couverts par le traité	- De forts coefficients de variations sont observables, notamment au niveau Caisse Régionale. - Au niveau national, ce constat est moins prononcé et montre l'un des attraits de la mutualisation des risques et portefeuilles.	Prise en compte de la réassurance sur les événements extrêmes modélisés

Le logiciel RMS ne prend pas en compte les différents programmes de réassurance interne et externe. Une resimulation des événements a donc été réalisée via le logiciel « Risk Explorer »³⁵.

De ce fait et avant de poursuivre, il est exposé brièvement le mécanisme de réassurance des Caisses Régionales, ainsi que le traité externe « Force de la Nature » de Groupama.

II.3. Prise en compte de la réassurance sur les événements extrêmes modélisés

Avant de développer le programme de réassurance³⁶ externe de Groupama SA et la façon dont il a été intégré, il faut savoir que les différentes Entités du Groupe, notamment les Caisses Régionales, se réassurent exclusivement auprès de Groupama SA.

Cet aspect permet de négocier vis-à-vis des réassureurs, des traités moins coûteux au niveau Groupe, tout en faisant bénéficier de traités avantageux au niveau Solo (Entité).

Seul le programme de réassurance des Caisses Régionales sera décrit dans le point suivant.

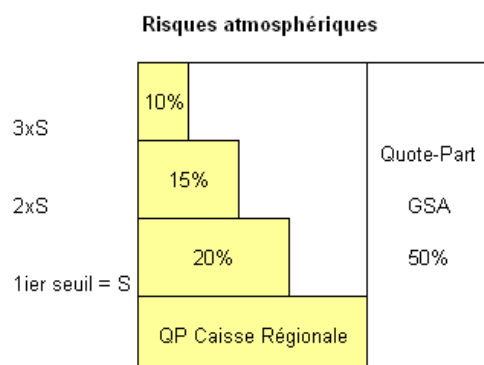
N.B. : Gan Assurances, Gan Eurocourtage et Groupama Insurance possèdent des programmes spécifiques (différents de celui des Caisses Régionales).

³⁵ Logiciel permettant de modéliser et quantifier les différents risques auxquels les entreprises sont confrontées

³⁶ Il est rappelé brièvement en Annexe [9.], les 4 principaux produits de réassurance.

II.3.1. Description du fonctionnement de la réassurance interne pour les Caisses Régionales

Pour les risques atmosphériques, les Caisses Régionales ont un traité en « *Stop Loss* », décrit graphiquement comme suit :



Le programme se compose d'un traité en Quote-Part à 50% et de « *Stop Loss* » successifs. Le seuil du « *Stop Loss* » est à un seuil S de 140%.

Ce traité s'applique notamment sur la garantie « événements climatiques » décrite dans ce rapport.

Graphie 2. 15. Description du traité de réassurance interne des Caisses Régionales

II.3.2. Description du programme de réassurance externe

Concernant le traité de réassurance externe de Groupama SA, le programme « Forces de la Nature Groupe » est un traité en excédent de sinistres par événement. Il peut être traduit par le graphique ci-dessous :

XS Groupe T 7 200M EUR xs 1 980M EUR 1 @ 100%
XS Groupe T 6 150M EUR xs 1 830M EUR 0 rec
XS Groupe T 5 250M EUR xs 1 580M EUR 1 @ 100%
XS Groupe T 4 430M EUR xs 1 150M EUR 1 @ 100%
XS Groupe T 3 350M EUR xs 800M EUR 2 @ 100%
XS Groupe T 2 300M EUR xs 500M EUR 2 @ 100%
XS Groupe T 1 260M EUR xs 240M EUR 2 @ 100%
XS Groupe Sous-jacente complémentaire 40M EUR xs 200M EUR xs 80M EUR 0 rec
XS Groupe Sous-jacente 40M EUR xs 200M EUR 1 @ 25%
XS Groupe Aggregate 100M EUR xs 100M EUR xs 50M EUR 1 @ 0%
Rétention

Comme il est possible de le constater, ce programme se décompose en plusieurs tranches, avec des reconstitutions différentes selon les tranches.

Dans un XS sans reconstitution, le Réassureur s'engage à payer au maximum une fois la portée de l'XS. Mais, en général, il est prévu dans un traité, un certain nombre n de reconstitutions (voire « illimité »). Dans ce cas, cela signifie que le réassureur paiera au maximum $(n+1) \times \text{portée}$ où n désigne le nombre de reconstitutions.

De plus, en cas de survenance d'un sinistre, si cette « garantie » supplémentaire est accordée, une prime dite de reconstitution peut être activée. La Cédante (ici Groupama SA) doit verser au réassureur une prime supplémentaire en cas de survenance du sinistre.

Le traité prévoit le taux de prime de ces reconstitutions (qui peuvent être gratuites). Ce taux correspond à un pourcentage par rapport à la prime de réassurance originale. Ainsi, par exemple, pour l'« XS Groupe Sous-jacente », une prime de reconstitution de 25% de la prime de réassurance de départ sera versée par Groupama SA en cas de mise en jeu de la garantie.

Ces reconstitutions peuvent être calculées au *prorata capita* (proportionnellement au montant à reconstituer), au *prorata temporis* (proportionnellement au temps restant à couvrir jusqu'à la fin de la période de couverture) ou au *double prorata*.

Graphie 2. 16. Description du traité externe « Force de la Nature »

Dans ce programme, les reconstitutions sont calculées au *simple prorata capita*, quel que soit le moment de l'année où elles sont payées. Pour des raisons de confidentialités, les valeurs de ces primes de reconstitution ne sont pas présentées. Néanmoins, il figure en *Annexe [10.]* un « pseudo schéma » rappelant le mécanisme de coût des primes de reconstitution.

Pour palier ces fluctuations de primes de reconstitution, une protection « *RPP* » de prime de reconstitution a été souscrite sur les tranches T1 à T4 (parties grisées sur le *Graphie 2.16.*).

Pour les deux tranches les plus basses, une franchise annuelle de 50M€ doit être prise en compte. Enfin, il est important de savoir que les sinistres provenant des Caisses Régionales entrent à hauteur de 100% dans cette couverture (cf. chapitre II.3.1.), tandis que les sinistres de Gan Assurances et Gan Eurocourtage entrent à hauteur de 85% et ceux de Groupama Insurances à hauteur de 75%.

Rappel objectif	Points retenus	Étape suivante
Description du programme de réassurance externe	- Le programme de réassurance externe Groupe « Force de la Nature » couvre plusieurs Entités et plusieurs risques - Ce programme est constitué de plusieurs traités en excédent de sinistres (dont 4 sont couverts par un traité <i>RPP</i>).	Resimulation des événements tempêtes sur la plateforme Risk Explorer pour intégration des différents programmes de réassurance

Afin de prendre en compte ces différents traités, des simulations de Monte Carlo³⁷ ont été réalisées via l'outil Risk Explorer.

II.3.3. Resimulation des événements tempêtes sur la plateforme Risk Explorer

Le logiciel Risk Explorer est une interface qui permet d'utiliser les tables issues du logiciel de modélisation d'événements naturels RMS. Cette plate-forme va permettre de prendre en compte l'architecture spécifique du Groupe, ainsi que les programmes de Réassurance de chaque Entité.

Etant donné que la fréquence des événements suit une loi de Poisson, il est possible, à partir du champ « *RATE* », d'évaluer le paramètre de cette loi (le nombre annuel moyen d'événements) en sommant les fréquences annuelles de chaque événement.

Pour une loi de Poisson définie telle que $P(N = n) = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$, le paramètre λ peut être défini tel que :

$$\lambda = \sum_i rate_i, \text{ où } rate_i \text{ est la fréquence annuelle de l'évènement } i$$

Dès lors, il est possible de générer de nouveau des fréquences de sinistralités tempêtes, à partir des fréquences annuelles déterminées et extraites des tables « *Event Info Set* ».

Pour un événement donné, il faut également resimuler un taux d'endommagement *DR* sur tous ces événements. Sachant que la loi utilisée par RMS est une loi Bêta³⁸, il est également possible de calculer les paramètres α et β de cette loi.

Pour obtenir les paramètres de la loi Bêta, les sorties du logiciel RMS sont de nouveau utilisées :

- le coût moyen de l'évènement :

$$CM_{Total} = \sum_{k=1}^K CM_{Entité k}, \text{ où } CM \text{ représente le coût moyen de l'évènement et } K \text{ le nombre d'Entités touchées par cet évènement.}$$

- les écart-types indépendants et corrélés :

$$\sigma_{i,Total} = \sqrt{\sum_{k=1}^K \sigma_{i,Entité k}^2} \text{ où } \sigma_i \text{ représente l'écart-type indépendant}$$

$$\sigma_{c,Total} = \sum_{k=1}^K \sigma_{c,Entité k} \text{ où } \sigma_c \text{ représente l'écart-type corrélé}$$

³⁷ La méthode de Monte Carlo consiste à évaluer une quantité à partir de tirages (expériences) aléatoires répétés. Une simulation de Monte Carlo a l'avantage de fournir les écarts-types des estimations produites.

³⁸ Rappel : Si X suit une loi $Beta(\alpha, \beta)$ alors sa fonction de densité s'écrit $f_{(\alpha, \beta)}(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$,

avec $x \in [0;1]$, $\alpha > 0$, $\beta > 0$ et $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$.

- l'écart-type total est donné par :

$$\sigma_{Total} = \sigma_{I,Total} + \sigma_{C,Total}$$

- l'exposition totale est obtenue en sommant les expositions par Entité :

$$Expo_{Total} = \sum_{k=1}^K Expo_{Entité\ k}$$

Ainsi, la moyenne et l'écart-type du taux d'endommagement valent respectivement :

$$E(DR) = \frac{CM_{Total}}{Expo_{Total}} = \frac{\alpha}{\alpha + \beta}$$

$$\sigma(DR) = \frac{\sigma_{Total}}{Expo_{Total}} = \sqrt{\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}}$$

Il vient les paramètres de la loi Bêta :

$$\alpha = \frac{E(DR)^2[1 - E(DR)]}{\sigma(DR)^2} - E(DR)$$

$$\beta = \frac{\alpha}{E(DR)} - \alpha$$

Le taux d'endommagement est multiplié à l'exposition totale afin d'obtenir un coût total d'évènement. Ce coût est ensuite réparti entre les Entités proportionnellement et conformément à la contribution de chacune d'elles.

Par le biais des identifiants des tempêtes, la corrélation entre les Entités est directement prise en compte, y compris la corrélation France / UK. En effet, les identifiants « *EVENT ID* » sont les mêmes pour les Entités françaises et la filiale britannique.

Pour les 12 Entités françaises (les 10 Caisses Régionales, Gan Assurances et Gan Eurocourtage), une distribution d'OEP³⁹ (par évènement) et d'AEP⁴⁰ globale (charge annuelle) est ainsi obtenue.

Résultats :

- L'ensemble de ces simulations a permis d'estimer le nombre de sinistres qui dépasserait le seuil de 100M€. Ce nombre est en moyenne de 0,15 sinistre.

- Le second constat effectué est que 1 fois tous les 200 ans, 2 sinistres dépassent le seuil des 100M€. Le traité de réassurance « Forces de la Nature » de Groupama prévoit une reconstitution complémentaire (l'« XS Groupe Sous-jacente complémentaire »), ce qui permet la couverture d'un 3^{ème} évènement au cas où celui-ci surviendrait.

- Enfin, pour conclure, sur une période bicentenaire, un évènement pourrait atteindre le montant maximum des 2180M€. Dès lors, la capacité maximale du traité « Forces de la Nature » a été ajustée à cette valeur.

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Resimulation des évènements tempêtes sur la plateforme Risk Explorer (pour intégration des programmes de réassurance)	- Détermination des paramètres des lois de modélisation de l'outil RMS - Resimulation des évènements avec intégration des programmes de réassurance via l'outil Risk Explorer.	- L'ensemble des simulations a permis d'estimer le nombre de sinistres qui dépasserait le seuil de 100M€. Ce nombre est en moyenne de 0,15 sinistre. - Une fois tous les 200 ans, 2 sinistres dépassent le seuil des 100M€. - D'un point de vue prudentiel le traité de réassurance « Forces de la Nature » de Groupama prévoit une reconstitution complémentaire, ce qui permet la couverture d'un 3 ^{ème} évènement. - Sur une période bicentenaire, un évènement pourrait atteindre le montant maximum des 2180M€. La capacité maximale du traité a été ajustée à cette valeur.	Synthèse des résultats après application des programmes de réassurance

Il convient maintenant de présenter les résultats de ces simulations avec intégration du programme de réassurance.

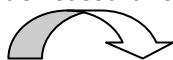
³⁹ OEP : « Occurrence Exceedence Probability » : Distribution d'un évènement individuel

⁴⁰ AEP : « Agregate Exceedence Probability » : Distribution de la charge sinistre annuelle

II.3.4. Synthèse des résultats après application des programmes de réassurance

Après application des programmes de réassurance, il advient des événements modélisés précédemment, les résultats suivants (pour les deux Entités GCA et GSUD, et Caisses Régionales (Hors GLBR)).

Application des programmes
de réassurance



Perte GCA	Sinistres Bruts	Sinistres Nets
Espérance	16,2	12,5
Ecart-type	35,6	16,9
Coefficient de variation	219,7%	134,7%
Période de retour 1/10	25,6	22,8
Période de retour 1/20	47,5	33,2
Période de retour 1/50	101,1	51,4
Période de retour 1/100	152,5	70,1
Période de retour 1/200	230,3	87,2

Constats :

- Pour l'Entité GCA, les effets de la réassurance atténuent non seulement de 22% l'espérance des coûts des sinistres, mais il diminue considérablement l'écart-type.
- Pour une période de retour de 200 ans, sur l'évènement majeur, l'impact serait réduit de 62,1%.

Tableau 2. 30. Résultat des modélisations après application des programmes de réassurance pour GCA

Perte GSUD	Sinistres Bruts	Sinistres Nets
Espérance	7,0	7,0
Ecart-type	15,0	6,1
Coefficient de variation	212,7%	86,6%
Période de retour 1/10	9,4	10,2
Période de retour 1/20	12,6	12,9
Période de retour 1/50	26,2	20,4
Période de retour 1/100	48,2	29,1
Période de retour 1/200	83,6	37,3

Constats :

- Pour l'Entité GSUD, les effets de la réassurance atténuent uniquement de 0,2% l'espérance des coûts des sinistres. Néanmoins, il est également possible de remarquer une diminution drastique du coefficient de variation.
- Pour une période de retour de 200 ans, sur l'évènement majeur, l'impact serait réduit de 55,5%.

Tableau 2. 31. Résultat des modélisations après application des programmes de réassurance pour GSUD

Perte Toutes CR (hors GLBR)	Sinistres Bruts	Sinistres Nets
Espérance	126,8	106,4
Ecart-type	189,5	81,6
Coefficient de variation	149,4%	76,7%
Période de retour 1/10	189,2	165,9
Période de retour 1/20	317,8	195,1
Période de retour 1/50	617,2	231,9
Période de retour 1/100	927,4	265,2
Période de retour 1/200	1 335,6	322,5

Constats :

- Pour l'ensemble des Caisses Régionales hors GLBR, les effets de la réassurance atténuent de 16,1% l'espérance des coûts des sinistres.
- Le coefficient de variation et l'écart-type sont très nettement réduits et sont le résultat du programme de réassurance.
- Pour une période de retour de 200 ans, sur l'évènement majeur, l'impact serait réduit de 75,8%.

Tableau 2. 32. Résultat des modélisations après application des programmes de réassurance pour les CR (Hors GLBR)

La tarification actuelle du produit étant une tarification élaborée au niveau national, la géolocalisation n'avait pas été prise en compte à ce stade.

Aujourd'hui, de par l'espérance des charges calculée par Entité (cf. ci-dessus), les Caisses Régionales ont la possibilité d'avoir une mutualisation adaptée à leur portefeuille et de connaître, notamment, leur exposition au risque Tempête.

Rappel objectif	Points retenus	Étape suivante
Synthèse des résultats après application des programmes de réassurance	- Réduction de la charge sinistre espérée grâce au programme de réassurance « Force de la nature » - Réduction de la volatilité du risque	Intégration de la facturation du coût de la réassurance dans la prime pure

Comme il est possible de le constater, l'espérance de charge sinistre est réduite (tout comme la volatilité), mais ces « réductions » ont un coût.

Le chapitre suivant présente la démarche qui a été utilisée pour intégrer ce coût de facture de réassurance.

II.3.5. Intégration de la facturation du coût de la réassurance

Dans le compte de participation aux résultats, une rubrique intitulée « Solde de réassurance cédée » est prévue.

Ce « solde de réassurance cédée » correspond à la différence entre le montant des sinistres à la charge des cessionnaires (les différentes Entités couvertes par le programme) et celui des cotisations cédées aux réassureurs.

Selon le cas, ce solde est donc inscrit au débit (valeur négative) ou au crédit (valeur positive) du compte de participation aux résultats.

Ce solde de réassurance doit donc être pris en compte. En effet, il est nécessaire de financer les primes cédées du programme « Force de la Nature ». Ainsi, deux cas se présentent quant à la réinjection de ce coût :

- Soit le même solde est affecté à toutes les Entités,
- Soit, une notion géographique est prise en compte, avec une allocation de la prime de réassurance et de la charge cédée, en fonction de la contribution de chaque Entité.

C'est cette deuxième option qui a été privilégiée.

N.B. : Il n'est pas spécifiquement opportun de répartir ces coûts au prorata des cotisations. Certaines Entités peuvent avoir volontairement sous-tarifé la garantie « événements climatiques » pour être compétitif sur certains profils de risque, ou autres raisons.

Rappel objectif	Points retenus	Étape suivante
Intégration de la facturation du coût de la réassurance	- Nécessité de financer le coût de la facture de réassurance - Allocation des factures de réassurance selon la contribution de chaque Entité des charges simulées.	Clés de répartition des événements modélisés entre les risques et produits couverts par le programme « Force de la Nature »

Pour rappel, Groupama paie une prime de réassurance contre environ de 50M€ de charge conservée (« *XS Groupe Aggregate* »). Au quantile 1/200 ans, les portées de la plupart des tranches du programme sont consommées au moins une fois. Il est donc nécessaire de répartir également les différentes primes de reconstitution, dont notamment le traité *RPP* payé annuellement (traité qui couvre les primes de reconstitution des tranches T1 à T4).

Ces différentes primes de reconstitution font là aussi l'objet d'une réallocation conformément à la contribution de chacune des Entités dans les événements simulés.

Le coût de facture de réassurance est traité sous forme de chargement de la prime pure. Un aperçu de ces propos sera traité au chapitre III.3..

Les événements modélisés (dans les chapitres II.2. et II.3.) intègrent l'ensemble des périls couverts par le programme de réassurance. Il convient donc d'appliquer des clés de ventilations sur les charges estimées, afin d'avoir le coût du risque propre au produit MRH Privatis.

Cette ventilation de charge est traitée dans le point II.4. qui suit.

II.4. Clés de répartition des événements modélisés entre les risques et produits couverts par le programme « Force de la Nature »

Comme il l'a été précédemment dit, le traité « Force de la Nature » prend en compte l'ensemble des périls couverts (le risque MRH, le risque Agricole, Collectivité, Serres, ...). Or, ici, la modélisation de la prime pure porte essentiellement sur la garantie « événements climatiques » du produit MRH Privatis. C'est pourquoi des clés d'allocation ont été créées.

Dans un premier temps, ces clés sont élaborées à partir des sorties brutes du logiciel RMS (et non pas les sorties retraitées après application du programme de réassurance « Force de la Nature »). Cette première clé servira également à répartir le coût de facturation de la réassurance entre les différentes Entités (cette réallocation est possible car le logiciel RMS fournit des tableaux pour tous les périls et Entités).

Dans un second temps et pour plus de stabilité, une seconde clé est élaborée sur la combinaison sinistres/cotisations des différents produits MRH. En effet, chaque Caisse Régionale n'a pas la même proportion de produit Privatis dans son portefeuille car les dates de commercialisation de ce produit ont été différentes pour certaines Entités.

Suite à ces précisions, il en ressort les pondérations suivantes pour les Caisses Régionales :

RMS (DCER/DAG)	Répartition selon les risques	
	Risque MRH	Autres Risques
GAM	63,6%	36,4%
GCA	38,8%	61,2%
GCM	28,7%	71,3%
GGE	33,4%	66,6%
GNE	36,5%	63,5%
GOC	37,2%	62,8%
GRAA	41,4%	58,6%
GPVL	33,3%	66,7%
GSUD	45,1%	54,9%
Total	36,0%	64,0%

Tableau 2. 33. Répartition des charges sinistres espérées selon le péril

Bases Métiers (DADIP)	Répartition sur les produits MRH	
	Privatis	Autres produits
GAM	65,7%	34,3%
GCA	80,2%	19,8%
GCM	65,1%	34,9%
GGE	56,5%	43,5%
GNE	41,1%	58,9%
GOC	41,7%	58,3%
GRAA	80,6%	19,4%
GPVL	53,3%	46,7%
GSUD	67,4%	32,6%
Total	64,8%	35,2%

Tableau 2. 34. Répartition des charges sinistres espérées selon le produit MRH

Constats :

- Du *Tableau 2.33.*, il est possible de remarquer que le risque MRH représente 36% des charges sinistres brutes sur l'ensemble des périls modélisés. Sur ce même tableau, il est possible de remarquer que l'Entité GAM n'a pas beaucoup de risque autre que le risque MRH (région plutôt urbanisée). A l'inverse de GCM qui possède un portefeuille plus rural, la répartition des périls est nettement différente.

- Sur le *Tableau 2.34.*, la répartition « produit Privatis » / « produits autres MRH » est égale en moyenne à 65%. Il est possible de remarquer que des Caisses Régionales comme GCA et GRAA ont davantage souscrit ce produit par rapport aux autres Entités.

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Clés de répartition des événements modélisés entre les risques et produits couverts par le programme « Force de la Nature »	- Une première clé de répartition est élaborée à partir des sorties brutes du logiciel RMS - Une seconde clé est créée pour répartir les charges sinistres sur le produit Privatis uniquement	- En moyenne, 36% des charges sinistres sont affectés au risque MRH (les fluctuations sont importantes au niveau des Entités) - Les charges sinistres à allouer au produit Privatis sont en moyenne à près de 65%. L'affectation des charges est là aussi fluctuante selon les Entités.	Interaction des différents éléments dans la détermination de la prime pure et conséquence pour les Caisses Régionales sur le produit Privatis

A présent que les clés de répartition sont définies, il convient de faire la jonction de tous les travaux réalisés.

II.5. Interaction des différents éléments dans la détermination de la prime pure et conséquence pour les Caisses Régionales sur le produit Privatis

Pour conclure les travaux liés à la Section II., le chapitre suivant va s'articuler en deux points. Dans un premier temps, il sera expliqué de quelle façon le niveau de la prime pure va être calibré. Puis, il sera présenté une méthodologie d'aide à la décision pour que les Entités puissent se positionner face à ce nouveau calibrage.

II.5.1. Détermination du niveau de la prime pure

Ce chapitre permet de faire la jonction de toutes les étapes précédentes.

- Dans le chapitre II.1., il a été présenté la méthodologie de détermination des échelles de primes pures selon l'algorithme tarifaire de l'actuel produit ;
- Les modélisations effectuées sur le risque « tempêtes » dans le modèle interne, ont été présentées dans le chapitre II.2.. Celles-ci ont été réalisées sur l'ensemble des risques couverts par le traité de réassurance « Force de la Nature » ;
- Dans le chapitre II.3., le programme de réassurance a été appliqué sur les événements extrêmes modélisés du chapitre précédent ;
- Afin de ne cibler que le produit Privatis, des clés de répartition ont été appliquées et exposées dans le chapitre II.4..

A présent, pour déterminer le niveau de la prime pure de chaque Caisse Régionale, les étapes suivantes sont réalisées :

1^{ière} étape :

A partir des travaux de la Section II.1., une prime pure « unitaire » a pu être affectée à chaque ligne de contrat recensé dans la base 2010. De façon plus explicite, une multiplication des coefficients de chaque modalité de chaque critère tarifaire a été réalisée.

Les valeurs ont bien évidemment été proratisées à la durée d'exposition de chaque contrat.

N.B. : Le coefficient de multiplication déterminé par le logiciel Pretium a été forcé à 1. De même, les modalités des deux facteurs « Caisse Régionales » et « Exercice » ont également été forcées à 1.

Les travaux de modélisation dans le module « CAT » du modèle interne sont faits pour capter ces deux variables.

2^{ème} étape :

Les valeurs ainsi obtenues sont sommées par Entité⁴¹. Ces 10 sommes vont ensuite diviser les charges sinistres de la Section II.3. (charges sinistres tenant compte de tous les effets liés à la réassurance) auxquelles ont été appliquées les clés de répartition.

3^{ème} étape :

Dès lors, un coefficient multiplicatif est défini. Ce coefficient permet de déterminer le niveau de la prime pure finale que doit payer chaque sociétaire et ce, pour chaque Entité.

Pour faire suite à ce qui a été fait précédemment, seules les Caisses Régionales GCA et GSUD sont présentées, ainsi que le périmètre de l'ensemble des CR hors GLBR.

Entités	Espérance de sinistres en million d'euros				Sommes des échelles de prime pure proratisées	Déduction du coefficient multiplicateur (niveau de la prime pure en euro)	
	Avant application des clés de répartition		Après application des clés de répartition			Sinistres Bruts	Sinistres Nets
	Sinistres Bruts	Sinistres Nets	Sinistres Bruts	Sinistres Nets			
GCA	16,2	12,5	6,28	4,84	293 855	17,15	13,27
GSUD	7	7	3,16	3,16	164 777	12,99	12,97
Ensemble des CR (hors GLBR)	126,8	106,4	45,6	38,3	1 520 752	19,45	16,32

Tableau 2. 35. Mécanisme de calcul des coefficients multiplicateurs pour chaque Entité

Ce paragraphe clôt ainsi l'ensemble des travaux liés à la détermination de la prime pure de la garantie « événements climatiques ».

⁴¹ Voir Annexe : Programme SAS [11.1.]

Rappel objectif	Méthodologies employées	Étape suivante
Détermination du niveau de la prime pure	- Une multiplication des coefficients de chaque modalité de chaque critère tarifaire a été calculée afin d'obtenir une prime pure « unitaire » par contrat. - Les valeurs obtenues sont sommées par Entité afin de pouvoir diviser les charges sinistres espérées du chapitre II.3. après application des clés de répartition définies au chapitre II.4.. - Le coefficient multiplicatif obtenu permet de déterminer le niveau de la prime pure finale et ce, pour chaque Entité.	Aide à la décision pour les différentes Entités face à ce nouveau calibrage

A présent, les différentes Entités doivent se positionner face à ce nouveau calibrage de prime pure.

II.5.2. Aide à la décision pour les différentes Entités face à ce nouveau calibrage

Afin d'aider les Entités dans leur choix stratégique, deux méthodes leur sont proposées :

- Sur le principe de « *Back-Testing* », la première consistera à comparer le nouveau calibrage de la prime pure avec les charges sinistres réelles observées en 2009 et 2010 ;
- La deuxième méthode présentera le rapport entre le nouveau calibrage de la prime pure et les paramètres actuels implémentés.

Les résultats décrits par la suite sont transcrits⁴² et présentés sous la forme de tableau, croisement de deux des trois principaux critères tarifaires actuels SIGMA :

- Nombre de Pièces,
- Type de bâtiment.

Pour une question de lisibilité, la « limite de capital mobilier » garantie a été ôtée.

N.B.1 : Les deux options de capital mobilier : « 02 – sans capital mobilier » et « 03 : de 11 à 12,5 FFB » sont sous-estimées.

N.B.2 : Une comparaison du nouveau calibrage de la prime pure chargée, selon les normes Groupe, et les cotisations actuelles, aurait pu être intéressante, mais les taux de chargements et frais de gestions sont différents selon les Caisses Régionales. En effet, ces taux diffèrent de par :

- le coût de fonctionnement de chaque Entité (effectif, frais de gestion, partenariat avec divers organismes dans le traitement et la réparation des biens sinistrés,...),
- la politique commerciale de ces dernières.

Il peut être également intéressant de préciser que ces chargements peuvent être différents selon les profils de risque. Bien souvent, le « taux de chargement » se compose de frais fixes (plus précisément d'un montant fixe) et de frais variables.

(a) Comparaison entre le nouveau calibrage de la prime pure et les primes pures brutes observées en 2009 et 2010 (« *Back-Testing* »)

Comme il l'a été mentionné, il est présenté ci-dessous une comparaison entre le nouveau calibrage de la prime pure et les primes pures brutes observées en 2009 et 2010 :

Nombre de pièces	Collectif	Individuel	Total	poids
1	-76%	-7%	-35%	7.4%
2	-24%	-7%	-12%	11.9%
3	-52%	8%	0%	18.0%
4	-15%	4%	3%	20.7%
5	23%	12%	12%	17.7%
6	205%	24%	26%	11.5%
7	118%	54%	55%	6.3%
8	340%	48%	52%	3.4%
9	194%	68%	70%	1.6%
10	48%	75%	75%	0.8%
11	190%	54%	57%	0.3%
12	-57%	138%	133%	0.3%
Total	0%	26%	25%	
poids	25,5%	74,5%		

Tableau 2. 36. Comparaison entre le nouveau calibrage et la prime pure réelle 2009

Nombre de pièces	Collectif	Individuel	Total	poids
1	-58%	118%	48%	7,0%
2	-49%	65%	34%	12,0%
3	-50%	28%	18%	18,4%
4	-18%	18%	17%	21,1%
5	2%	10%	9%	17,7%
6	99%	9%	10%	11,4%
7	43%	12%	13%	6,2%
8	57%	20%	20%	3,3%
9	-33%	31%	30%	1,6%
10	21%	27%	27%	0,8%
11	59%	-10%	-8%	0,3%
12	-48%	5%	4%	0,3%
Total	-25%	17%	15%	
poids	25,7%	74,3%		

Tableau 2. 37. Comparaison entre le nouveau calibrage et la prime pure réelle 2010

⁴² Voir Programme SAS [11.2.] en Annexes

Interprétations des tableaux :

- Les valeurs en rouge indiquent que les primes pures observées sont supérieures au nouveau calibrage de la prime pure (les valeurs sont supérieures à 0%, c'est le cas d'une « sous-estimation »).
- Les valeurs en noir indiquent que les primes pures observées sont en phase avec le profil de risque à -5% ou +5%.
- Les valeurs en vert indiquent que les primes pures observées sont inférieures au nouveau calibrage de la prime pure (les valeurs sont inférieures à 0%, c'est le cas d'une « surestimation »).

N.B. : Avant de commenter ces résultats, il semble important d'apporter un regard sur la lecture des valeurs. Par exemple, sur le *Tableau 2.36.*, avec une valeur de 0% sur les habitations collectives et une population représentant près de 26%, le résultat global sur la garantie devrait mathématiquement avoisiner les 19,3% : $0\% \times 25,7\% + 26\% \times 74,3\% = 19,3\%$.

Mais la pondération du portefeuille et le niveau distinct des primes pures (ancien paramétrage et nouveau calibrage) engendrent un constat différent.

Pour ce *Tableau 2.36.*, les appartements ont une prime pure observée et une prime pure calibrée selon la nouvelle méthodologie de 3,8€. Dans le même temps, les habitations individuelles ont une prime pure observée de 36,3€, pour une prime pure selon le nouveau calibrage de 28,8€.

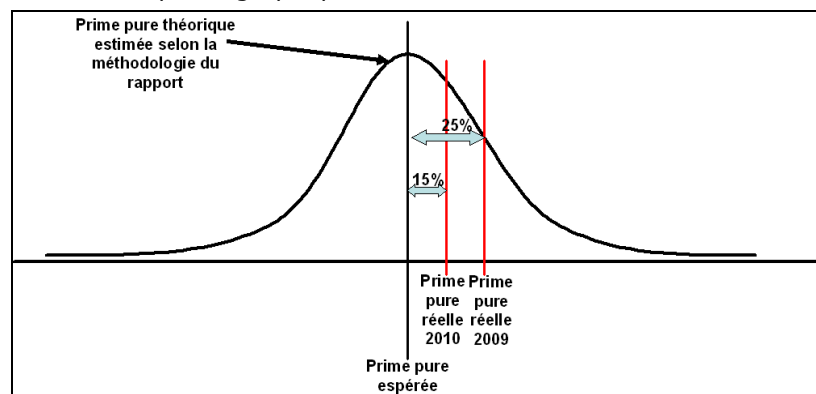
$\frac{3,8 \times 25,5\% + 36,3 \times 74,5\%}{3,8 \times 25,5\% + 28,8 \times 74,5\%} - 1 = 25\%$. Ce principe de pondération explique les résultats observés dans

les colonnes intitulées « Total » des deux tableaux.

Constat lié aux tableaux 2.36. et 2.37. :

- Ce « Back Testing » démontre que le nouveau calibrage de la garantie « événements climatiques » réalisé dans ce rapport est inférieur aux primes pures brutes réelles 2009 et 2010. C'est un point plutôt rassurant. En effet, les exercices 2009 et 2010 ont été marqués par deux événements identifiables. La tempête Klaus et la tempête Xynthia, dont les périodes de retour sont estimées respectivement à une fois tous les 10 ans et une fois tous les 5 ans en termes de charge (Klaus et Xynthia étaient vraisemblablement de même intensité, mais le régime « Cat. Nat. » a été utilisé pour cette dernière, cf. *Graphique 1.1.*).

Ce constat peut être traduit par le graphique suivant :



Graphique 2. 17. Interprétation des événements climatiques 2009 et 2010 par rapport au nouveau calibrage de la prime pure

Il est proposé en *Annexe [12.]*, un « back testing » sur l'exercice 2011. Sur cette annexe, il est possible de remarquer que le nouveau calibrage est 58% supérieur à la sinistralité observée 2011.

(b) Comparaison entre le nouveau calibrage de la prime pure et la prime pure actuellement paramétrée

Comme il l'a été mentionné, chaque Entité a la possibilité d'appliquer ses propres taux de chargements. En effet, ces taux peuvent fluctuer selon le coût de fonctionnement et la politique commerciale qui leur est propre.

Afin de les aider dans leur choix stratégique, la deuxième méthode va présenter le rapport entre le nouveau calibrage de la prime pure et les paramètres actuellement saisis.

Les résultats sont les suivants :

Nombre de pièces	Collectif	Individuel	Total	poids
1	-40%	15%	-16%	7,0%
2	-36%	21%	-3%	12,0%
3	-30%	29%	16%	18,4%
4	-35%	33%	27%	21,1%
5	-27%	41%	39%	17,7%
6	-14%	44%	42%	11,4%
7	-9%	42%	41%	6,2%
8	-1%	37%	36%	3,3%
9	-2%	44%	43%	1,6%
10	-8%	26%	26%	0,8%
11	-1%	25%	24%	0,3%
12	9%	41%	40%	0,3%
Total	-31%	38%	32%	
poids	25,7%	74,3%		

Tableau 2. 38. Comparaison entre le nouveau calibrage de la prime pure et l'ancien paramétrage

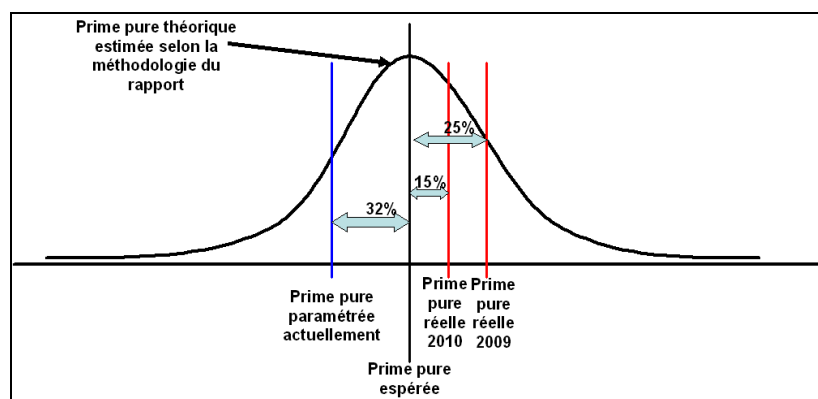
Interprétations des tableaux :

- Les valeurs en rouge indiquent que les primes pures paramétrées actuellement sont inférieures à ce qui a été établie avec le nouveau calibrage (les valeurs sont supérieures à 0%, c'est le cas d'une « sous-estimation »).
- Les valeurs en noir indiquent que les primes pures paramétrées sont en phase avec les profils de risque modélisés dans ce rapport à -5% ou +5%.
- Les valeurs en vert indiquent que les primes pures actuellement paramétrées sont supérieures à ce qui a été établie avec le nouveau calibrage (les valeurs sont inférieures à 0%, c'est le cas d'une « surestimation »).

Constats :

- Tout d'abord, au titre de la garantie « Evènements Climatiques », il est possible de voir qu'au niveau National, la prime pure paramétrée a été sous-estimée de 32%.
- Les primes pures paramétrées pour les habitations collectives ont été surestimées sur la garantie « évènements climatiques ». A l'inverse, les habitations individuelles ont été sous-estimées. Ainsi, la proportion des habitations individuelles dans le portefeuille tire la sous-estimation globale de la garantie.

De la même façon que précédemment, les propos ci-dessus sont de nouveau traduits par la représentation graphique suivante :



Graphie 2. 18. Interprétation du nouveau calibrage de la prime pure, par rapport à l'ancien paramétrage

Rappel objectif	Méthodologies employées	Points retenus
Aide à la décision pour les différentes Entités face à ce nouveau calibrage	- Comparaison entre le nouveau calibrage de la prime pure et les primes pures brutes observées en 2009 et 2010 (<i>Back-Testing</i>). - Comparaison entre le nouveau calibrage de la prime pure et la prime pure actuellement paramétrée.	- De par la nouvelle méthodologie, une sous-estimation globale de la garantie « évènements climatiques » est à constater. - Les résultats sont essentiellement tirés par les habitations individuelles. - Les résultats liés aux habitations collectives montrent une surestimation

Point de transition :

Face à l'impact des récents événements naturels dans les résultats du Groupe Groupama et l'arrivée de la nouvelle réglementation Solvabilité 2, une réflexion est actuellement menée sur les produits Multi Risques Habitation commercialisés à ce jour, notamment dans l'approche tarifaire existante de la garantie « événements climatiques ».

L'enjeu de cette section était d'analyser et de calibrer la prime pure de cette garantie.

L'approche actuarielle classique qui consiste à utiliser des données historiques pour évaluer les primes pures espérées futures (actualisées d'une projection à l'ultime et d'un indice d'inflation) ne semble pas appropriée.

En effet, les incertitudes liées au dérèglement climatique de la planète et les conséquences sur la fréquence et l'ampleur des événements futurs rendent délicate l'approche tarifaire actuarielle classique. Même avec une connaissance approfondie du passé, le dérèglement climatique ne laisse pas supposer un comportement identique dans les années à venir.

La mise en place du modèle interne a contribué à appréhender le risque différemment. L'élaboration du tarif n'est plus essentiellement basée sur les statistiques passées. Elle s'appuie parallèlement sur une modélisation d'événements atypiques extrêmes (dont la survenance est d'une fois tous les 200 ans comme l'exige la nouvelle réglementation Solvabilité 2) ce qui constitue un impact non négligeable dans le niveau et l'approche tarifaire connus jusqu'à présent.

Les résultats liés à cette nouvelle approche ont permis de montrer que la prime pure de la garantie actuellement paramétrée semblerait être insuffisante. Pour être en phase avec les risques portés, des revalorisations seraient donc nécessaires, notamment sur les habitations individuelles.

Néanmoins, avant d'effectuer toutes opérations, il est indispensable d'apprécier la justesse de la méthodologie. Pour cela, il convient de regarder exactement quelles sont les sensibilités liées à l'approche tarifaire proposée.

Ainsi, dans la Section III., il sera exposé l'ensemble des sensibilités liées à la méthodologie proposée. Des jeux d'hypothèses seront également mis en avant, afin de prendre en compte l'impact de diverses « implémentations externes » dans l'exigence tarifaire.

Face au projet de loi, déposé au Sénat, réformant le régime d'indemnisation des catastrophes naturelles, il sera présenté dans la Section IV., une réflexion sur ces risques et les différentes perspectives qui pourraient y être liées.

Section III. Sensibilités dans l'approche des modèles proposés et jeux d'hypothèses

Face à l'approche proposée dans la Section II., un regard critique doit être apporté. En particulier, il peut sembler naturel de s'interroger sur les sensibilités liées à la méthodologie qui a été décrite. Pour répondre à cet objectif, cette section se découpe en 4 chapitres :

- Dans le chapitre III.1., il est présenté les diverses sensibilités de construction liées à la détermination de la prime pure ;
- Il sera exposé dans le chapitre III.2., des sensibilités liées aux calculs et à la modélisation des sous-modules non-vie, avec notamment une comparaison entre les calculs du modèle standard et interne ;
- Le coût du risque étant sensible au programme de réassurance, il sera présenté dans le chapitre III.3., une méthodologie qui permet d'identifier les effets de possibles changements dans le programme de réassurance (« Force de la Nature ») ;
- Enfin, pour clore cette section, le chapitre III.4. proposera deux modèles économiques pour un produit d'assurance non-vie. Ces modèles traduiront les sensibilités liées à l'allocation du capital et l'exigence de rentabilité qui peut en découler.

III.1. Sensibilités liées à la méthodologie employée dans l'analyse de la prime pure

Afin de répondre à cette problématique, il sera particulièrement porté un regard sur :

- l'historique de la base de données étudiée ;
- le seuil d'exclusion entre sinistres attritionnels et sinistres non-récurrents ;
- le modèle de régression Gamma ;
- l'articulation entre les différentes Directions au sein de Groupama SA ;
- la valorisation des SMP des risques couverts par le traité « Force de la Nature » ;
- la modélisation des charges espérées par le modèle RMS.

Historique de la base de données

La détermination des échelles de fréquences et coûts moyens a été réalisée sur deux exercices à forte sinistralité. Cependant, les modélisations ont montré quelques sensibilités, notamment de la volatilité dans les échelles des coûts moyens. Ceci peut être constaté sur l'intervalle de confiance des habitations de 10 pièces et plus. Pour les deux exercices agrégés, l'exposition rencontrée sur ces modalités était inférieure à 739 sinistres.

L'intégration d'un troisième exercice pourrait donc être envisagée afin de réduire cette volatilité.

Seuil d'exclusion entre sinistres attritionnels et sinistres non-récurrents

À ce jour, il n'existe pas de réelle démarche pour définir un seuil entre les sinistres attritionnels et les sinistres à montant moins courant. Néanmoins, cette analyse est importante, voire fondamentale. Une acceptation de sinistres à montant trop élevé pourrait altérer le modèle, mais un rejet important de sinistres, dû à un seuil trop bas, réduit la discrimination du risque.

Ce seuil à 5700€ peut sembler peut-être trop faible puisqu'un « tassement » sur l'échelle des coûts moyens pour les habitations de 10 et 11 pièces a été observé (cependant et comme il l'a été évoqué, ce n'est peut être pas le seuil qui doit être remis en cause, mais le manque de sinistres dans la base étudiée sur ces habitations de 10 pièces et plus).

Modèle de régression Gamma

Les lois choisies pour la détermination des échelles tarifaires auraient pu être analogues aux lois du modèle interne. Dans la modélisation des échelles de risques des coûts moyens, c'est la loi Gamma qui a été privilégiée. Dans le modèle interne, c'est la loi Bêta.

L'usage du modèle Bêta suppose que la valorisation des biens assurés soit juste. Sans remettre en cause le travail réalisé par la DCER, l'intégration d'un paramètre supplémentaire rend l'approche plus sensible.

Par conséquent, se baser sur des montants déterminés et « réels » est une approche qui reste adéquate, même si une critique est à apporter sur le modèle Gamma. En effet, la loi Gamma est sensible aux valeurs « extrêmes ». La nécessité de créer ce seuil d'exclusion entre les sinistres attritionnels et non-récurrents était une démarche déterminante (voir *Annexe [8.3.]*).

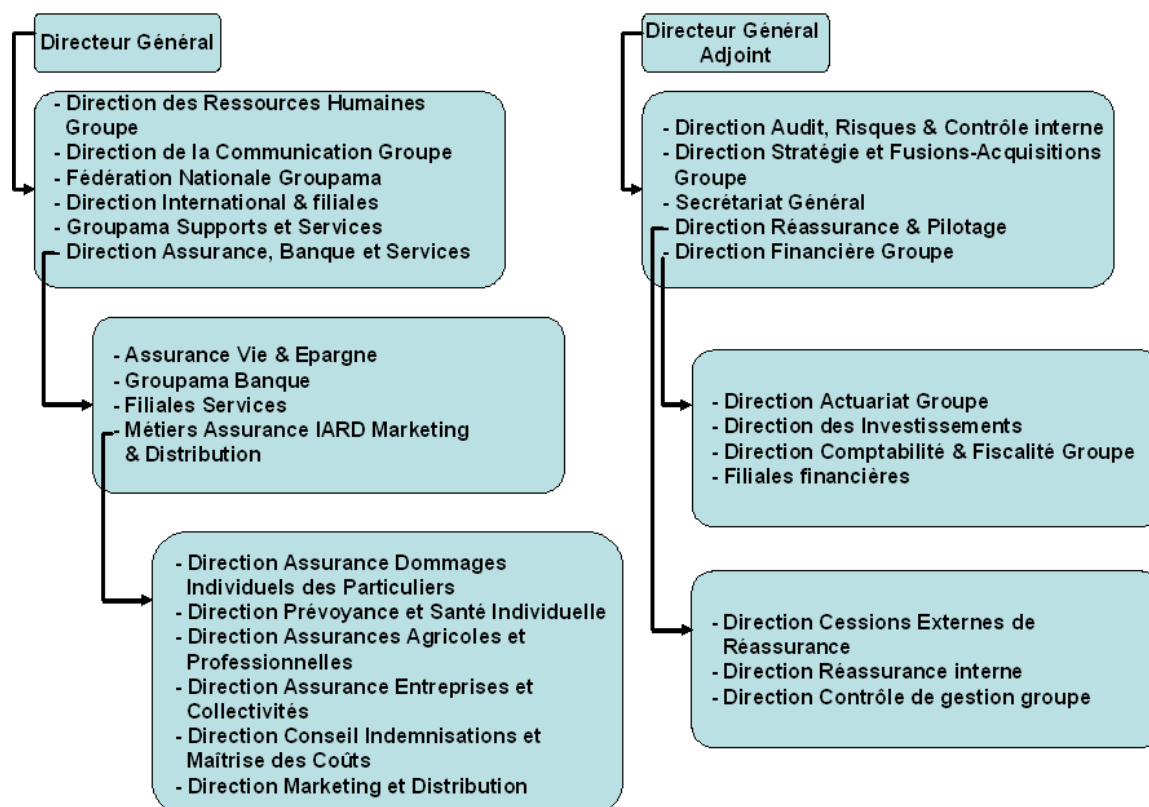
Enfin, dans la base de données, il apparaît pour chaque ligne (i.e. à chaque période d'exposition des polices d'assurance) le nombre total de sinistres et la charge totale observée correspondante. Face à ce type de données, il semblerait que les modèles Tweedie¹ apparaissent comme des modèles théoriques plus adéquats à ce type de données. En effet, ce type de modèle tient compte

¹ Une description plus précise sur son intérêt et ses faiblesses est énoncée à l'*Annexe [13.]*

du grand pourcentage de zéros dans la variable des coûts moyens. Ils donneraient donc une meilleure approche. Mettre en concurrence ce modèle (qui modélise finalement directement la prime pure) avec l'approche qui combine les modèles Poisson et Gamma pourrait être intéressant. Toutefois, malgré l'attrait que pourrait présenter ce type de modèle, il ne faut pas omettre, ici aussi, le côté interprétation. Comme il l'a déjà été noté, la compréhension des méthodes utilisées doit être assimilée par tous les acteurs de l'entreprise (notamment les instances dirigeantes).

Articulation entre les différentes Directions au sein de Groupama SA

L'articulation entre les différentes Directions au sein de Groupama SA peut sembler délicate. C'est pourquoi, il est proposé ci-dessous une architecture simplifiée de l'organigramme de l'entreprise. Les objectifs et responsabilités de chacune des Directions citées dans ce rapport seront exposés dans le *Tableau 3.1.*



Graphe 3. 1. Organigramme simplifié de Groupama SA

L'organigramme et donc le système de gouvernance décrit ici sont le résultat de la nouvelle organisation établie en octobre 2011. Cette architecture sous forme de « scission » permet d'éviter le principe de « juge et partie ». Le tableau proposé ci-dessous explique ces propos.

Directions	Responsabilités	Pour ce rapport	
		Données	Objectifs
DADIP (Direction Assurance Dommages Individuels des Particuliers)	- Gérer les risques et suivre les résultats des différentes Entités sur les produits Automobile, MRH, ... - Responsable du risque « Catastrophes Naturelles »	Tous les contrats présents proratisés à leur temps de présence entre le 1 ^{er} janvier et le 31 décembre sur plusieurs exercices	Calibrer la prime pure de la garantie « événements climatiques »
DCER (Direction Cessions Externes de Réassurance)	- Calibrer le niveau de réassurance externe sur différents périls - Responsable du risque majeur Tempête	Tous les contrats présents au 31 décembre d'un exercice <i>N</i>	Valoriser les sommes assurées de chaque contrat pour l'ensemble des risques couverts par le traité
DAG (Direction Actuariat Groupe)	- Responsable du modèle interne - Elaboration des rapports actuariels - Evaluation des provisions (conception et mise en œuvre des outils de modélisation du passif)	Tous les contrats présents au 31 décembre d'un exercice <i>N</i> valorisés selon les calculs de la DCER	Calculer les besoins en fonds propres de chaque péril, dont la MRH, après application des différents programmes de réassurance interne et externe.

Tableau 3.1. Lien entre les Directions dans l'approche réalisée

Valorisation des SMP des risques couverts par le traité « Force de la Nature »

Les valorisations réalisées par la DCER pourraient être « challengées » en s'appuyant sur la modélisation des coûts moyens. En effet, certaines Entités ont noué des partenariats rendant les coûts de réparations des sinistres moindres au regard d'une situation classique. Ces maîtrises des coûts ne sont pas prises en compte actuellement dans la valorisation des sommes assurées. L'approche de la DCER est donc plutôt prudente et a par conséquent un léger impact dans le niveau de la prime pure (cette prudence a donc un léger surcoût).

Modélisation des charges espérées par le modèle RMS

Dans la modélisation des charges sinistres espérées, des limites peuvent être citées quant à l'utilisation du logiciel RMS :

- La précision géographique pourrait être améliorée. En effet, le niveau de maille actuel est au code postal (or, il existe 6 272 codes postaux, pour 36 568 communes). Une localisation au code commune serait, *a minima*, plus pertinente (d'autant plus que les données sont disponibles. Un niveau plus fin existe également dans le système, il s'agit du niveau IRIS²).
- Dans la construction du modèle d'exposition, il existe peut être un manque d'exhaustivité dans les données (la présence d'arbre à proximité du risque pourrait par exemple être ajoutée, ou bien encore la possibilité de renseigner si une commune possède un « Plan de Prévention des Risques », dit « PPR »).

Rappel objectif	Points retenus	Étape suivante
Sensibilités liées à la méthodologie employée dans l'analyse de la prime pure	- Une année d'historique supplémentaire pourrait permettre de réduire la volatilité observée sur certaines modalités. - Le seuil d'exclusion entre sinistres attritionnels et sinistres non-récurrents pourrait être revu avec l'intégration d'un exercice supplémentaire. Néanmoins, ce seuil semble être pertinent aux vues d'études antérieures. - La combinaison des modèles de régression Poisson/Gamma pourrait être mise en concurrence avec le modèle Tweedie. - L'articulation entre les Directions n'est pas aisée, mais conforme à une politique de non ingérence. - La valorisation des sommes assurées réalisée par la DCER pourrait être revue légèrement à la baisse. - Le modèle d'exposition de RMS pourrait être encore plus fin, notamment dans la localisation des risques.	Sensibilité du SCR MRH

Face à ces divers éléments exposés, la méthodologie proposée peut être validée. Seule la détermination du niveau du coût du risque est « discutable », au sens où les SMP actuels sont peut-être surestimés. Le calibrage de la prime pure relève donc un caractère prudentiel. A noter, de par la méthodologie proposée, que ce niveau pourra être rapidement actualisé à un degré plus adapté (une fois les valorisations mises à jour et effectuées sur l'exercice 2011).

Pour compléter l'approche, les sensibilités liées aux calculs et aux modélisations des sous-modules non-vie sont présentées dans le chapitre suivant.

² La définition est donnée dans la Section IV.

III.2. Sensibilités du SCR MRH

Pour exercer une activité assurantielle en France, les compagnies « doivent respecter une marge de solvabilité calculée selon des modalités définies par décret en Conseil d'Etat » (Article L.334-1).

Afin d'étayer ces propos, il sera présenté successivement :

- Le contexte réglementaire dans lequel évolue le monde assurantiel actuellement (et le calcul de marge lié), ainsi que les perspectives engendrées par la nouvelle Directive Solvabilité 2 (dont l'application est prévue en 2014) ;
- Dans le chapitre III.3.2., les modalités de calcul de la formule standard et interne sous Solvabilité 2 ;
- Dans le chapitre III.3.3., les limites et sensibilités du modèle interne.

III.2.1. Contexte réglementaire actuel et perspectives

Aujourd'hui, chaque pays de l'Union Européenne a la possibilité d'élaborer un cadre de solvabilité propre à sa politique prudentielle. En France, ce calcul de marge prudentielle est défini par l'Article R.334-5 du Code des assurances (qu'il s'agisse d'une compagnie vie ou non-vie).

(a) Contexte prudentiel actuel en France

En assurance non-vie, la marge de solvabilité correspond (en simplifiant) à des fonds propres, dont le montant est défini par la valeur la plus élevée des deux valeurs suivantes (l'une basée sur les cotisations ou primes perçues, l'autre sur les charges sinistres déterminées en moyenne au cours des trois derniers exercices) :

- i) $(16\% * \text{primes brutes} * S)$ où
 $S = (\text{sinistres bruts} - \text{sinistres cédés à la réassurance}) / \text{sinistres bruts}$
 S ne peut être inférieur à 50%.
- ii) $(23\% * \text{sinistres bruts} * S)$ où
 $S = (\text{sinistres bruts} - \text{sinistres cédés à la réassurance}) / \text{sinistres bruts}$
 S ne peut être inférieur à 50%.

Dans cette formule de marge de solvabilité, 4 points de faiblesses peuvent être repérés :

- Les risques spécifiques et propres à chaque compagnie d'assurance n'ont pas d'impact sur la marge de solvabilité ;
- Aucun effet de diversification des risques n'est pris en compte ;
- Il existe une incohérence avec les normes *IFRS*³ en vigueur car les capitaux (Fonds propres) sont actuellement évalués au coût historique ;
- Une grande disparité réglementaire existe au sein de l'Union Européenne.

N.B. : à noter qu'il existe en France un autre caractère prudentiel : la provision pour égalisation. Il est renseigné en *Annexe [12.]* quelques précisions sur cette provision très utilisée pour des risques telle que la Tempête.

La mise en place de la future Directive Solvabilité 2 devrait permettre de combler ces faiblesses.

(b) Contexte prudentiel en 2014 avec Solvabilité 2

La Commission Européenne et l'*EIOPA* (« *European Insurance and Occupational Pensions Authority* »⁴) ont élaboré et préconisé la mise en place d'un nouveau référentiel destiné, avant tout, à protéger les assurés d'un éventuel risque de faillite de leurs compagnies d'assurance. Sur les risques couverts par ces dernières, un niveau de Capital (Fonds propres) adéquat doit être constitué afin d'assurer la solvabilité de l'institution financière.

Là, où les organismes d'assurance devaient utiliser des méthodes basées sur des pourcentages de primes ou sinistres, une approche plus pertinente est élaborée, avec pour objectifs principaux :

- de réduire les risques de faillite à 0,5% par an,
- de mettre en adéquation les exigences de solvabilité avec les risques portés,
- d'harmoniser les pratiques au niveau européen.

³ *IFRS* : « *International Financial Reporting Standards* », Normes internationales d'information financière (vision des différents postes du bilan en valeur de marché).

⁴ Anciennement appelé *CEIOPS* : « *Committee of European Insurance and Occupational Pensions Supervisors* »

Solvabilité 2 place la gestion des risques au cœur des préoccupations des organismes. Les entreprises sont ainsi encouragées à mettre en place une gestion saine dans leurs organisations, notamment par la mise en place :

- d'un système de gouvernance,
- d'un système de mesure des risques (« ORSA »),
- d'une plus grande transparence et fiabilité des informations communiquées (à la fois aux autorités de tutelles et au marché).

Le principe de contrôles est donc l'un des concepts clés de Solvabilité 2.

Un autre concept clé de Solvabilité 2 (S2) est l'évaluation des différents postes du bilan à leurs valeurs de marché :

- Concernant l'actif, l'évaluation résulte d'une lecture directe des prix du marché. Si cela n'est pas possible, l'estimation se base sur des modèles de valorisations.
- Pour le passif, notamment pour l'évaluation des provisions techniques, l'estimation est réalisée selon la méthode du « *best estimate* », à laquelle doit être ajoutée une « marge de risque » :

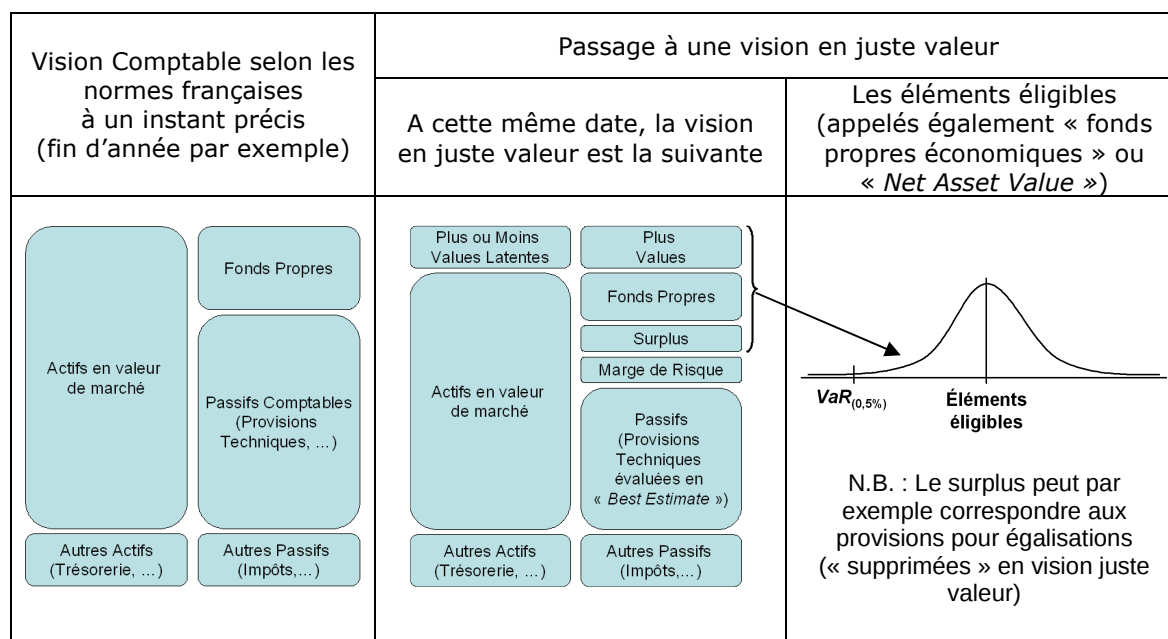
- Le « *best estimate* » peut être traduit comme étant la valeur actuelle probable des flux de trésoreries futurs, avec une prise en compte de la valeur temporelle de l'argent (estimée sur une base de courbe de taux sans risque).

Dans les spécifications techniques (QIS5), aucune méthode n'est explicitement nommée quant à la méthode de détermination de ces provisions techniques, mais de nombreuses formules de calculs sont proposées. Ces dernières prennent, de plus, en compte, un certain nombre de paramètres. La réassurance en est, par exemple, l'un d'eux.

Les provisions techniques doivent également intégrer une « marge de risque ». Cette marge pour risque permet de définir une marge de sécurité sur l'estimation des engagements pris par l'organisme d'assurance. Ainsi, en cas de défaillance de l'assureur, les provisions techniques pourrait être transférées vers un autre acteur du marché sans que celui-ci n'ait besoin de se financer⁵.

Dans le QIS5, la marge pour risque est associée à la notion de « *Cost-of-Capital rate* » (littéralement en français, coût du capital), dont le taux actuellement retenu est de 6%. Cette marge pour risque doit également être pondérée selon la contribution du capital requis pour chaque métier au sens Solvabilité 2 (voir *Tableau 3.3.*).

Le passage entre la vision comptable française actuelle et la vision en juste valeur (IFRS ou sous Solvabilité 2) est proposé ci-dessous (les formes géométriques n'engagent en rien un quelconque poids des rubriques).



Graph 3. 2. Passage entre la vision comptable française et la vision du bilan en norme IFRS (juste valeur)

⁵ THERON Pierre (2008) *IFRS, Solvabilité 2, Embedded Value : Quel traitement du risque ?*, Bulletin Français d'Actuariat, Vol. 8, n°15, janvier-juin 2008, pp. 67-96

L'architecture de ce nouveau référentiel S2 a été créée sous forme de 3 piliers : exigence quantitative, exigence qualitative et publication d'information. Seul le premier pilier (et l'exigence de capital liée) est présenté ci-après.

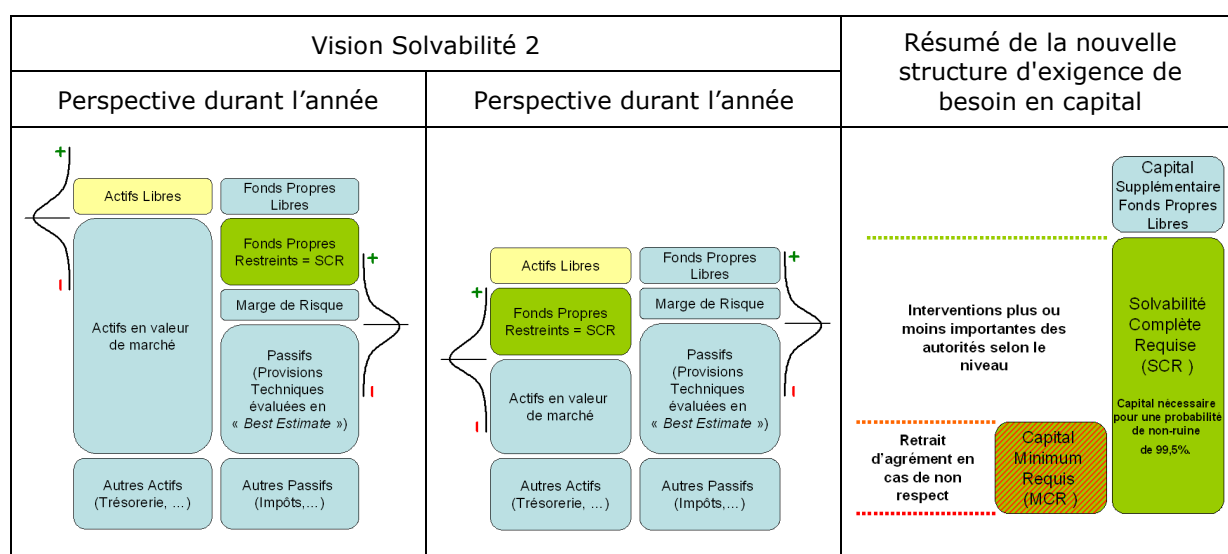
(c) Pilier I : Exigences quantitatives

L'exigence quantitative et le concept de besoin en capital réglementaire (appelé également « besoin en fonds propres ») sont le premier pilier clé introduit par Solvabilité 2.

Ce besoin en capital réglementaire (ou marge de solvabilité, selon les textes) est défini par la création de deux niveaux de capitaux distincts :

- le **MCR** (« *Minimum Capital Requirement* ») : niveau minimal de capital en deçà duquel une compagnie d'assurance ne peut fonctionner (les autorités, en l'occurrence l'ACP⁶ intervient) ;
- le **SCR** (« *Solvency Capital Requirement* ») : capital requis pour faire face à une probabilité de ruine de 0,5% (ou, selon la lecture souhaitée, une probabilité de non faillite de 99,5% à horizon un an).

Pour étayer l'ensemble de ces propos, une représentation graphique est présentée ci-dessous :



Graph 3. 3. Représentation graphique de l'exigence de besoin en capital sous Solvabilité 2

Ces deux niveaux de capitaux sont déterminés, soit à l'aide d'une formule standard, soit par l'intermédiaire d'un modèle interne (intégral ou partiel). Groupama est dans cette dernière situation. Les structures de calculs sont données dans les spécifications techniques du QIS5.

Le calcul du **MCR** est spécifié tel que $MCR = \max\{MCR_{Combined}; AMCR\}$ avec **AMCR** (« *Absolute Floor of the MCR* »), correspond à un montant de 2,2M€ en assurance non-vie et

$MCR_{Combined} = \min\{\max[MCR_{Linear}; 0,25 \times SCR]; 0,45 \times SCR\}$ et $MCR_{Linear} = MCR_{Life} + MCR_{Non Life}$ où :

MCR_{Life} = Niveau minimal de capital pour le risque vie,

$MCR_{No life}$ = Niveau minimal de capital pour le risque non-vie.

En ce qui concerne le calcul du **SCR**, celui-ci est spécifié tel que $SCR = BSCR + Adj + SCR_{Op}$ avec :

BSCR = Fonds propres requis de base,

Adj = Ajustement lié aux effets des provisions techniques et impôts différés,

SCR_{Op} = Besoin en capitaux pour le risque opérationnel.

La détermination du **BSCR** repose sur une structure modulaire de risques (cf. Graph 2.13.). En l'occurrence, ce sont six modules de risques, composés eux-mêmes de sous-modules, qui permettent d'établir ce capital.

Chaque sous-module subit une approche basée sous forme de « choc » et permet ainsi de leur déterminer un capital requis propre. L'agrégation de ces sous-modules, selon des matrices de

⁶ ACP : Autorité de Contrôle Prudential

corrélations, traduisant la dépendance des risques et leur diversification, aboutit au capital requis final réglementaire.

Dans les spécifications techniques, le *BSCR* est spécifiquement décrit tel que :

$$BSRC = \sqrt{\sum Corr_{ij} \times SCR_i \times SCR_j} + SCR_{intangibles}, \text{ avec}$$

$Corr_{i,j}$ = Facteur de corrélation entre *SCR* défini dans la matrice du *QIS5* (cf. *Tableau 3.1.*),

SCR_i, SCR_j = *SCR* de sous-module en ligne *i* et colonne *j* de la matrice de corrélation,

$SCR_{intangibles}$ = Besoin en capitaux pour risque d'immobilisations incorporelles.

La matrice de corrélation $Corr_{i,j}$ est définie comme suit :

i \ j	Market	Default	Life	Health	Non-life
Market	1				
Default	0.25	1			
Life	0.25	0.25	1		
Health	0.25	0.25	0.25	1	
Non-life	0.25	0.5	0	0	1

Tableau 3. 1. Matrice de corrélation extraite du *QIS5*

Le *SCR* (et par effet le *MCR*) vie et non-vie sont fonction :

- des provisions « *best estimate* »,
- des primes émises nettes sur les douze derniers mois,
- de paramètres calibrés par l'*EIOPA*,
- et de différentes lignes de métiers (appelées dans les spécifications « *LoB : Line of Business* »⁷).

Rappel objectif	Points retenus	Prochaine étape
Contexte réglementaire actuel et perspectives	Aujourd'hui : - Pas d'appréciation et de compréhension des risques propres à chaque compagnie d'assurance ; - Pas de prise en compte des effets de diversifications des risques pour le calcul d'exigence de marge ; - Incohérence avec les normes <i>IFRS</i> en vigueur ; - De grandes disparités réglementaires existent au sein de l'Union Européenne. Demain avec Solvabilité 2 : - Principe de contrôles et vision économique du bilan (évaluée en juste valeur) ; - Suppression des provisions pour égalisation (spécificité française) ; - Réduction des risques de faillite par la matérialisation de 2 niveaux réglementaires (le <i>SCR</i> qui est le capital requis pour faire face à une probabilité de ruine de 0,5% par an et le <i>MCR</i> en deçà duquel un organisme perd son agrément) calculés, soit à l'aide d'une formule standard, soit par l'intermédiaire d'un modèle interne (intégral ou partiel) ; - Mise en adéquation des exigences de solvabilité avec les risques portés ; - Harmonisation des pratiques au niveau européen.	Modalités de calcul du <i>SCR</i> selon le modèle standard et interne

Ce contexte réglementaire étant à présent établi, il est décrit, dans le chapitre ci-dessous, le mode de calcul qui a été élaboré pour les besoins en fonds propres selon les modules et sous-modules de risque inclus dans le modèle interne.

N.B. : Pour déterminer le niveau de capitaux propres que doivent posséder les différents organismes face à leurs expositions aux risques, les régulateurs utilisent généralement la « *VaR* ». La valeur exposée au risque (appelé parfois Valeur-à-Risque, ou selon l'appellation anglo-saxonne « *Value at Risk* » dite « *VaR* ») est une tentative pour résumer en un seul nombre l'exposition au risque d'un portefeuille.

En l'occurrence avec Solvabilité 2, l'objectif d'un calcul de *VaR* est de pouvoir assurer qu'au seuil *s* de 0,5%, la perte sur un portefeuille ne dépassera pas un montant défini par la $VaR_{(0,5\%)}$ dans les 365 prochains jours.

⁷ Un tableau descriptif est donné au chapitre suivant (cf. *Tableau 3.3.*)

III.2.2. Modalités de calcul du SCR selon le modèle standard et interne

Dans le cadre non-vie, le besoin de fonds propres (le SCR) découle des éléments suivants :

- Les volatilités associées au Risque de Réserve (sous estimations des provisions pour sinistres en « *best estimate* » : par exemple, un mali⁸ sur le stock de provisions pour sinistres),
- Les volatilités associées au Risque de Prime y compris Risque de Catastrophe (une insuffisance de primes ou cotisations pour payer les sinistres relatifs à une nouvelle année de production).

Dans les spécifications techniques (QIS5), la formule du SCR non-vie (notée SCR_{nl}) est décrite telle que $SCR_{nl} = \sqrt{\sum CorrNL_{r,c} \cdot NL_r \cdot NL_c}$ avec :

$CorrNL_{r,c}$ = Facteur de corrélation du risque défini dans la matrice ci-dessous (cf. *Tableau 3.2.*),
 $NL_r = NL_{pr}$ = Besoins en capitaux pour le risque de prime et réserve en non-vie,
 $NL_c = NL_{CAT}$ = Besoins en capitaux pour le risque de catastrophe non-vie.

La matrice de corrélation $CorrNL$ est définie comme suit :

$CorrNL$	NL_{pr}	NL_{lapse}	NL_{CAT}
NL_{pr}	1		
NL_{lapse}	0	1	
NL_{CAT}	0.25	0	1

Tableau 3. 2. Matrice de corrélation du risque Non Vie extraite du QIS5

N.B. : NL_{lapse} = Besoin en capitaux pour le risque de disparition non-vie.

Les besoins en capitaux pour le risque de catastrophe (NL_{CAT}) ayant été calculés dans la Section II., il est décrit, ci-dessous, les modalités de calcul concernant les besoins en capitaux pour le risque de prime et réserve non-vie (noté NL_{pr}).

Ces besoins résultent de la fluctuation dans le temps, de la fréquence et de la sévérité des événements assurés. Le cadre réglementaire Solvabilité 2 fixe ces fluctuations par une segmentation au niveau « *Line of Business* » (« *LoB* ») décrite dans le tableau ci-dessous :

Line of Business	Ligne métier	Standard deviation for (Volatilité liée au)	
		Premium risk (Risque de prime)	Reserve risk (Risque de réserve)
Motor vehicle liability	RC automobile	10%	9,5%
Motor, other classes	Dommages automobile	7%	10%
Marine, aviation, transport (MAT)	Marine, aviation, transport	17%	14%
Fire and other property damage	Incendie et autres dommages	10%	11%
Third-party liability	RC générale	15%	11%
Credit and suretyship	Crédit et caution	21,5%	19%
Legal expenses	Protection juridique	6,5%	9%
Assistance	Assistance	5%	11%
Miscellaneous	Divers	13%	15%
Non-proportional reinsurance - property	Réassurance non-proportionnelle - Dommages	17,5%	20%
Non-proportional reinsurance - casualty	Réassurance non-proportionnelle - RC	17%	20%
Non-proportional reinsurance - MAT	Réassurance non-proportionnelle - MAT	16%	20%
Medical expense	Remboursement de frais de soins	4%	10%
Income protection	Perte de revenus	8,5%	14%
Workers compensation	Accident du travail	5,5%	11%
Non-proportional reinsurance - health	Réassurance non-proportionnelle - Santé	17%	20%

Tableau 3. 3. Table des volatilités de primes et de réserves dans le cadre du modèle standard

Dans le modèle standard, il semble important de préciser que les volatilités liées à ces segments de risque Solvabilité 2 sont indépendantes des fluctuations historiques du portefeuille. Dans le modèle interne, ce n'est pas le cas, les volatilités ont été ajustées aux fluctuations historiques.

Soient V et σ :

- V désigne le volume correspondant aux provisions pour sinistres « *best estimate* » nettes de réassurance et aux primes prévisionnelles nettes de réassurance,
- σ désignant la volatilité globale calculée à partir des volatilités spécifiques pour le risque de réserve et pour le risque de réserve par segment Solvabilité 2

⁸ Mali : Estimation insuffisante des coûts de sinistres.

Dans les spécifications techniques, le besoin de capital global (Réserves et Primes) noté NL_{pr} est défini tel que : $NL_{pr} = \rho(\sigma).V$ où $\rho(\sigma) = \frac{\exp\left(N_{0.995} \times \sqrt{\ln(\sigma^2 + 1)}\right)}{\sqrt{\sigma^2 + 1}} - 1$ décrit une distribution log-normale du risque sous-jacent et où $N_{0.995}$ correspond au quantile (également appelé VaR) à 99,5% de la loi normale standard.

De manière simplifiée, $NL_{pr} \cong 3.V\sigma$.

Dans le modèle interne, pour le risque de Réserves, le besoin de capital dépend ainsi des :

- **Volumes de risques** (notés $V_{(res,lob)}$) par « *Line of Business* » Solvabilité 2, correspondant aux provisions pour sinistres en « *best estimate* » nettes de réassurance (meilleure estimation du stock des provisions, dossier à dossier, nettes de réassurance sans aucune marge ou excédent implicite) ;
- **Volatilités spécifiques** (notées $\sigma_{(res,lob)}$) par « *Line of Business* » Solvabilité 2, calculées à partir des données historiques sous forme de triangles de liquidation⁹. Cette volatilité doit capter le risque de Boni¹⁰ / Mali sur le stock de provisions, dossier à dossier, hors marge implicite.

De même, dans le modèle interne, pour le risque de Primes, le besoin de capital dépend des :

- **S/P historiques** (notés $\sigma_{(prem,lob)}$) qui servent à calculer la volatilité observée sur la chronique de S/P par « *Line of Business* » Solvabilité 2 et la projection sur l'année N+1 du S/P Moyen attendu (en vision prudentielle). Le S/P moyen attendu peut prendre en compte des évolutions tarifaires N+1 validées par le management ;
- **Cotisations et frais prévisionnels** (notées $V_{(prem,lob)}$) définis dans le Plan Stratégique Opérationnel (PSO) ;

Ainsi, pour chaque « *LoB* » Solvabilité 2, selon les spécifications techniques, il vient :

$$\sigma_{(lob)} = \frac{\sqrt{(\sigma_{(prem,lob)} V_{(prem,lob)})^2 + 2\alpha\sigma_{(prem,lob)}\sigma_{(res,lob)} V_{(prem,lob)} V_{(res,lob)} + (\sigma_{(res,lob)} V_{(res,lob)})^2}}{V_{(prem,lob)} + V_{(res,lob)}} \text{ avec } \alpha = 0,5^{11} \text{ et}$$

$$V_{(lob)} = (V_{(prem,lob)} + V_{(res,lob)}) \times (0.75 + 0.25.DIV_{lob}) \text{ où } DIV_{lob} = \frac{\sum_j (V_{(prem,j,lob)} + V_{(res,j,lob)})^2}{\left(\sum_j (V_{(prem,j,lob)} + V_{(res,j,lob)})\right)^2} \text{ intègre une}$$

diversité géographique j (limitée aux segments stipulés dans l'Annexe M du QIS5) pour chaque *LoB*.

Rappel objectif	Points retenus	Prochaine étape
Modalités de calcul du SCR selon le modèle standard et interne	- Prise en compte des effets de diversifications : - au niveau géographique ; - au niveau des volatilités spécifiques à chaque métier « <i>LoB</i> » (pour le modèle interne, les volatilités ont été ajustées aux spécificités observées au niveau Groupe) - Le SCR « CAT » a une corrélation de 0,25 avec le SCR de prime et de réserve	Analyse des sensibilités et comparaisons liées aux formules de calcul des modèles standard et interne

III.2.3. Analyse des sensibilités et comparaisons liées aux formules de calcul des modèles standard et interne

Les calculs prudentiels élaborés dans le modèle interne suivent l'approche réglementaire et ont été réalisés à fin 2010 dans les conditions réelles de production d'un calcul à cette date.

Dans les modalités de calcul des SCR, il est apparu plusieurs sensibilités. Celles-ci sont de nature à impacter fortement le besoin de fonds propres. Elles sont décrites ci-dessous.

(a) Sensibilités liées au risque de primes et de réserves

Le risque de réserve s'appuie sur l'estimation du stock des provisions dossier/dossier nettes de réassurance sans aucune marge/excédent implicite (« *best estimate* ») à partir des triangles de liquidation.

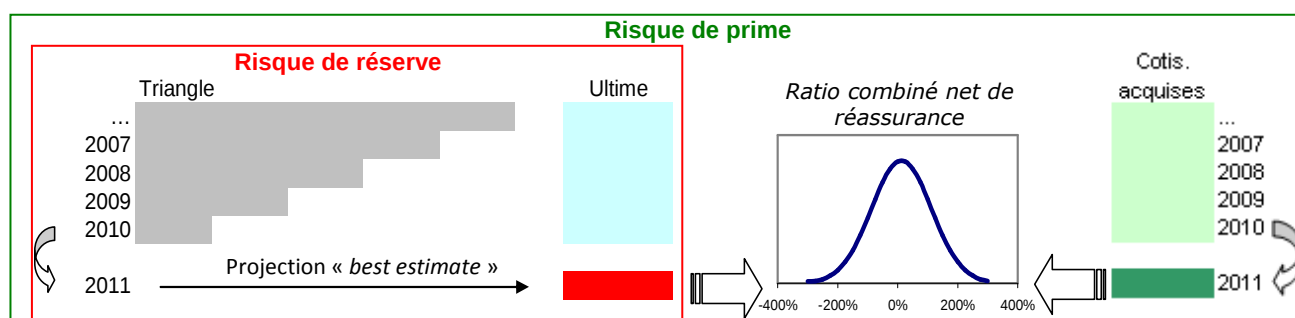
⁹ Cf Annexe [15.] : Triangle de liquidation

¹⁰ Boni : Sur-évaluation du coût de certains sinistres.

¹¹ α désigne le coefficient de corrélation entre le risque de réserve et le risque de prime

Le risque de prime en modèle interne s'appuie sur la distribution du ratio combiné par « *Line of Business* » Solvabilité 2 en net de réassurance. Pour rappel, le ratio combiné correspond aux charges de l'entreprise (« *best estimate* » des sinistres + frais de gestion de l'Entité) divisé par les cotisations acquises (par construction, ce risque est lié au risque de réserve).

Le graphique proposé ci-dessous illustre la mécanique de construction des volatilités associées aux métiers (« *LoB* ») S2 dans le modèle interne :



Graph 3. 4. Représentation graphique de la sensibilité liée au risque de prime et de réserve

Constats :

- Les volatilités S2 ont été ajustées aux volatilités observées sur l'historique, mais les calibrages réalisés sont fortement dépendants :
 - des profondeurs d'historiques,
 - des changements de profil de portefeuille.
- La sinistralité moyenne projetée est extrapolée à partir des tendances observées dans les historiques, avec un poids important des exercices 2008 à 2010 sur lesquels les ratios de sinistralité sont élevés (haut de cycle) et intègrent notamment l'occurrence de tempêtes.

(b) Sensibilités liées aux gains de diversification

Dans les spécifications techniques du QIS5, le NL_{pr} de primes et de réserves tient également compte des effets de diversification. Cette diversification s'échelonne tant au niveau du risque de prime et de réserve avec une corrélation fixée à 50% (cf. $\alpha = 0,5$) qu'au niveau segment Solvabilité 2 avec la corrélation $CorrLob_{r,c}$ définie selon la matrice ci-dessous :

Correlations	1	2	3	4	5	6	7	8	9	10	11	12
1: M (3 rd party)	1											
2: M (other)	0,5	1										
3: MAT	0,5	0,25	1									
4: Fire	0,25	0,25	0,25	1								
5: 3 rd party liab	0,5	0,25	0,25	0,25	1							
6: credit	0,25	0,25	0,25	0,25	0,5	1						
7: legal exp.	0,5	0,5	0,25	0,25	0,5	0,5	1					
8: assistance	0,25	0,5	0,5	0,5	0,25	0,25	0,25	1				
9: misc.	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1			
10: reins. (prop)	0,25	0,25	0,25	0,5	0,25	0,25	0,25	0,5	0,25	1		
11: reins. (cas)	0,25	0,25	0,25	0,25	0,5	0,5	0,5	0,25	0,25	0,25	1	
12: reins. (MAT)	0,25	0,25	0,5	0,5	0,25	0,25	0,25	0,25	0,5	0,25	0,25	1

Tableau 3. 4. Matrice de corrélation extraite du QIS5 pour les métiers non-vie

Ainsi, la corrélation entre « *LoB* » est définie telle que $\sigma = \sqrt{\frac{1}{V^2} \sum_{r,c} CorrLob_{r,c} \times \sigma_r \times \sigma_c \times V_r \times V_c}$

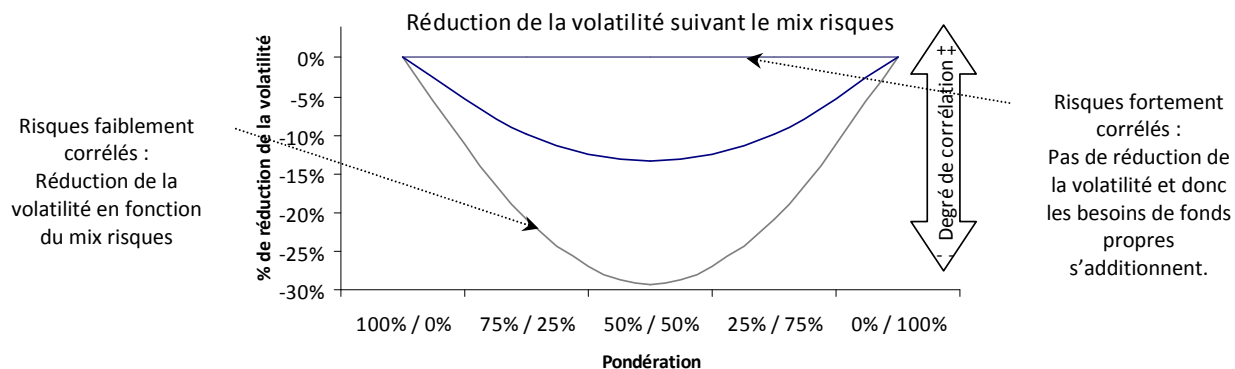
Constats :

Un métier contenant plusieurs risques (au sens Solvabilité 2) bénéficie donc d'une réduction de sa volatilité :

- Il y a des effets de compensation entre risques : Par exemple, le S/C à la probabilité 1/200 en RC Auto ne se réalise pas simultanément la même année que le risque à 1/200 en Dommage Auto (gains de diversification inter risques). La probabilité que ces deux risques se réalisent simultanément est au-delà de la probabilité 1/200 ;
- Le gain en termes de volatilité dépend du degré de corrélation entre les risques ;
- Un métier mono risque est pénalisé.

Ainsi, le « mix risques » a un impact direct sur le besoin de capital.

Afin d'étayer ces propos, le graphique ci-dessous illustre la diminution de volatilité (et donc l'effet de diversification) suivant la pondération de 2 risques (de même volatilité) :



Graph 3. 5. Représentation graphique de l'effet de diversification

(c) Sensibilités liées au risque de catastrophe

Le module (SCR) spécifique aux risques catastrophes prend en compte en formule standard :

- Différents risques catastrophes :
 - Les périls naturels : Tempête, Grêle, Inondation, Tremblement de terre, Subsidence ;
 - Les scénarios d'origine humaine : RC automobile, RC générale, conflagration ;
- Les montants en brut de réassurance obtenus par application de coefficients standard aux expositions Groupama renseignées par département ;
- Les couvertures de réassurance.

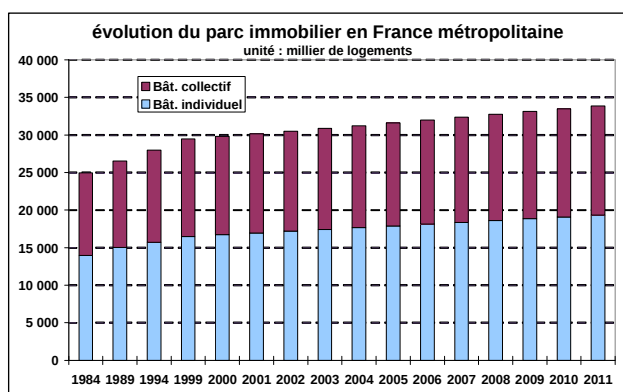
Dans le modèle interne, les limites (en lien avec l'étude) peuvent être décrites en 3 points :

- Sur la partie Tempête / Serres :
 - La sinistralité « Serres » ne fait pas l'objet d'une modélisation spécifique. L'application de la réassurance spécifique Serres des Caisses Régionales n'est donc pas possible. Néanmoins, l'impact est limité compte tenu des volumes mis en jeu.
 - La modélisation s'appuie entièrement sur le logiciel RMS. Or, comme il l'a été noté dans le chapitre III.1., il y a peut être un manque d'exhaustivité aux niveaux des données.
- Sur la partie « Cat. Nat. » :

Les risques « inondation », « tremblement de terre » et « sécheresse » ne font pas l'objet d'une modélisation catastrophe fine comme en « Tempête ». Ces risques sont appréhendés de manière macro par un S/C global pour l'ensemble des catastrophes couvertes par le dispositif « Cat. Nat. » de la Caisse Centrale de Réassurance (CCR). La couverture de réassurance offerte par la CCR est telle que les quantiles extrêmes de la distribution brute n'impactent pas le besoin de fonds propres¹².
- Sur le SCR MRH :

En 2011, le parc immobilier a progressé de 1,11% (cf. graphe ci-contre) pour s'établir à près de 33,9 millions de logements.

Les parts de logement individuel et de logement collectif s'établissent respectivement à 56,9% et 43,1% (à noter que 84,3% des logements sont des résidences principales, 9,5% des résidences secondaires et 6,2% des logements vacants).



Graph 3. 6. Évolution du parc immobilier en France métropolitaine (Source SOeS – CGDD, Comptes du logement)

Pour les Caisses Régionales, cette répartition est bien différente. En effet, sur le produit Privatis, la répartition logement individuel / logement collectif est respectivement de 74% / 26%. Dans le calcul du SCR MRH, l'impact est relativement important.

¹² Néanmoins, si les inondations liées à une tempête ou sécheresses ne sont pas assez importantes, la garantie mise en jeu est la garantie « Evénements Climatiques ». La prime pure proposée dans ce rapport tient compte de ces risques probables.

(d) Comparatif entre le modèle standard et le modèle interne

Pour conclure ce chapitre, il est proposé ci-dessous un comparatif entre les spécificités du modèle standard et celles du modèle interne :

Comparaison	Modèle standard	Modèle interne
Les facteurs de volatilité par segment de risque	<u>Risque de prime</u> - Identique quelle que soit la taille de l'Entité - Réassurance non prise en compte (quel que soit le niveau de réassurance, le facteur de volatilité est identique) <u>Risque de réserve</u> - Être en mesure de justifier la volatilité des portefeuilles (indicateurs de suivi des risques) pour éviter tous risques de « <i>Capital Add on</i> »	<u>Risque de prime</u> - Calibrage spécifique par « <i>business unit</i> » et type de sinistre (attritionnel, graves) - Prise en compte de la réassurance <u>Risque de réserve</u> - Calibrage des volatilités sur triangles Groupe avec prise en compte de l'effet volume (au niveau du Groupe)
Le résultat de l'année	- Ratio combiné moyen non pris en compte	- Prise en compte du résultat de l'année (capital supplémentaire en cas de Ratio combiné moyen supérieur à 100%) en diminution du besoin en fonds propres réglementaires à immobiliser
Les risques catastrophes	- Définis en formule standard - Application de la réassurance	- Calibrage spécifique sur les principaux risques - Prise en compte de la réassurance - En net de réassurance : proche formule standard
Les corrélations entre segments de risque	- 50% entre risque de prime et risque de réserve - Matrice de corrélation entre segments	- Idem formule standard

Tableau 3. 5. Table de comparaison entre le modèle standard et le modèle interne

Constats :

- Le modèle interne est plus précis et capte davantage les effets des portefeuilles non-vie.
- Les résultats actuels sont plutôt « conservateurs » par rapport aux ratios réellement constatés. Les mesures engagées par le Groupe en 2011 et 2012 et au-delà (tarif et réduction des frais) ne sont pas intégrées dans les résultats du besoin de fonds propres.
- Les différentes sensibilités pourront être reprises dans le cadre du dispositif ORSA.

Rappel objectif	Points retenus	Étape suivante
Analyse des sensibilités et comparaisons liées aux formules de calcul des modèles standard et interne	- Les volatilités de primes et de réserves pour chaque « <i>LoB</i> » S2 ont été calibrées sur l'historique portefeuille et sont donc fortement dépendantes : - des profondeurs d'historiques, - des changements de profil de portefeuille. - Un métier contenant plusieurs risques (au sens Solvabilité 2) bénéficie d'une réduction de volatilité selon le degré de corrélation entre les risques ; - Le SCR « <i>Cat</i> » lié au programme « Force de la Nature » : - s'appuie essentiellement sur le logiciel RMS (or, le positionnement est actuellement revu) ; - les risques « inondation », « tremblement de terre » et « sécheresse » ne font pas l'objet d'une modélisation fine ; - la répartition du portefeuille MRH est atypique. - Les décisions stratégiques opérationnelles ont un impact direct sur le besoin de capital économique.	Hypothèses liées à des modifications du solde de réassurance dans l'évaluation du tarif et des résultats techniques de la garantie « événements climatiques »

Comme il l'a été montré dans la Section II. et noté dans ce chapitre, le besoin en capital est sensible au programme de réassurance. De même, le programme de réassurance a un impact direct sur l'exigence tarifaire.

Afin d'en prendre davantage conscience et d'en percevoir les effets, il est développé, dans le chapitre suivant, des jeux d'hypothèses montrant l'impact direct de modifications du solde de réassurance dans l'exigence tarifaire.

III.3. Hypothèses liées à des modifications du solde de réassurance dans l'évaluation du tarif et de la rentabilité pour la garantie « événements climatiques »

Afin d'évaluer l'impact d'un changement du solde de réassurance, il est proposé ci-dessous une méthodologie basée sur une évaluation d'indicateurs de performances et rentabilités pour un produit d'assurance non-vie :

- Le chapitre III.3.1. présente les grandes lignes de ce principe.
- Par la suite, les différents indicateurs seront « stressés » afin d'observer les différentes stratégies de rentabilité possibles.

III.3.1. Evaluation de la rentabilité d'un produit non-vie

La méthodologie présentée ici, peut être déclinée pour chaque métier non-vie et pour l'ensemble des Entités du Groupe. L'approche consiste à définir des indicateurs de « performance cibles ».

L'évaluation de ces paramètres tient compte de caractéristiques propres à chaque métier, notamment :

- le benchmark externe de frais généraux et/ou de S/C (S/P),
- l'incidence des besoins de réassurance (dans le solde et les fonds propres),
- le rôle des produits financiers associés dans la formation des résultats, ...

La démarche proposée ici permet également de prendre en considération l'objectif global de rentabilité financière du Groupe et/ou de ses Entités (pour chaque métier, la contribution aux résultats des Entités est exprimée sous forme de ratio combiné).

Pour le métier MRH, le tableau se présente comme suit (les données présentées ci-dessous sont fictives de par leurs aspects confidentiels). Pour le moment, ces valeurs peuvent être lues, hors précisions, aussi bien en euro qu'en pourcentage :

	Col. A	Col. B	Col. C	Col. D	Col. E	Col. F	Col. G	Col. H	Col. I	Col. J
	Cotisations Origines (Base 100)	S/C brut espéré	Solde probable cédé en réassurance externe	Frais généraux	Résultat technique	Ratio combiné %	Produits financiers taux sans risque	Résultat économique	Fonds Propres allouées	ROE %
Indicateurs	100,0	-65,0	-6,0	-27,0	2,0	98,0	2,0	4,0	40	10,0

Tableau 3. 6. Evaluation de la rentabilité d'un produit non-vie

Interprétation du Tableau 3.5. :

Les cotisations « Col. A » (et les produits financiers calculés au taux sans risque, « Col. G ») doivent permettre de couvrir l'ensemble des charges suivantes :

- les sinistres et frais liés correspondent au S/C (décrits dans la « Col. B »). Ce S/C permet d'atteindre un objectif de résultat (dès lors que les frais généraux respectent le niveau cible prédéfini). Dans cette approche, le S/C cible est décliné afin d'être intégré à une démarche tarifaire,
- le solde probable cédé en réassurance externe (« Col. C »). Dans cette optique, les récupérations probables de sinistres sont calculées conformément au modèle interne, c'est-à-dire en les limitant aux tranches basses « travaillantes » des « Excess of Loss » et en écartant l'hypothèse d'événements majeurs ;
- les frais généraux techniques (« Col. D »). Ils correspondent à l'ensemble des frais de gestion des sinistres, d'acquisition, d'administration et autres produits et charges techniques. Ces frais généraux sont fixés à un niveau cible compatible avec les pratiques du marché ;

Hormis le résultat technique « Col. E », l'évaluation de la rentabilité d'un produit d'assurance non-vie peut également être spécifiée selon :

- le ratio-combiné (« Col. F ») qui est l'un des indicateurs les plus connus du marché. Il correspond à l'ensemble des frais de l'assureur rapporté aux cotisations acquises (ce qui se traduit sous forme mathématique, (« Col. B » + « Col. C » + « Col. D »)/ « Col. A ») ;
- le résultat économique (« Col. H ») qui correspond à la somme du résultat technique (« Col. E ») et des produits financiers calculés au taux sans risque (« Col. F ») ;
- le « ROE » (« Return On Equity » appelé également Rendement du Capital) qui est quant à lui le rapport entre le résultat économique de l'exercice (« Col. H ») et les fonds propres alloués disponibles à l'ouverture (« Col. I »). Le ROE permet d'intégrer tous les éléments qui constituent le résultat de la compagnie (techniques, financiers, gestion).

N.B. Il existe « normalement » une colonne supplémentaire intégrant le solde des autres postes du compte de résultats (éléments du compte non technique, impôts, retraitements de consolidation). Cette colonne ne figure pas dans le tableau afin d'en simplifier l'approche.

La méthode proposée ici privilégie la référence à des frais généraux cibles actualisés de manière à tenir compte des contraintes de compétitivité par métier, issues de l'observation du marché.

Sur les principes de « *Stress Test* », il va être à présent proposé des jeux d'hypothèses sur le solde de réassurance. Face à ces modifications, les indicateurs du *Tableau 3.6.* fluctueront et permettront d'en conclure les divers effets et opportunités qui peuvent se présenter.

III.3.2. Modifications du solde de réassurance et sensibilités liées

Dans ce chapitre, le « solde de réassurance » va être stressé afin d'évaluer son impact sur les divers indicateurs précédemment cités. Selon l'évolution des paramètres, différentes stratégies de rentabilité pour un produit d'assurance non-vie pourront être identifiées.

La ligne du *Tableau 3.6.* sera intitulée « Etat origine ». Elle reste toujours factice de par son caractère confidentiel, mais elle permet d'avoir une valeur de comparaison avec les scénarios proposés.

Les variables « Frais généraux » et « Produits financiers alloués » seront des indicateurs figés.

Suite à cette brève introduction, il vient le tableau (les valeurs peuvent être lues, hors précisions, aussi bien en euro qu'en pourcentage) :

	Col. A	Col. B	Col. C	Col. D	Col. E	Col. F	Col. G	Col. H	Col. I	Col. J
	Cotisations Origines (Base 100)	S/C brut espéré	Solde probable cédé en réassurance externe	Frais généraux	Résultat technique	Ratio combiné %	Produits financiers taux sans risque	Résultat économique	Fonds Propres alloués	ROE %
Etat origine	100,0	-65,0	-6,0	-27,0	2,0	98,0	2,0	4,0	40	10,0
Scénario 1	100,0	-65,0	-5,0	-27,0	3,0	97,0	2,0	5,0	52,6	9,5
Scénario 2	100,0	-65,0	-6,6	-27,0	1,4	98,6	2,0	3,4	40	8,5

Tableau 3. 7. Evolution des indicateurs de performances selon des modifications du solde de réassurance

Interprétations pour les tableaux 3.7 et 3.8 :

- Les valeurs sous fond jaune sont les paramètres de scénarios stressés et contraints ;
- Les valeurs sous fond blanc sont les valeurs cibles non modifiables ;
- Les valeurs sous fond rose sont les indicateurs qui ont été impactés par les modifications ;
- Les valeurs sous fond vert sont des ajustements possibles afin d'atteindre les valeurs imposées.

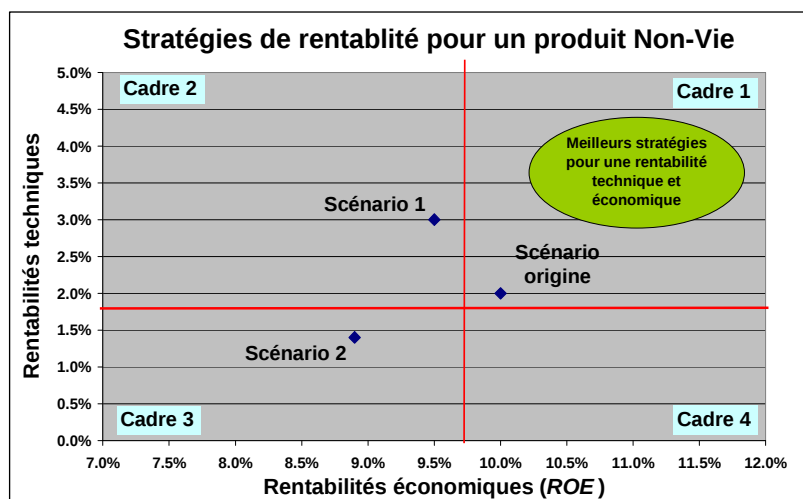
Descriptif des deux scénarios du Tableau 3.7. :

Scénario 1 : Le solde probable cédé en réassurance externe est fixé à -5,0€ (hypothèse d'une priorité plus haute, ou par exemple une suppression de la tranche T1 dans le traité). Le résultat technique ressort en progression et va faire évoluer le ratio combiné favorablement. Cependant, le ROE va évoluer de façon inverse. En effet, avec un tel scénario, que ce soit en modèle standard ou interne, le besoin en fonds propres est plus conséquent.

Scénario 2 : A présent, le solde probable cédé en réassurance externe est fixé à -6,6€ (hypothèse d'une augmentation de 10% des cotisations cédées aux réassureurs, pour un programme équivalent). Dans de telles conditions, si aucun changement n'est effectué, une baisse du résultat technique est à prévoir. Le ratio combiné passe à 98,6%. Le besoin en fonds propres n'ayant pas évolué, le ROE se contracte à un taux de 8,5%.

Face à de telles modifications, plusieurs stratégies de réassurance existent. Pour un produit d'assurance non-vie, ces options stratégiques sont définies selon la rentabilité attendue par le Groupe.

Ces politiques peuvent être transcrites sous forme de graphique en 4 typologies, comme suit :



Graph 3. 7. Représentation graphique des stratégies de rentabilité pour un produit Non-Vie.

Interprétations :

Cadre 1 : Cette partie supérieure droite correspond à la rentabilité optimale d'un produit d'assurance non-vie. Elle coïncide à l'« état origine ».

Cadre 2 : Ce cadre correspond au « Scénario 1 » et décrit une rentabilité technique importante, mais une rentabilité économique plus faible. Les besoins en fonds propres étant plus conséquents, le ROE se contracte.

Cadre 3 : La rentabilité technique et économique seraient insuffisantes au regard de la stratégie définie par le Groupe. C'est le cas du « Scénario 2 ».

Cadre 4 : Ce cadre correspond à un produit qui dégagerait peu de rentabilité technique, mais où les fonds propres alloués seraient relativement faibles.

En réponse aux scénarios précédents et pour illustrer les effets de la réassurance, il est proposé les deux cas de figures suivants (les valeurs doivent être lues en euro, hors précisions) :

	Col. A	Col. B	Col. C	Col. D	Col. E	Col. F	Col. G	Col. H	Col. I	Col. J
	Cotisations Origines (Base 100)	S/C brut espéré	Solde probable cédé en réassurance externe	Frais généraux	Résultat technique	Ratio combiné %	Produits financiers taux sans risque	Résultat économique	Fonds Propres allouées	ROE %
Etat origine	100,0	-65,0	-6,0	-27,0	2,0	98,0	2,0	4,0	40	10,0
Scénario 3	100,6	-65,0	-6,6	-27,0	2,0	98,0	2,0	4,0	40	10,0
Scénario 4	100,0	-76,0	5,0	-27,0	2,0	98,0	2,0	4,0	40	10,0

Tableau 3. 8. Scénarios complémentaires selon les modifications du solde de réassurance

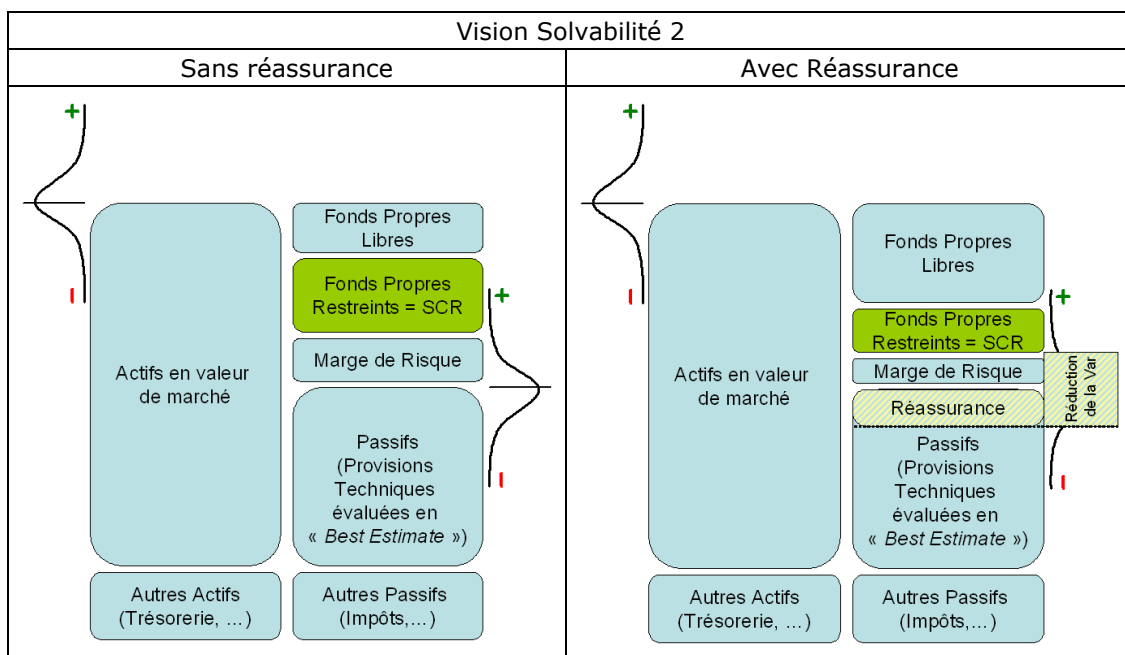
Descriptif des deux scénarios du Tableau 3.7. :

Scénario 3 : Le solde probable cédé en réassurance externe est de nouveau fixé à -6,6€ (hypothèse d'une augmentation de 10% des cotisations cédées aux réassureurs, pour un programme équivalent). Les résultats techniques cibles sont imposés à 2% afin de faire ressortir un ratio combiné de 98%. Puisque les frais généraux et les produits financiers ne peuvent être des variables d'ajustement, « jouer » sur les cotisations est le seul levier possible. Dans ce sens, une augmentation tarifaire de 0,6% est nécessaire pour atteindre les résultats objectifs fixés.

N.B. : Face à un environnement concurrentiel important, la plupart des acteurs du marché avait octroyé de nombreuses baisses tarifaires et compensé ces dernières par les produits financiers. Aujourd'hui, de par l'évolution des marchés financiers observée, les assureurs réajustent les cotisations à des niveaux plus proches de la réalité technique.

Scénario 4 : Dans ce scénario, le solde probable cédé en réassurance externe ressort à 5€. Ce scénario correspond à la survenance d'événements extrêmes qui ont fait jouer le programme de réassurance « Force de la Nature » (les possibles primes de reconstitution ont été intégrées dans ce solde). De par la survenance de cet événement, le S/C a atteint la valeur des 76€. Cependant, le traité de réassurance a joué tout son rôle de lissage, le résultat technique reste inchangé, tout comme le ROE.

L'ensemble de ces scénarios et le rôle de la réassurance peut être traduit graphiquement de la façon suivante :



Graph 3. 8. Représentation graphique des effets de la réassurance dans les besoins en fonds propres

Face à la survenance d'évènements majeurs, la réassurance joue donc un rôle prépondérant dans :

- les résultats techniques d'un organisme d'assurance,
- la tarification (nécessité de calibrer le niveau de la prime selon l'absorption du programme de réassurance et le solde probable lié),
- la réduction des besoins en fonds propres.

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Hypothèses liées à des modifications du solde de réassurance dans l'évaluation du tarif et de la rentabilité pour la garantie « événements climatiques »	- Descriptif d'indicateur d'évaluation - Modification des valeurs du solde de réassurance	La réassurance joue un rôle prépondérant dans : - le lissage des résultats techniques d'un organisme d'assurance, - la tarification (nécessité de calibrer le niveau de la prime selon l'absorption du programme de réassurance et le coût lié), - la réduction des besoins en fonds propres.	Modèles économiques : Jeux d'hypothèses dans l'allocation du capital et conséquences sur l'exigence de rentabilité

Un traité de réassurance montre donc tout son intérêt dans le lissage des résultats techniques et dans la réduction du besoin en fonds propres.

Malgré ces réductions de fonds propres avec un programme de réassurance, l'exigence de capital a fortement progressé avec l'arrivée de Solvabilité 2. De ce fait, la notion d'immobilisation du capital par « LoB » va entraîner une nouvelle perception de la rentabilité économique pour un produit d'assurance.

Un aperçu possible de ces perceptions est réalisé dans le chapitre suivant.

III.4. Modèles économiques : Jeux d'hypothèses dans l'allocation du capital et conséquences sur l'exigence de rentabilité

Hormis un cadre prudentiel renforcé, Solvabilité 2 implique une modification des pratiques actuelles en renforçant des spécificités telles que :

- l'intégration des risques,
- la considération du coût du capital,
- la création de valeur pour les actionnaires ou sociétaires.

La performance des métiers ne se basera plus uniquement sur les marges nettes réalisées. Elle devra parallèlement être analysée au regard des capitaux alloués sur chaque métier.

En effet, l'utilisation du capital n'est pas gratuite. Une certaine discipline capitalistique va donc s'imposer afin d'intégrer le coût du capital dans la mesure de la performance.

De nombreux modèles de capital économique vont ainsi s'imposer progressivement dans les organismes assurantiels, y compris dans le monde des mutuelles (Groupama n'échappe pas à cette règle), afin d'optimiser l'allocation des ressources et la gestion des risques portés.

Comme il l'a été dit auparavant, pour un exercice comptable donné, le *ROE* se calcule en faisant le rapport entre le résultat net comptable de l'exercice et les fonds propres disponibles à l'ouverture. Ce taux de rendement intègre tous les éléments qui constituent le résultat de la compagnie (techniques, financiers, gestion).

L'allocation de capital dépend bien évidemment du capital de « départ » à allouer et du modèle utilisé :

- Les modèles les plus simples se contentent de répartir le capital social entre les diverses Entités et/ou métiers.
- Les modèles les plus complexes permettent de répartir le capital sur des bases de constitution du capital global, en fonction de la volatilité propre à chaque Entité et/ou métier.

Afin de comprendre les enjeux liés à l'allocation du capital et son coût, deux types de modèles économiques, basés sur ces principes sont proposés ci-dessous.

La répartition des fonds propres sera réalisée en fonction de trois lignes métiers au sens Solvabilité 2 (en *Annexe [16.]*, il est proposé les mêmes modèles, mais avec quatre lignes métiers).

Hypothèses de départ pour les deux modèles :

La société commercialise trois produits, correspondant aux trois premiers métiers au sens Solvabilité 2. À partir du modèle standard, l'entreprise a été en mesure de déterminer pour chaque « *LoB* », le besoin en fonds propres nécessaires. Les hypothèses sont les suivantes :

	Métier 1 (LoB 1)	Métier 2 (LoB 2)	Métier 3 (LoB 3)	SCR total individuel
Hypothèses SCR	2 300	3 500	2 500	8 300
Répartition	27,7%	42,2%	30,1%	

Tableau 3. 9. Hypothèses de départ pour les modèles économiques

De plus, de par la formule standard, une diversification est à intégrer. Ainsi, le SCR total n'est pas de 8 300, mais de 6 478 (les valeurs sont exprimées en millions d'euros).

De manière détaillée, ce calcul est établi grâce à la matrice de corrélation du *Tableau 3.2.* et l'équation $SCR_{nl} = \sqrt{\sum CorrNL_{r,c} \cdot NL_r \cdot NL_c}$.

Par étape, il vient :

métier		SCR métier		(SCR corrélé) ²	Règle de comptage	« Impact »
M _r	M _c	NL _r	NL _c	CorrNL _{r,c}		
M1	M2	2 300	3 500	4 025 000	Compter 2 fois de par la matrice	8 050 000
M1	M3	2 300	2 500	2 875 000		5 750 000
M2	M3	3 500	2 500	2 187 500		4 375 000
M1	M1	2 300	2 300	5 290 000	Compter 1 fois	5 290 000
M2	M2	3 500	3 500	12 250 000		12 250 000
M3	M3	2 500	2 500	6 250 000		6 250 000
$\sum CorrNL_{r,c}.NL_r.NL_c =$						41 965 000

Tableau 3. 10. Etape de calcul dans le SCR final réglementaire

Enfin, dernière hypothèse à prendre en compte, l'entreprise souhaite une rentabilité économique de 10% (*ROE*=10%). Le résultat cible est donc de 648 (=6 478x10%).

Premier modèle économique :

Le modèle proposé ici est le plus simple et se contente de répartir proportionnellement les résultats cibles et les fonds propres réglementaires entre les trois métiers du dessus au prorata du ratio

$$\frac{SCR \text{ de la formule standard}}{SCR \text{ total individuel}} = \frac{6\,478}{8\,300}$$

Ces propos peuvent être traduits par le tableau suivant :

SCR déduit selon formule standard		6 478
Hypothèse ROE		10%
Résultat cible		648

	Métier M1	Métier M2	Métier M3	Somme des valeurs
Hypothèses SCR	2 300	3 500	2 500	8 300
Clé de répartition des contributions	27,7%	42,2%	30,1%	
SCR réalloué selon la clé	1 795	2 732	1 951	6 478
Résultat cible réalloué selon la clé	180	273	195	648
Contribution <i>a posteriori</i>	27,7%	42,2%	30,1%	
ROE <i>a posteriori</i>	7,8%	7,8%	7,8%	

Tableau 3. 11. Résultats du premier modèle de capital économique

Constats :

- La répartition du SCR standard et du résultat économique cible est répartie de manière identique selon la clé de répartition des contributions.
- Le ROE *a posteriori* est identique à chaque métier. Celui-ci n'est pas de 10% par métier, mais seulement de 7,8% (rapport entre le résultat cible et l'hypothèse de SCR initial par métier).

Deuxième modèle économique :

Le second modèle « plus complexe » permet de répartir les fonds propres réglementaires en fonction de la volatilité propre à chaque métier. Ainsi, le modèle économique suivant prend en compte les effets de diversification. Plus un métier demande des fonds propres importants et a des effets de diversification faibles, plus il sera demandé à celui-ci une rentabilité importante.

Pour réallouer les 6 478 de fonds propres réglementaires, la méthodologie est la suivante :

Si demain la société ne commercialisait que les métiers M1 et M2, l'exigence de fonds propres, selon la formule standard, serait de 5 059.

Par conséquent, l'intégration du métier M3 n'engendre qu'un surcroît de capital de 1 419.

Cette mécanique de retrait de métier, puis d'intégration, est calculée successivement et permet de déterminer des contributions « ajustées » à chaque métier (ou d'allocation marginale selon la littérature, le « tout sauf un »). En répétant la méthodologie, les résultats sont les suivants :

		Contributions ajustées				
M _i	M _j	SCR réglementaire	Contrib SCR All - x	Métier concerné	Rappel SCR initial	Baisse du SCR due à la diversification
M1	M2	5 059	1 419	M3	2 500	- 43,2%
M1	M3	4 158	2 320	M2	3 500	- 33,7%
M2	M3	4 783	1 695	M1	2 300	- 26,3%

Tableau 3. 12. Calculs des contributions ajustées

Ainsi, le modèle économique est le suivant :

	Métier M1	Métier M2	Métier M3	
Contribution initiale	27,7%	42,2%	30,1%	Somme des contributions ajustées
SCR "ajustée" (effet de diversification)	1 695	2 320	1 419	5 435
Effet de diversification (baisse du SCR initial)	- 26,3%	- 33,7%	- 43,2%	
Clé de répartition des contributions ajustées	31,2%	42,7%	26,1%	
				Somme des valeurs
SCR initial	2 300	3 500	2 500	8 300
SCR formule standard réalloué selon la clé	2 021	2 765	1 692	6 478
Résultat cible (10% sur le SCR réalloué)	202	277	169	648
Gains de SCR après application de la clé de répartition	-12,1%	-21,0%	-32,3%	
ROE <i>a posteriori</i> (calculé sur SCR initial)	8,8%	7,9%	6,8%	
Vérification de la contribution <i>a posteriori</i>	31,2%	42,7%	26,1%	

Tableau 3. 13. Résultats selon le second modèle de capital économique

Constats :

A présent, sur ce modèle économique, l'analyse est différente :

- L'activité M1 occupe un poids plus important que ne le laisse croire son *SCR* initial. Ce métier peut même être considéré comme l'activité la plus « risquée ». En effet, l'effet de diversification (rapport du « *SCR* ajusté » sur le « *SCR* initial » - 1) n'est que de 26,3%. Avec ce modèle, la rentabilité *a posteriori* exigée est de 8,8%, soit un taux bien supérieur à ce qui a été constaté dans le premier modèle économique.
- Pour l'activité M2, que ce soit sur la contribution *a posteriori* ou sur le *ROE*, les résultats sont quasiment similaires au premier modèle économique. Cependant, l'effet de diversification est intéressant, puisque les besoins en fonds propres initiaux baissent de 33,7%.
- L'activité M3 peut être considérée comme l'activité la plus « rentable ». L'exigence de résultat est la moins élevée et les effets de diversification sont importants (baisse du « *SCR* initial » de plus de 32%).

En mesurant la performance des métiers non plus selon le rapport marge nette sur cotisations, mais selon le rapport marge nette sur capital alloué (*ROE*), le métier le plus rentable peut devenir le métier qui possède la contribution la plus faible en termes de *SCR* (l'activité qui dégage la plus petite marge de la société peut ainsi être l'activité la plus rentable du Groupe).

Pour améliorer le modèle économique de la société, deux solutions sont encore possibles :

- Réduire les capitaux alloués au regard des sous-modules consommateurs de fonds propres est une première possibilité. Ce qui implique de se dégager de certains profils de risques ou de se développer sur des risques moins consommateurs.
- Gérer l'excédent de capitaux est la seconde solution. En effet, si Z% des capitaux disponibles de la société sont alloués à l'activité d'assurance, (1-Z)% « détruisent » de la valeur.

Les deux modèles présentés ici peuvent être nommés respectivement de :

- modèle économique avec approche proportionnelle,
- modèle économique avec approche marginale.

Rappel objectif	Méthodologies employées	Points retenus
Modèles économiques : Jeux d'hypothèses dans l'allocation du capital et conséquences sur l'exigence de rentabilité	- Deux types de modèles économiques ont été proposés	- Certaines « <i>LoB</i> » consomment beaucoup plus de capital que d'autres. Selon les gains de diversifications, des désengagements de certaines branches se feront, ou inversement un développement plus soutenu pourrait être souhaité. - En fonction du mixe « <i>Line of Business</i> » S2 qui compose une ligne métier, l'immobilisation de fonds propres à mettre en face du risque peut être plus ou moins importante.

Point de transition :

Avant d'effectuer toutes opérations relatives aux conclusions de la Section II., il était nécessaire d'apprécier la justesse de la méthodologie proposée.

L'analyse réalisée dans cette section a permis de détecter un ensemble de sensibilités.

Au regard de la détermination des échelles de risques, la seule sensibilité semble être le manque d'exhaustivité des données sur les échelles des coûts moyens (pour les habitations de plus de 10 pièces). L'intégration d'un exercice supplémentaire permettrait d'ajuster définitivement l'approche, en rendant l'échelle de risque plus robuste.

Pour ce qui est de la détermination du niveau du coût du risque, deux sensibilités sont à retenir :

- les valorisations des SMP réalisées par la DCER sont probablement prudentes. En effet, certaines Entités ont noué des partenariats rendant les coûts de réparations des sinistres moindres au regard d'une situation classique.
- les modifications possibles dans le programme de réassurance (que ce soit sur l'ensemble des risques couverts par le traité, les Entités, les priorités ou le coût lié à la facture de réassurance) ont un impact direct sur le niveau de tarification.

Face à ces divers éléments, la méthodologie proposée peut être validée. Seule la détermination du niveau du coût du risque est « discutable », mais ce paramètre actuel ne relève qu'un caractère prudentiel. A noter, de par l'approche proposée, que ce niveau pourra être rapidement actualisé à un degré adapté (une fois les mises à jour effectuées).

Hormis un cadre prudentiel renforcé, Solvabilité 2 implique également une modification des pratiques actuelles en renforçant des spécificités telles que :

- la considération du coût du capital,
- la création de valeur pour les actionnaires ou sociétaires.

L'intégration du coût du capital et les effets de diversifications des risques sont des aspects encore peu exploités, mais qui vont être de nature à influencer fortement l'exigence tarifaire dans les années à venir.

La performance des métiers ne se base plus uniquement sur les marges nettes réalisées. Elle doit parallèlement être analysée au regard des capitaux alloués sur chaque métier. L'utilisation du capital n'est pas gratuite. Une certaine discipline capitalistique va donc s'imposer afin d'intégrer le coût du capital dans la mesure de la performance.

De nombreux modèles de capital économique vont s'immiscer progressivement dans les organismes assurantiels, y compris dans le monde des mutuelles (Groupama n'échappe pas à cette règle) afin d'optimiser l'allocation des ressources et la gestion des risques portés.

Le projet de loi déposé au Sénat (et enregistré le 3 avril 2012) réformant le régime d'indemnisation des catastrophes naturelles laisse présumer des désengagements de l'Etat, ce qui devrait avoir un impact direct sur la prise en charge des assureurs. Par conséquent, la réflexion qui a été réalisée jusqu'à présent sur la garantie « événements climatiques » va être portée plus loin.

Face aux perspectives liées à la réforme du système « Cat. Nat. », il sera présenté dans la Section IV., une réflexion sur ces risques et les différentes mesures qui pourraient être prises pour y répondre.

Section IV. Les perspectives face aux risques naturels

Face à l'impact probable de la réforme du système « Cat. Nat. », des questions d'« assurabilité » vont se poser :

- Dans quelles proportions, l'Etat va-t-il se désengager de certains périls ?
- Les assureurs devront-ils prendre à leur charge ces risques ou s'en séparer ?

La présence humaine sur des zones à risque doit induire une réflexion sur les possibles aménagements, tant par les secteurs de l'assurance et de la réassurance que les Pouvoirs publics : En effet, est-il encore raisonnable d'accepter des constructions et habitations sur des zones connues pour avoir déjà été sinistrées (en inondations, mouvements de terrain, ...) ?

Pour un organisme d'assurance, avec une cartographie détaillée des zones sinistrées et « sinistrables », un rôle de prévention ou d'incitation (moralisation, selon la littérature), par une sélection ou segmentation tarifaire plus fine, pourrait être élaborée. Pour des produits d'assurance, tels que les produits MRH, les risques économiques sauraient être réduits.

Pour répondre à cet objectif, cette section se découpe en 4 chapitres :

- Le premier chapitre exposera certains points du projet de loi portant sur la réforme du régime d'indemnisation des catastrophes naturelles enregistré à la Présidence du Sénat le 3 avril 2012 ;
- Les possibles désengagements de l'Etat énoncés dans ce projet inciteront peut être les réassureurs à changer leurs garanties. Face à ces probables modifications, il sera présenté, dans le deuxième chapitre, un autre dispositif alternatif de couverture : Les dérivés climatiques.
- De par les probables exigences supplémentaires des réassureurs et les dispositions énoncées dans le projet de loi, la possibilité d'ajouter des critères tarifaires dans l'actuel algorithme sera traitée dans le troisième point ;
- Pour clore les travaux, le quatrième chapitre traitera le côté prévention. Un des aspects souvent omis, mais qui peut permettre de réduire significativement les coûts et l'exposition aux risques ;

Les travaux relatifs au chapitre IV.3.2. sont de nature confidentielle et sont en cours de réalisation au sein de la DADIP¹. Par conséquent, il ne sera exposé que des aspects dont le marché a connaissance, mais qui ne sont pas encore employés au sein du produit Privatis.

IV.1. Projet de loi portant sur la réforme du régime d'indemnisation des catastrophes naturelles

Le régime d'indemnisation des catastrophes naturelles a été institué par la loi n° 82-600 du 13 juillet 1982. Ce régime hybride (mi-publics, mi-privés, via un partage des risques assureur-Etat grâce à la CCR²), sans véritable équivalent en Europe et concentré sur la fonction de réassurance, a démontré, depuis sa création, son efficacité. Il procure une large couverture contre les dommages résultant de catastrophes naturelles (avec une garantie illimitée de l'Etat) pour un coût « modéré ».

Dans un contexte mondial et national marqué ces dernières années par une récurrence accrue de catastrophes naturelles, il a cependant révélé deux faiblesses :

- un cadre juridique imprécis, dommageable à la transparence et à l'équité du régime, et
- des mécanismes d'incitation à la prévention insuffisants.

Aujourd'hui, si le régime est financièrement équilibré (pour le moment), il reste toutefois insuffisamment provisionné au regard d'une catastrophe naturelle d'ampleur exceptionnelle, de type crue centennale de la Seine.

Le projet de loi présenté et enregistré à la Présidence du Sénat le 3 avril 2012 apporte des modifications au Code :

- des assurances
- de la construction et de l'habitation.

¹ DADIP : Direction Assurances Dommages Individuels et Particuliers

² CCR : Caisse Centrale de Réassurance

Ce projet a pour ambition de moderniser durablement le régime d'indemnisation des catastrophes naturelles. Les quelques dispositions susceptibles d'impacter fortement les compagnies d'assurance sont brièvement présentés ci-dessous :

Disposition du **chapitre I^{er}** renforçant la transparence et l'équité du régime

L'**article 1^{er}** modifie les articles L.125-1, L.125-5 et L.125-6 du Code des assurances. Il précise d'une part, le cadre juridique du régime (en particulier son périmètre d'intervention) et améliore d'autre part, le fonctionnement et la transparence de la procédure de reconnaissance de l'état de catastrophe naturelle :

- en prévoyant la fixation d'une liste des phénomènes éligibles au régime d'indemnisation et la définition de la méthodologie permettant aux experts scientifiques d'apprécier l'intensité anormale des agents naturels qui sont causes de ces phénomènes ;
- en délimitant clairement l'intervention de l'assurance construction et de la garantie contre les effets des catastrophes naturelles en matière d'indemnisation des dommages résultant d'un mouvement de terrain différentiel consécutif à la sécheresse et à la réhydratation des sols ;
- en actualisant les conditions du bénéfice de la garantie contre les effets des catastrophes naturelles à l'encontre des biens construits et activités exercées sur des terrains classés inconstructibles par un plan de prévention des risques naturels prévisibles.

Dans l'article L.125-6, il est prévu que les trois premiers alinéas soient remplacés (il n'est précisé ici que les modifications liées au premier alinéa) :

« Dans les terrains situés dans des zones inconstructibles d'un plan de prévention des risques naturels prévisibles approuvé ou appliqué par anticipation dans les conditions fixées par les dispositions du chapitre II du titre VI du livre V du Code de l'environnement, l'assuré est déchu du bénéfice de la garantie contre les effets des catastrophes naturelles s'agissant de dommages causés à des biens mentionnés à l'article L.125-1 par un phénomène sur lequel porte le plan, lorsque ces biens ont été construits postérieurement à la publication du plan, sans autorisation administrative de construire ».

Constat :

La proposition suivante est plutôt satisfaisante, néanmoins, d'un point de vue juridique, elle laisse supposer qu'il est possible de construire sur une zone inconstructible, sous réserve d'une « autorisation administrative ». Ce genre de propos est dommageable. Certaines pratiques actuelles (autorisation de construire en zones à risque) vont donc pouvoir perdurer. Dans un tel cadre, une responsabilité sur les autorités délivrant cette autorisation devrait être exigée.

Disposition du **chapitre II** renforçant les mécanismes incitant à la prévention

L'**article 2** modifie l'article L.125-2 du Code des assurances. Il instaure la possibilité d'une modulation encadrée des primes versées par les assurés pour la garantie contre les effets des catastrophes naturelles, en la réservant aux acteurs qui disposent des outils en matière de renforcement de la prévention : les collectivités locales et les entreprises à partir d'une certaine taille.

Il est prévu dans l'article L.125-2 que le troisième alinéa soit modifié par les dispositions suivantes :

« La garantie est accordée en contrepartie du paiement d'une prime ou cotisation additionnelle, individualisée dans l'avis d'échéance du contrat mentionné à l'article L.125-1 et calculée à partir d'un taux défini pour chaque catégorie de contrat. Ce taux, fixé par arrêté du ministre chargé de l'économie, est appliqué au montant de la prime ou cotisation principale ou au montant des capitaux assurés, selon la catégorie de contrat. Toutefois, un décret en Conseil d'État détermine les limites minimales et maximales entre lesquelles, pour les contrats qui couvrent des risques autres que ceux des particuliers et qui portent sur un montant de capitaux assurés dépassant un seuil qu'il fixe, ce taux peut varier. » ;

Constat :

La proposition suivante laisse la possibilité de réaliser une tarification plus discriminante (modulé par un taux pour les collectivités locales et les entreprises, ce qui est une avancée relativement intéressante).

Cependant, le taux de 12% actuellement appliqué sur les contrats MRH pour la couverture de la garantie « catastrophes naturelles » ne pourra probablement pas encore être modulé selon le risque de l'entité assuré.

Il est toutefois possible de « moduler » la cotisation « catastrophes naturelles » en MRH selon le capital assuré. En effet, selon l'Article L.125-2, « Les entreprises d'assurance doivent insérer dans les contrats [...] une prime ou cotisation additionnelle, individualisée dans l'avis d'échéance du contrat visé à l'article L. 125-1 et calculée à partir d'un taux unique défini par arrêté pour chaque catégorie de contrat. Ce taux est appliqué au montant de la prime ou cotisation principale ou au **montant des capitaux assurés**, selon la catégorie de contrat. »

Cet aspect de modulation selon le capital assuré est actuellement peu utilisé en assurance MRH. Cependant, c'est une perspective intéressante dans le cadre d'une moralisation du risque.

Les dispositions suivantes n'impactent pas le Code des assurances, mais le caractère des articles L.112-23. et L.112-25 (destinés Code de la construction et de l'habitation) est intéressant.

Disposition du **chapitre II Sous-Section 2** renforçant la prévention des risques de mouvement de terrain différentiel consécutif à la sécheresse et à la réhydratation des sols

L'**article 4** complète l'article L.112-19 du Code de la construction et de l'habitation par 6 autres articles.

Sur ces 6 articles, en complément d'une étude de sol annexé à tout logement, il est prévu par l'article L.112-23. que *« Lorsqu'un contrat a pour objet des travaux de construction ou la maîtrise d'œuvre d'un ou de plusieurs immeubles à usage d'habitation ou à usage professionnel et d'habitation ne comportant pas plus de deux logements, le constructeur de l'ouvrage est tenu, sauf si l'absence de risque de mouvement de terrain différentiel consécutif à la sécheresse et à la réhydratation des sols est en l'espèce avérée :*

« 1° Soit de respecter des techniques particulières de construction définies par voie réglementaire ;

« 2° Soit de suivre les recommandations d'une étude géotechnique fournie par le maître d'ouvrage ou que le constructeur fait réaliser, qui prend en compte l'implantation et les caractéristiques du bâtiment.

L'article L.112-25. prévoit quant à lui *« Un décret en Conseil d'État définit les modalités d'application de la présente sous-section. Il précise notamment :*

« 1° Les contrats entrant dans le champ d'application des articles L. 122-22 et L. 122-23 qui, en raison de la nature ou de l'ampleur limitée du projet, ne sont pas soumis aux dispositions de ces articles ;

« 2° Les conditions dans lesquelles les études réalisées en application des articles L. 112-21, L. 112-22 et L. 112-23 sont transmises aux entreprises d'assurance ainsi qu'à des organismes scientifiques et d'expertise technique. »

Constats :

- Les compagnies d'assurances pourront donc demander lors de la souscription d'une entité assurée, les études de sols et de constructions réalisées. De par ces informations, la mesure du risque assurée devrait être fine, notamment :

- plus de fiabilité dans le calcul du capital assuré par la connaissance des matériaux de construction utilisés ;
- meilleure connaissance du risque sous-jacent au sol.

- Il est également probable que les recours soient plus nombreux sur les contrats « risque décennal » en cas de non respect des normes de constructions.

Rappel objectif	Points retenus	Étape suivante
Projet de loi portant sur la réforme du régime d'indemnisation des catastrophes naturelles	- Sous réserve d'autorisation, il est toujours possible de construire sur des zones inconstructibles. C'est un point dommageable. - Les taux actuellement fixés par la législation pourront être a priori modulable selon le risque couvert. - Les compagnies d'assurances auront l'opportunité de demander, lors de la souscription d'une entité assurée, les études de sols et de constructions réalisées.	Les dérivés climatiques

De par les dispositions énoncées dans le projet de loi, les probables désengagements de l'Etat sur certains périls entraîneront des exigences supplémentaires de la part des réassureurs. Il sera donc traité dans le point suivant une alternative face aux possibles impacts des réassureurs.

IV.2. Les dérivés climatiques

Depuis quelques années, le marché de la réassurance subit un phénomène de concentration assez important. En effet, en 1990, les 10 principaux réassureurs mondiaux représentaient 32% du marché et les 5 premiers s'accaparaient 23% du marché. En 2007, les 10 premiers réassureurs détiennent 61% du marché de la réassurance, dont 50% attribuables aux 5 principaux³.

De par les dispositions énoncées dans le projet de loi, les probables désengagements de l'Etat sur certains périls entraîneront des exigences supplémentaires de la part des réassureurs. Face à cette concentration du marché, les possibilités de concurrence et de couverture vont se restreindre. Les réassureurs pourraient donc être amenés à :

- augmenter les primes de réassurance,
- exiger davantage d'information,
- baisser certaines couvertures.

En parallèle, dans les années qui viennent de s'écouler, les diverses compagnies d'assurances ont constaté que les besoins de réassurance dépassaient ce qui pouvait être proposé par les moyens classiques.

De nouvelles possibilités ont donc été envisagées, en particulier par le biais des marchés financiers : la titrisation.

Le plus courant des produits connus est l'obligation catastrophique (appelé « *CAT bond* »). Il s'agit d'une obligation émise par une compagnie d'assurance, payant un taux d'intérêt supérieur au taux usuel du marché. En contrepartie de cette rémunération supplémentaire, le détenteur d'obligation s'engage à fournir de la réassurance en excédent de sinistre. Selon les clauses de l'obligation, les intérêts et/ou le principal pourront servir à indemniser les sinistres⁴.

À la différence des dérivés classiques, l'« actif » sous-jacent n'a donc aucune valeur puisqu'il peut représenter la température, la quantité de neige tombée, les précipitations,

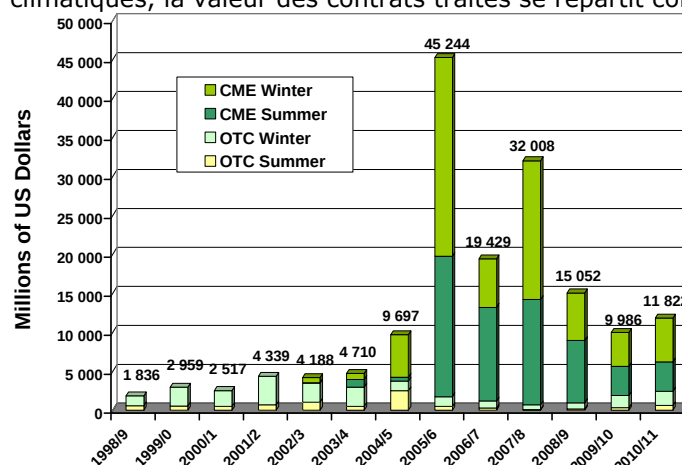
L'absence de corrélation significative entre ces titres et la rentabilité du marché est une source intéressante de diversification pour un investisseur.

Pour ce dernier, les obligations « *CAT* » engendrent (*a priori*) :

- une rentabilité un peu plus élevée que celle proposée par le marché avec une forte probabilité, et,
- une perte conséquente avec une probabilité faible.

Le prix du « *Cat. Bond* » et les coupons sont déterminés de telle sorte qu'ils compenseront l'immobilisation de l'argent investi à laquelle sera ajoutée une prime de risque liée à une probable survenance d'évènement(s) exposant le principal et/ou les coupons à une diminution (ou annulation).

Les premiers dérivés climatiques ont été négociés de gré à gré (« *OTC* ») à partir de 1997 aux Etats-Unis. Sur le Chicago Mercantile Exchange (« *CME* »), seul marché organisé de dérivés climatiques, la valeur des contrats traités se répartit comme suit entre les années 2000 et 2011 :



Graph 4. 1. Répartition des différents types de contrats climatiques négociés de gré à gré au cours de la période 2000/2011.

Source : Weather Risk Management Association et Price Water Cooper

³ JAL Pierre (2010), *Réassurance Non-Vie*, ISFA, Université Claude Bernard, Lyon 2

⁴ Ces propos s'appuient sur l'ouvrage de HULL John (2007), *Options, futures et autres actifs dérivés*, 6^{ième} édition, Pearson Education et l'article de CHARPENTIER Arthur (2002), *Titrisation des risques catastrophes : les Cat Bonds*, FFSA – Direction Etudes, Statistiques & Systèmes d'information

Constats :

- Entre 2005 et 2006, le marché des dérivés climatiques a dépassé 45 milliards de dollars, un montant 5 fois plus important que la saison précédente.
- Sur les années 2006 et 2008, il est possible d'observer un ralentissement. Celui-ci est lié à la hausse des taux d'intérêt et donc une diminution de la liquidité.
- En 2008/9, un repli important est de nouveau observable. Cette chute correspond à la « tempête financière » liée à la crise des « subprimes » (en lien avec l'augmentation des taux d'intérêt citée précédemment).

Malgré l'attrait que peut présenter ce type de couverture pour une société d'assurance, un inconvénient est à soulever. Afin de répondre à ses engagements (pourvoir les coupons), la compagnie devra « sortir » ou « dégager » des résultats plus importants. Un des moyens d'y parvenir est de réaliser des augmentations tarifaires supplémentaires. Dans un contexte concurrentiel accru, où les prix de l'assurance sont tendus, l'approche reste délicate.

Rappel objectif	Points retenus	Etape suivante
Les dérivés climatiques	- Les dérivés climatiques offrent une couverture supplémentaire à la réassurance. - La diversification de couverture peut inciter la réduction des primes exigées par les réassureurs. - Les exigences supplémentaires de rentabilité influenceront directement sur les prix des produits d'assurances.	Un nouvel algorithme dans l'approche tarifaire des risques naturels

Face aux probables exigences supplémentaires des réassureurs et de par les dispositions énoncées dans le projet de loi, les compagnies se devront de réduire leurs risques. Un des premiers leviers est de réaliser une segmentation tarifaire plus poussée avec l'intégration de critères tarifaires supplémentaire. C'est ce qui est traité dans le point suivant.

IV.3. Un nouvel algorithme dans l'approche tarifaire des risques naturels

Il est rappelé, selon les conditions générales du produit Privatis, que la garantie « événements climatiques » couvre les détériorations accidentelles subies par le bâtiment ou le mobilier usuel résultant de l'action directe :

- du vent ou du choc d'un corps renversé ou projeté par le vent,
- de la grêle sur les toitures,
- du poids de la neige ou de la glace sur les toitures,
- d'une avalanche,
- de glissement ou affaissement de terrain (lorsqu'un arrêté interministériel reconnaît l'état de catastrophes naturelles, c'est la garantie « catastrophes naturelles » qui intervient).

La probable réforme du système « Cat. Nat. » laisse présumer des désengagements de l'Etat, ce qui devrait donc avoir un impact direct sur la prise en charge des assureurs, notamment pour Groupama au vu des conditions générales du produit Privatis.

Néanmoins, ce projet de loi laisse également une perspective intéressante : un premier pas vers un taux modulable est envisagé (pour les collectivités locales et les entreprises à partir d'une certaine taille).

Ainsi, un des premiers aspects envisagé pourrait être de décomposer tarifairement les aléas naturels dans leurs globalités, en créant des garanties distinctes propres. Par exemple, des garanties libellées :

- « risques météorologiques » couvrant le risque tempête (vents violents), les risques grêles et neiges pourraient peut-être faire l'objet là aussi d'une tarification distincte ;
- « risques climatiques » couvrant le risque de sécheresse (et donc le risque de mouvement de terrain également appelé subsidence : différentiel des sols consécutif à la sécheresse et à la réhydratation) ;
- « risques d'inondation » couvrant les débordements, les ruissellements et les submersions.

De même, l'opportunité de revoir la détermination des échelles de risque par l'intégration de variables supplémentaires est également soulevée, notamment avec des critères de géolocalisations (zoniers).

Les échelles déterminées dans la Section II. (additionnées de variables pertinentes et discriminantes) pourraient vraisemblablement porter la même information, mais avec un zonier géographique différent selon le risque couvert. La détermination des échelles des zoniers pourrait par exemple être établie sur les résidus des modèles déterminés précédemment.

Cette perception géographique vis-à-vis des risques est présentée dans le chapitre IV.3.1. suivant.

IV.3.1. Géolocalisations des risques naturels via les données de la MRN et présentées via le logiciel MAP info

Après une année particulièrement sinistrée par des événements catastrophiques naturels (inondations et tempêtes), la FFSA (Fédération Française des Sociétés d'Assurance) et le GEMA (Groupement des Entreprises Mutuelles d'Assurances) ont créé, début 2000, la Mission des Risques Naturels (MRN).

Cette association propose à ses membres des services d'analyse et d'interprétation des données publiques qu'elle collecte et restitue. Afin d'en faciliter leurs utilisations, la MRN les adapte par le biais de méthodes et d'outils spécifiques (base de données, Système d'Information Géographique de la MRN, études spécifiques).

La MRN précise néanmoins que ses travaux sont contraints par :

- L'hétérogénéité et le manque de fiabilité des données publiques, qui ne permettent pas de construire des modèles d'aléas climatiques fiables.
- La réticence des assureurs français à partager au sein de la MRN les informations de sinistralité concernant leurs portefeuilles. La MRN n'a pas encore pu réaliser des études de vulnérabilité (l'endommagement consécutif aux événements) des portefeuilles d'assurance tel que l'outil RMS peut le proposer.

N.B. : Alors que les initiatives de mise en commun des données de sinistralité de chaque portefeuille sont quasi inexistantes en France, d'autres pays d'Europe, comme la Suisse et l'Allemagne, pratiquent ce partage d'informations de manière plus généralisée et encadrée.

Dans le cadre des échanges de données, la MRN a été en mesure de fournir une géolocalisation d'un ensemble de tables portant sur les risques :

- inondations,
- sécheresses,
- mouvements de terrain,
- sismiques,
- avalanches.

Dans ces tables fournies, il est à noter la présence du code IRIS.

Une définition de cet indicateur est donnée par l'INSEE :

Les communes d'au moins 10 000 habitants et la plupart des communes de 5 000 à 10 000 habitants sont découpées en IRIS⁵. Ce découpage, maille de base de la diffusion de statistiques infra-communales, constitue une partition du territoire de ces communes en "quartiers" dont la population est de l'ordre de 2 000 habitants. La France compte plus de 56 000 IRIS sur la métropole.

Par extension, afin de couvrir l'ensemble du territoire, un IRIS est assimilé à chacune des communes non découpées en IRIS.

Ce découpage a été élaboré en partenariat avec les interlocuteurs locaux, notamment les communes, selon des règles précises définies en concertation avec la Commission Nationale Informatique et Libertés (CNIL). Il est construit à partir de critères géographiques et statistiques et, chaque IRIS doit être, autant que possible, homogène du point de vue de l'habitat. Les IRIS offrent l'outil le plus élaboré à ce jour pour décrire la structure interne de près de 1 900 communes d'au moins 5 000 habitants.

Dans la suite de ce chapitre, seuls les quatre premiers risques sont cartographiés.

(a) Géolocalisation du risque inondation

Quel que soit le type de cours d'eau considéré, la position générale retenue en France, et par la MRN, est de prendre en compte une crue de référence correspondant à la plus forte crue historique connue, dans la mesure où sa durée de retour est au moins centennale.

Dans le cas contraire, la crue centennale est définie de manière théorique, par extrapolation des informations historiques disponibles, et en tenant compte du fonctionnement actuel du bassin versant et des aléas pluviométriques susceptibles de l'affecter.

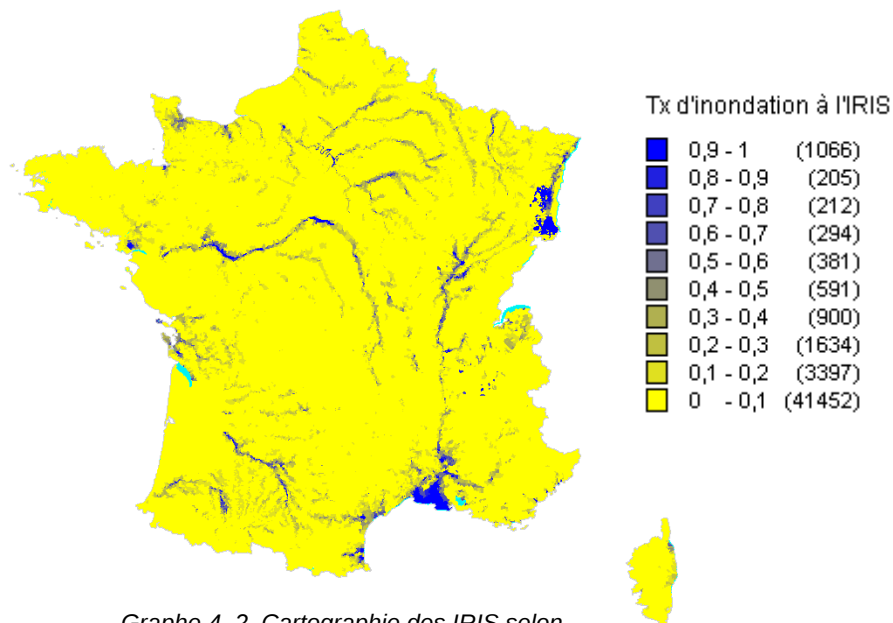
⁵ Source INSEE

Dans les tables « inondations » fournies par la MRN, il est à noter la présence :

- du Code IRIS,
- du m^2 de surface de l'IRIS,
- du m^2 de surface inondable sur cette IRIS.

A partir de ces valeurs, un taux de surface inondable par IRIS peut être aisément calculé et cartographié via l'outil MAP info.

Des informations à disposition, il en ressort la cartographie suivante :

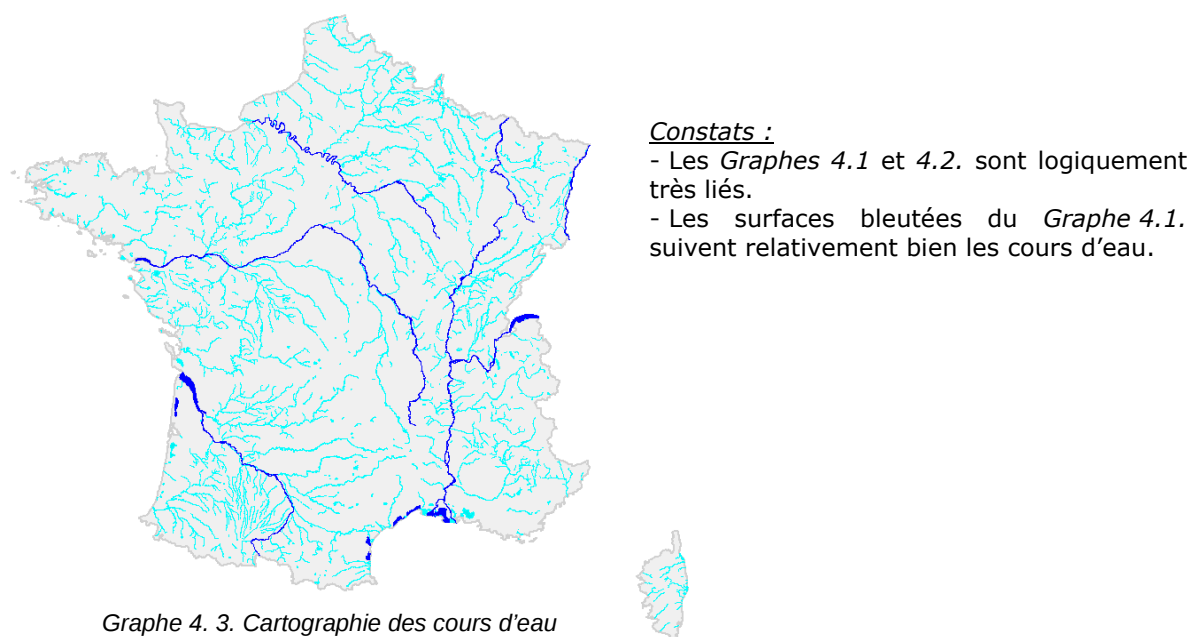


Graphe 4. 2. Cartographie des IRIS selon leur taux de surface inondable

Constats :

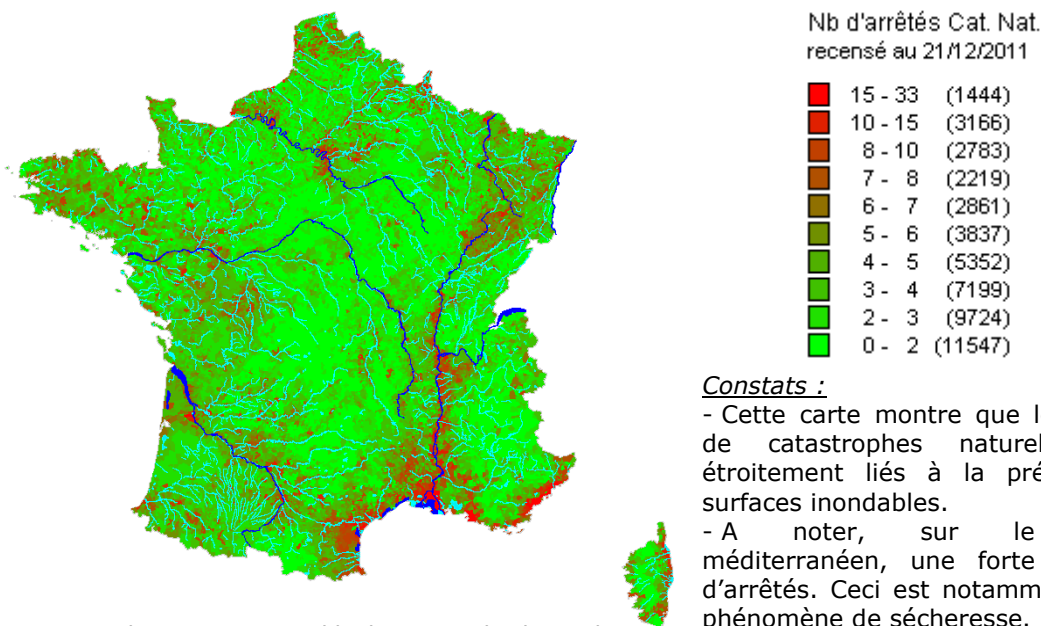
- Les taux de surfaces inondables par IRIS délimitent assez nettement les cours d'eau cartographiés sur le *Graphe 4.2*.
- De fortes expositions sont remarquables sur les bassins alsaciens et méditerranéens
- Une coupure brutale de la Gironde est observable. Lors de l'envoi des expositions, la MRN revoyait les aléas sur cette région.

Grâce au logiciel MAP info et les cartes fournies par ce dernier, il est possible de visualiser les principaux cours d'eau en France.



Graphe 4. 3. Cartographie des cours d'eau en France selon les données de MAP info

La MRN a également fourni dans son fichier, le nombre d'arrêtés « Cat. Nat. » recensés jusqu'au 21/12/2011. Ces arrêtés ont également pu être cartographiés avec le logiciel MAP info.



Graph 4. 4. Cartographie des IRIS selon le nombre d'arrêtés catastrophes naturelles recensé jusqu'au 21/12/2011.

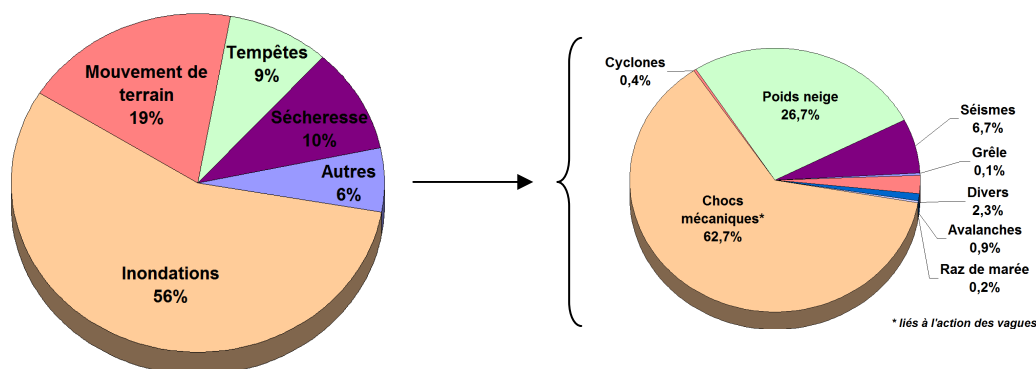
Constats :

- Cette carte montre que les arrêtés de catastrophes naturelles sont étroitement liés à la présence de surfaces inondables.
- A noter, sur le bassin méditerranéen, une forte présence d'arrêtés. Ceci est notamment dû au phénomène de sécheresse.

Sur la période 1982-2010, les inondations constituent en France l'évènement ayant donné le plus lieu à indemnisation par le régime « Cat. Nat. ». En termes de répartition, les évènements se décomposent ainsi :

- 55% d'inondations,
- 41% de sécheresses,
- 4% d'autres aléas⁶.

En termes de fréquence d'arrêtés, la répartition est la suivante :



Graph 4. 5. Répartition des arrêtés par nature de périls (cumul 1982-2010)

Constats :

- Là où la charge sinistre « inondation » représente 55% des indemnisations, il y a en nombre 56% d'arrêtés « Cat. Nat. » liés à ce risque.
- Pour la sécheresse, le second poste de sinistre en indemnisations (41%), la fréquence n'est que de 10%. Il semblerait donc que le coût du risque « sécheresse » soit plus élevé que celui des inondations.
- Les mouvements de terrain représentent 19% des arrêtés « Cat. Nat. ». C'est le second poste en termes de fréquence.

N.B. : Le phénomène de subsidence (mouvement de terrain) est un phénomène lié au risque sécheresse. Ce risque est présenté dans le paragraphe (c).

⁶ Chiffres extraits de l'article, *Synthèse de l'Etude relative à l'impact du changement climatique et de l'aménagement du territoire sur la survenance d'évènements naturels en France*, avril 2009, Fédération Française des Sociétés d'Assurance

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Géolocalisation du risque inondation	- Utilisation des données de la MRN - Utilisation du logiciel MapInfo pour géolocaliser les risques	- Les taux de surfaces inondables par IRIS délimitent nettement les cours d'eau - Les arrêtés de catastrophes naturelles sont étroitement liés à la présence de surfaces inondables. - Le coût du risque sécheresse est plus important que celui de l'inondation.	Géolocalisation du risque sécheresse

Face à ces derniers propos, il est présenté dans le chapitre suivant une cartographie du risque sécheresse (2^{ème} charge sinistre dans le régime « Cat. Nat. » en terme de charges sinistres).

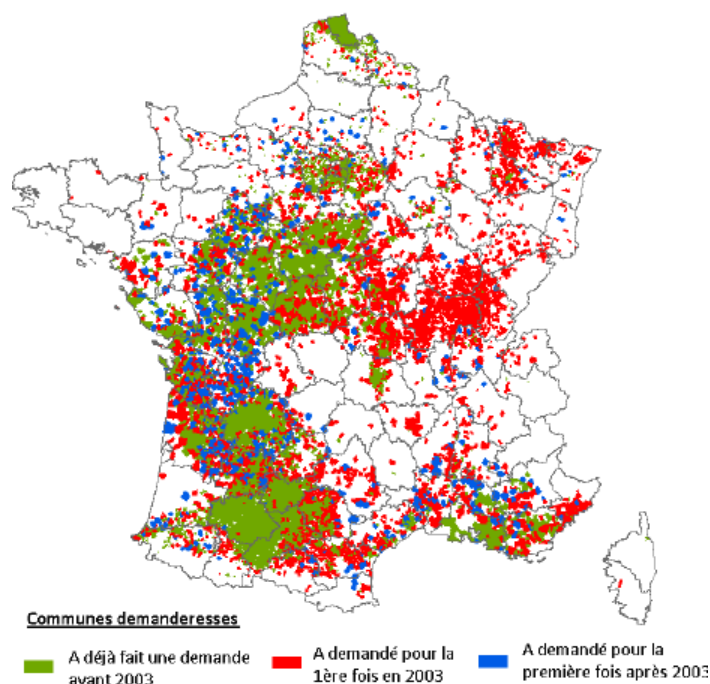
(b) Géolocalisation du risque sécheresse

A fin 2010⁷, l'évaluation du coût de la sécheresse atteignait 4,9 milliards d'euros, ce qui représente 31% du total des indemnités versées (hors effet subsidence) par le régime catastrophes naturelles depuis sa mise en place en 1982.

La géolocalisation du risque sécheresse peut sembler plus erratique que la géolocalisation du risque inondation puisque celle-ci dépend de la pluviométrie au cours d'une année.

Des travaux sont en cours d'élaboration au sein de la MRN, afin d'obtenir des modèles aussi pertinents qu'en inondation. Il ne sera donc pas présenté les résultats proposés.

Lors de la « journée CCR Catastrophes Naturelles » (3^{ème} édition), une cartographie très intéressante a été présentée :



Graphe 4. 6. Carte de cumul des demandes d'arrêtés de catastrophes naturelles « sécheresse » en France de 1982 à 2011

Constats :

- Avant 2003, la plus part des demandes d'arrêtés « Cat. Nat. » sécheresses avait été réalisée sur la partie Est du pays. Plus précisément, une concentration est visible sur :
 - les départements du Nord, du Var et des Bouches du Rhône ;
 - les régions : Ile de France, Centre, Midi-Pyrénées, Limousin et Poitou-Charentes.
- Lors de la canicule de 2003, ces demandes se sont quasiment étendues à l'ensemble du territoire ;
- Depuis 2003, le phénomène a progressé et il est possible de constater une extension des demandes sur le reste du pays.

Afin d'étayer un peu plus ces propos, le ministère de l'Ecologie vient de mettre en ligne sur son site Internet⁸ le résultat de travaux de recherches scientifiques permettant d'apprécier localement le dérèglement climatique et son impact sur la métropole.

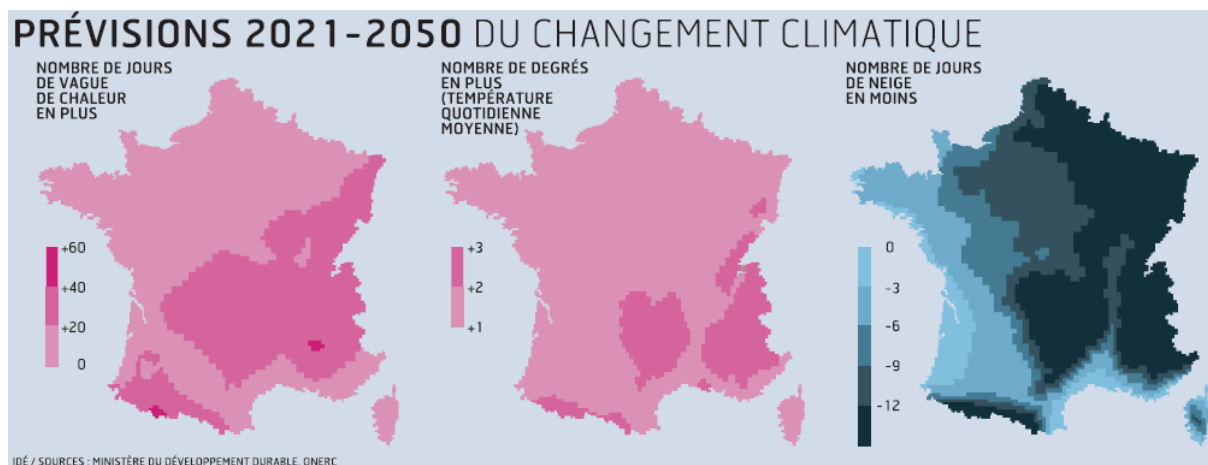
Les projections climatiques par région, basées sur les travaux du GIEC (Groupe d'experts Intergouvernemental sur l'Evolution du Climat) ont une résolution de 8 km sur la France métropolitaine.

⁷ Chiffres tirés de l'article *L'assurance des catastrophes naturelles en 2010 et premières tendances 2011*, réalisé par l'Association Française de l'Assurance, étude Mai 2012.

⁸ Site internet : www.developpementdurable.gouv.fr/-Le-climat-de-la-France-au-XXIe-.html

Les deux rapports publiés par l'Onerc (Observatoire national sur les effets du réchauffement climatique) proposent trois scénarios de réchauffement :

- hausse de 1,4 à 1,9 degrés des températures d'ici à la fin du siècle, pour le plus optimiste,
- de 2,2 à 2,7 degrés, pour le plus mitigé,
- et de 2,9 à 3,5 degrés, pour la vision plus pessimiste.



Graph 4. 7. Prévisions sur le changement climatique en France sur les années 2021-2050

Constats :

- Pour ce qui est du nombre de vagues de chaleur sur les années 2021-2050, le GIEC envisage pour le Limousin, l'Auvergne, la région Rhône-Alpes et le bassin pyrénéen des épisodes beaucoup plus nombreux.
- La hausse des températures pour ces mêmes régions est estimée à plus de 2 degrés.
- Pour ce qui est du nombre de jours de neige en moins, les régions qui étaient jusqu'à présent propices à ces précipitations, le seront beaucoup moins (-12 jours, par exemple, sur les Alpes et les Pyrénées). Cette dernière information n'est pas sans conséquence. En effet, il y a tout un pan de l'économie qui risque d'être affecté, notamment les stations de ski, mais aussi certains assureurs qui pourraient couvrir ces temps de non enneigement dans des garanties vacances.

En conclusion, les projections climatiques confirment :

- une expansion spatiale du phénomène sécheresse en France Métropolitaine ;
- une augmentation de la fréquence et de la sévérité des sécheresses géotechniques.

Avec ce genre de travaux, l'appréciation du risque « sécheresse » pourra être de plus en plus efficace, même si des interrogations peuvent persister (à savoir quel est l'impact de sécheresses printanières ou automnales sur le bâti ? Ou bien encore quel est l'impact de sécheresses successives ?)

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Géolocalisation du risque sécheresse	- Documentation MRN, puis utilisation prochaine des données après mises à jour - Utilisation des prévisions du GIEC sur les effets du changement climatique	- Une concentration d'arrêtés visibles tout d'abord sur : - les départements du Nord, du Var et des Bouches du Rhône ; - les régions : Ile de France, Centre, Midi-Pyrénées, Limousin et Poitou-Charentes. - Le phénomène progresse et les demandes touchent l'ensemble du territoire. - Les perspectives 2021-2050 confirment les constats remarqués sur les demandes d'arrêtés.	Géolocalisation des mouvements de terrain

Le risque « sécheresse » est un risque étroitement lié au risque de « subsidence » (ce phénomène est d'ailleurs intégré dans les 41% de charges sinistres imputables au risque « sécheresse »). Pour rappel, le risque de « subsidence » correspond à un différentiel des sols consécutif à une sécheresse puis à une réhydratation des sols.

Dans le point qui suit, il sera présenté, dans une approche plus générale, la géolocalisation des mouvements de terrains.

(c) Géolocalisation des mouvements de terrains

Les mouvements de terrain sont des phénomènes régis par la force que la gravité terrestre exerce sur des sols instables. Ils peuvent être d'origine naturelle (gel/dégel, séisme, fortes précipitations ou chutes de neige, inondations, etc.) ou anthropique (déboisement, terrassement, etc.). Ils sont classés selon leurs caractéristiques dynamiques comme suit :

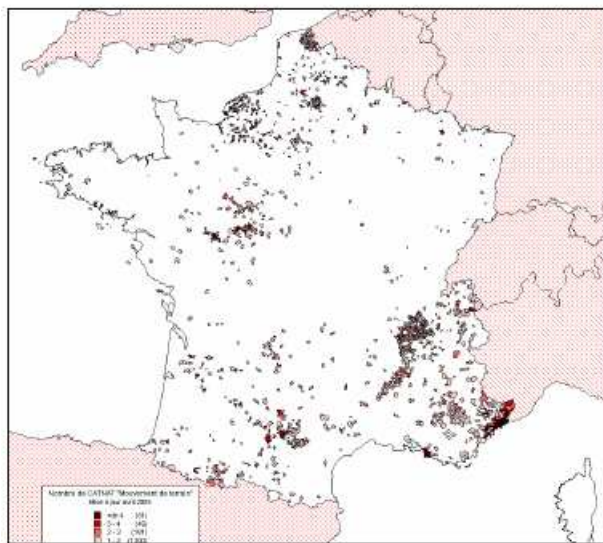
Phénomènes dus à un processus lent et continu	Phénomènes dus à un processus rapide et discontinu
- Phénomènes de gonflement ou de tassement des sols liés aux changements d'humidité des sols argileux (effets de la sécheresse sur le bâti ou sécheresse géotechnique) - Modifications de berges de cours d'eau, érosion du littoral - Avancées de dune	- Affaissements ou effondrements plus ou moins brutaux de cavités souterraines naturelles ou artificielles (mines, carrières) - Glissements de terrain par rupture de versant instable - Coulées boueuses et torrentielles - Ecoulements et chutes de blocs

Tableau 4. 1. Classement des mouvements de terrain selon leurs caractéristiques dynamiques

Les conséquences de tels mouvements sont d'ordres économiques, humaines et environnementales. Cependant, seuls les mouvements rapides peuvent mettre en péril les vies humaines.

Les mouvements de terrain de grande ampleur peuvent drainer des arbres, des rochers ou encore des voitures et être à l'origine de phénomènes induits tels des inondations ou un tsunami ayant pour origine les chutes de masses de terrains dans l'eau par exemple.

La MRN poursuit des travaux sur ce type de risque, tout comme pour la « sécheresse » et n’a pu fournir de vision récente. Néanmoins, il est extrait du rapport *Mémento pratique du particulier Risque « mouvements de terrain »*, une cartographie du cumul des arrêtés de catastrophes naturelles « mouvement de terrain » depuis 1982.



Graph 4. 8. Carte de cumul d'arrêtés de catastrophe naturelle
« mouvement de terrain » en France de 1982 à 2005

Constats :

Même si les arrêtés semblent un peu plus épars, certains bassins se dégagent dans la localisation des arrêtés « Cat. Nat. ». Une concentration d'arrêté est visible sur :

- les départements du Nord, des Alpes-Maritimes et du Tarn ;
- la région Rhône-Alpes.

Les dommages induits suite au retrait et au gonflement des argiles (dues à la sécheresse et la réhydratation des sols) devraient encore augmenter à l'avenir sous l'effet du changement climatique.

Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Géolocalisation des mouvements de terrain	- Documentation MRN, puis utilisation prochaine des données après mises à jour	- Une concentration d'arrêtés visibles sur : - les départements du Nord, des Alpes-Maritimes et du Tarn, - la région Rhône-Alpes. - Une progression du risque de subsidence due au dérèglement climatique	Géolocalisation du risque sismique

En reprenant la définition donnée du mouvement de terrain, à savoir « force [...] exercée sur des éléments de terrain », il peut sembler naturel de poursuivre cette définition en s'intéressant aux événements sismiques, un risque relativement négligé en France.

(d) Géolocalisation du risque sismique

Extrait du « 20 Minutes » du 28 février 2012 :

« "Le sud-est de la France touché par des secousses le 26 février 2012 à 23h37". Un séisme qualifié de "fort" par le Bureau Central Sismologique Français (BCSF) et dont la magnitude a atteint 4,9 sur l'échelle de Richter. L'épicentre était situé à 10km sous terre, au nord-est de Barcelonnette, dans les Alpes-de-Hautes-Provence. Aucun blessé, ni dégâts n'a été constaté, hormis quelques bâtiments jugés "vulnérables". Une réplique de magnitude plus faible (3,7) a suivi deux minutes après la première secousse ».

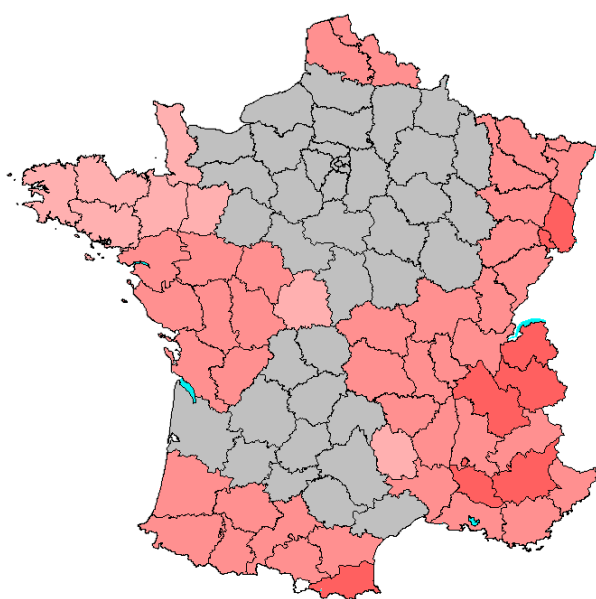


Graphe 4. 9. Localisation du séisme du 26 février 2012

Voici un aspect peu connu par la population et probablement négligé, aussi bien par les Pouvoirs publics que par les assureurs :

En France, le Réseau National de Surveillance Sismique (RéNaSS) comptabilise chaque année plus de 1 500 tremblements dont la magnitude est supérieure à 2,0. Seulement 3 ou 4 de ces séismes se situent dans une magnitude comprise entre 4 et 5 sur l'échelle de Richter.

La MRN a également réalisé et fourni une base permettant de cartographier le risque « sismique ». Ces géolocalisations restent relativement « imprécises » puisqu'elles sont exprimées au niveau départemental.



Alea sismique au niveau IRIS

Très faible < 0,7 m/s ²	(22607)
Faible 0,7 à 1,1 m/s ²	(3227)
Modéré 1,1 à 1,6 m/s ²	(21362)
Moyen 1,6 à 3,0 m/s ²	(2934)

Constats :

- Globalement, les départements frontaliers sont des régions sensibles au risque sismique (en particulier, les régions Rhône-Alpes et Provence-Alpes Côtés d'Azur et le bassin Pyrénéen).
- L'ancien massif armoricain est également sensible, mais dans de moindres proportions.

Graphe 4. 10. Cartographie des départements selon leur aléa sismique

Le Bureau Central Sismologique Français (BCSF) localisé à Strasbourg a mis en place une échelle d'intensité macrosismique. Cette échelle permet de percevoir la « gravité » des séismes.

Echelle d'intensités macrosismiques (EMS-98)

Intensités EMS98	I	II	III	IV	V	VI	VII	VIII	IX	X+
Dégâts potentiels bâtiments vulnérables	aucun	aucun	aucun	aucun	très légers	modérés	quelques effondrements partiels	nombreux effondrements partiels	nombreux effondrements	effondrements généralisés
Dégâts potentiels bâtiments peu vulnérables	aucun	aucun	aucun	aucun	aucun	aucun	très légers	modérés	effondrements partiels	nombreux effondrements
Perception humaine	non ressenti	très faible	faible	modéré	forte	brutale	très brutale	sévère	violente	extrême

Tableau 4. 2. Echelle d'intensité macrosismique définie par le Bureau Central Sismologique Français

Constats :

L'échelle d'intensité est classée en 10 niveaux et déterminée selon la gravité :

- des dégâts subis sur les bâtiments vulnérables,
- des dégâts subis sur les bâtiments peu vulnérables,
- de la perception du séisme par les habitants.

L'évènement du 26 février dernier a ainsi été catalogué entre le niveau IV et V de l'échelle d'intensité.

Le risque sismique étant un risque de nature extrême, rare et localisé à une cité lorsque celui-ci se produit (hors effet tsunami), il semblerait tout à fait naturel que ce type de risque soit mutualisé à un niveau national (principe de solidarité).

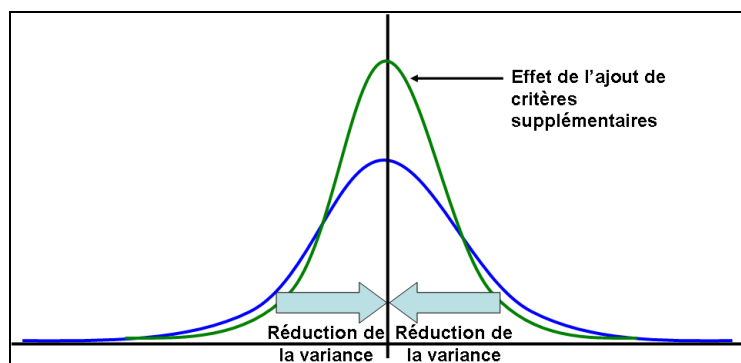
Rappel objectif	Méthodologies employées	Points retenus	Étape suivante
Géolocalisation du risque sismique	- Utilisation des données de la MRN - Utilisation du logiciel MapInfo pour géolocaliser les risques	- Globalement, les départements frontaliers sont des régions sensibles au risque sismique (en particulier, les régions Rhône-Alpes, Provence-Alpes Côtés d'Azur et le bassin Pyrénéen). - L'ancien massif armoricain est également sensible, mais dans de moindres proportions. - Une mutualisation au niveau national est préférable pour ce type de risque	Introduction de variables supplémentaires pour mieux appréhender le risque

Ces cartographies des risques naturels et l'intérêt de mettre un critère de géolocalisation dans une tarification (ou sélection) étant à présent réalisés, il convient d'instruire la réflexion plus loin, notamment dans l'intégration de variables externes supplémentaires au modèle tarifaire actuel.

IV.3.2. Introduction de variables supplémentaires pour mieux appréhender le risque

L'introduction de variables supplémentaires est une manière d'appréhender le risque plus finement, à condition toutefois que ces facteurs soient discriminants et pertinents (et dans ce cas, ne pas se fier uniquement au coefficient de détermination R^2 cf. *Annexe [6.]*).

L'ajout de critères est une piste intéressante qui permet de réduire les variations sur les mesures du risque, quelle que soit la composante considérée (fréquence, coût moyen, et, par effet prime pure). Cette réduction de variance peut être traduite graphiquement comme suit :



Graph 4. 11. Réduction de la variance des résidus par l'ajout de critères supplémentaires

Ces travaux étant confidentiels, il ne sera énuméré que des critères dont le marché a connaissance et qui ne sont pas pris en compte dans l'algorithme actuel du produit Privatis.

Concernant l'habitation :

- La notion d'étage : *a priori*, un aspect relativement pertinent dans l'élaboration d'un tarif, notamment pour un risque d'« inondation ».
- La notion d'habitation principale ou secondaire : hypothétiquement, les habitations secondaires pourraient être moins bien entretenues que des habitations principales. Par conséquent, les probabilités de sinistres pourraient être relativement importantes, engendrant un coût du risque plus élevé pour ces dernières.
- Les matériaux utilisés pour la construction des bâtiments pourraient également être analysés. Derrière ce principe, il y a les notions de robustesse du bâtiment assuré et les coûts de réparations possibles en cas de sinistre (intérêt du projet de loi énoncé au chapitre IV.1.).

Concernant la localisation :

- La localisation du bien assuré est la donnée la plus fiable. De ce fait, tous les indicateurs de géolocalisation doivent être étudiés, en particulier :
 - Les informations (fréquence inondation, sécheresse, ...) liées, au plus fin à l'IRIS, *a minima* au niveau commune INSEE ;
 - Les critères socio-démographiques (nombre d'habitant, taux d'emploi,...).

Concernant le profil de l'entité assurée :

- Des variables selon le profil de l'assuré pourraient être étudiées. En effet, il est possible qu'un certain type de profils soit plus assujéti à déclarer ces sinistres (cf. Chapitre II.1.1. avec les profils de risques ayant une fréquence de sinistres de 22,5%).
- Il pourrait être intéressant d'étudier l'impact de la sinistralité passée dans la sinistralité future (application du modèle de Bühlmann-Straub).

N.B. : Comme il l'a été signalé, la détermination des échelles des zoniers peut être établie sur les résidus des modèles élaborés. Or, avec les diverses informations à disposition, il pourra être mis en concurrence ces données avec la méthode élaborée sur les résidus.

Rappel objectif	Méthodologies employées	Étape suivante
Introduction de variables supplémentaires pour mieux appréhender le risque	Etudier un ensemble de critères internes et externes sur : - l'habitation - la localisation - le profil du souscripteur	La prévention pour réduire les effets des aléas naturels

Jouer uniquement sur les cotisations n'est pas l'unique solution dans une amélioration des résultats techniques et maîtrises des risques. Il est toujours possible d'agir en amont. La prévention est un de ces principes.

Il est proposé ci-dessous les actions de prévention mises en place par les Pouvoirs publics et par Groupama.

IV.4. La prévention pour réduire les effets des aléas naturels

Historiquement, l'implantation de la population et le maintien des activités humaines ont toujours eu lieu à proximité de cours d'eau (malgré les risques de débordements et d'importants dégâts épisodiques). De même, le littoral a toujours été recherché malgré les tempêtes. Il est donc nécessaire d'apprendre à vivre avec ces risques.

Au-delà des pratiques développées par les populations locales, la croissance de l'urbanisation s'est accompagnée d'une démarche de prévention qui s'est formalisée au fil des ans.

Dans le projet de loi enregistré actuellement au Sénat, il est également prévu d'engager des réformes afin de prendre davantage en compte la nécessité de renforcer la prévention et d'encourager les comportements responsables.

Une démarche de prévention des risques a pour but de réduire l'impact des phénomènes naturels prévisibles sur les personnes, les biens et l'environnement. Elle s'articule sur un ensemble de mesures à mettre en œuvre pour réduire les conséquences économiques, sociales et environnementales d'éventuels épisodes catastrophiques⁹.

Face à de tels constats, les Pouvoirs publics ont un rôle primordial à jouer. Il est exposé ci-dessous les actuels dispositifs mis en place et encadrés par les instances publiques.

⁹ Ce chapitre s'inspire de deux sources :

- Le moniteur des travaux publics et du bâtiment, n°5652, 23 mars 2012 ;
- <http://www.eaufrance.fr/>

IV.4.1. Les différents dispositifs des Pouvoirs publics face aux aléas naturels

Les Pouvoirs publics ont un rôle essentiel qui s'appuie sur plusieurs grands axes :

- La réduction de la vulnérabilité,
- La maîtrise de l'urbanisation,
- L'organisation des secours,
- L'information préventive.

Réduction de la vulnérabilité et Maîtrise de l'urbanisation

Les plans de prévention des risques (PPR) ont pour but de maîtriser l'urbanisation dans les zones exposées à un aléa. Etablis par l'État, ils font servitude d'utilité publique et doivent être annexés aux Plans locaux d'urbanisme (PLU) des communes concernées.

Les PPR permettent de délimiter des zones à l'intérieur desquelles des prescriptions peuvent être imposées aux constructions et à l'usage du sol. Les plans de prévention des risques naturels issus de la loi n°95-101 du 2 février 1995, définissent des zones d'interdiction et des zones de prescription.

Organisation des secours

Au niveau national, le Centre Opérationnel de Gestion Interministérielle des Crises (COGIC) est à disposition du gouvernement et du ministre de l'Intérieur.

Ce centre permet une gestion des secours essentiellement pour les départements et en particulier :

- du Service Départemental d'Incendie et de Secours qui sont des établissements publics départementaux autonomes,
- des Services Interministériels de Défense et de Protections Civiles.

Néanmoins, de nombreux autres acteurs participent à la sécurité civile :

- les citoyens,
- les maires,
- les préfets,
- les services opérationnels permanents,
- les services médicaux d'urgence,
- les administrations locales et les ONG.

Pour les citoyens, une loi de modernisation de la sécurité civile 2004-811 du 13 août 2004 prévoit l'organisation de la sécurité civile.

Le plan ORSEC départemental mis en place par les services préfectoraux détermine, compte tenu des risques existant dans le département, l'organisation générale des secours et recense l'ensemble des moyens publics et privés susceptibles d'être mis en œuvre.

Information préventive et « éducation »

Selon l'Article L.125-2 du Code de l'Environnement, la législation française impose une information de la population en matière de risques majeurs. Cette information, qui permet de réduire la vulnérabilité des populations face aux risques, est réalisée au travers de plusieurs documents :

- Le dossier départemental des risques majeurs est établi par le Préfet et appuyé par les services déconcentrés de l'Etat. Il présente l'ensemble des risques majeurs, naturels et technologiques auxquels sont exposées les communes d'un département.
- Le Dossier d'Information Communale sur les Risques Majeurs est établi par le Maire avec son conseil municipal. Ce rapport ajoute des actions de protection ou de sauvegarde pour la commune, les événements significatifs et les dispositions spécifiques à la gestion du risque au niveau communal. Il est librement consultable en mairie.

Enfin, depuis le 1er Juin 2006, les propriétaires vendeurs ou bailleurs ont l'obligation d'informer leurs acquéreurs, ou locataires, sur les risques naturels et technologiques auxquels leur bien immobilier est exposé. De même, les propriétaires doivent informer les sinistres ayant donné lieu à indemnisation au titre de la garantie catastrophe naturelle et technologique.

Rappel objectif	Points retenus	Étape suivante
Les différents dispositifs des Pouvoirs publics face aux aléas naturels	- Les Pouvoirs publics ont mis de nombreux plans et organisations pour faire face à des événements extrêmes - La réduction de la vulnérabilité est un point essentiel pour l'avenir. Mais celui-ci n'est peut être pas encore suffisamment incitatif. Il est important de ne plus édifier des bâtiments en zones à risques.	Les différents dispositifs de Groupama pour accompagner ses assurés

Groupama a été précurseur, il y a plus de 50 ans, dans le domaine de la prévention des risques. Pour citer quelques exemples :

- En matière de sécurité routière, le réseau des « Centres Centaure » rassemble sur l'ensemble du territoire français 12 centres de formation à la conduite ;

- Le Groupe mène également de nombreuses actions dans le domaine de la prévention agricole, comme par exemple l'opération « Dix de conduite rurale » qui consiste à former les jeunes des lycées agricoles à la conduite d'engins spécifiques au métier ;
- En matière de santé, de nombreuses prestations sont proposées aux assurés, dont un site Internet dédié à l'alimentation et la santé.

IV.4.2. Accompagnement des assurés face aux risques naturels

Groupama a mis en place depuis 2011, un dispositif d'alerte pour les communes assurées au sein de son portefeuille.

Si un évènement naturel est perçu comme sensible et si celui-ci est de nature à impacter fortement la commune, l'outil Predict informe par anticipation les élus et les gestionnaires identifiés. Suite à cette alerte, les personnes suivent un protocole établi.

Depuis son lancement 2011, cet outil de prévention a permis d'avertir à plusieurs reprises différentes entités assurées contre des évènements naturels risqués.



Graph 4. 12. Portal d'accueil de Groupama Predict 

Exemple de l'évènement pluvieux-orageux du samedi 28 au dimanche 29 avril 2012

- Le samedi 28 avril 2012 à 12h, Météo France place 9 départements en vigilance « orange : VENT VIOLENT ».

Météo-France prévoit des vents possibles atteignant les 90 km/h en vallées et jusqu'à 150km/h sur les sommets. Il s'agit des départements suivants : 07, 15, 26, 38, 42, 43, 69, 73 et 74. Cette vigilance est relayée aux communes par les préfetures.

- Suite à cette vigilance orange annoncé par Météo France, l'outil Predict informe les parties concernées par des messages d'alertes (sms, mail, téléphone et portail) de vents forts, et alimente en « info-risques » les zones exposées aux vents les plus violents.

- Dès le samedi, une dépression (nommée « QUEENIE ») centrée sur le pays génère un important flux du sud depuis la Méditerranée jusque sur le centre-est du pays. Des vents forts soufflent sur l'est et sont plus importants sur les reliefs (Alpes, Massif central). Dimanche, cette dépression se décale vers le nord-ouest et génère d'importants orages entre Toulouse et Aurillac.

Le dimanche 29/04, une ligne de cellules orageuses, donnant des précipitations très intenses et de la grêle se met en place entre Toulouse et Aurillac. Lors de cet épisode soutenu, une tornade se matérialise également aux abords de Toulouse.

Predict alimente en continu et en info-risques les communes alors exposées.

Constats :

- Un suivi précis et permanent au travers de « Wiki Predict » a pu être réalisé. L'espace client a été alimenté et tenu à jour en temps réel 24h/24h.

- Grâce au suivi et à l'analyse localisée, Predict Services a informé (par messages successifs) les 64 communes concernées par les orages violents de la soirée du 29/04, pour risque de ruissellement.

Face à ces messages, il est malheureusement difficile d'estimer les « gains » qui ont pu être réalisés. En effet, réduire les accès de certains lieux, rehausser ou déplacer du mobilier, ... sont des aspects difficilement quantifiables.

Rappel objectif	Points retenus
Les différents dispositifs de Groupama pour accompagner ses assurés	- Les élus et les gestionnaires identifiés sont informés lors d'évènements atypiques susceptibles d'impacter fortement des localités - Difficultés à mesurer les « gains » réalisés sur des évènements atypiques

Point de transition :

La probable réforme du système « Cat. Nat. » laisse présumer des désengagements de l'Etat, ce qui devrait avoir un impact direct sur la prise en charge des assureurs. Par conséquent, la réflexion qui a été réalisée jusqu'à présent sur la garantie « événements climatiques » a été portée plus loin dans cette section.

Depuis quelques années, le marché de la réassurance subit un phénomène de concentration assez important. De par les dispositions énoncées dans le projet de loi et les désengagements de l'Etat dictés sur certains périls, les réassureurs seront vraisemblablement enclins à changer les modalités des couvertures actuellement proposées.

Au regard de ces probables impacts, les nouvelles possibilités de couverture que sont les dérivés climatiques vont certainement prendre un nouvel essor. Néanmoins, pour répondre aux clauses du dérivé, il sera nécessaire de « dégager » plus de rentabilité. Dans cette optique, des augmentations tarifaires devront être réalisées, ce qui est un relatif inconvénient dans un marché de l'assurance relativement tendu.

Ce projet de loi laisse toutefois une perspective intéressante : le taux actuellement appliqué de 12% sur le risque MRH pourra probablement à l'avenir être revu et modulé selon le risque (comme il est actuellement envisagé pour les collectivités locales et les entreprises).

A cette nouvelle flexibilité, il pourrait être envisagé de décomposer tarifairement les aléas naturels dans leur globalité, en créant des garanties distinctes propres. Par exemple, des garanties libellées :

- « risques météorologiques » couvrant le risque tempête (vents violents), les risques grêles et neiges pourraient peut-être faire l'objet là aussi d'une tarification distincte ;
- « risques climatiques » couvrant le risque de sécheresse et de subsidence ;
- « risques d'inondation » couvrant les débordements, les ruissellements et les submersions.

En parallèle à cette décomposition, il pourrait également être pertinent de revoir la détermination des échelles de risque par l'intégration de variables supplémentaires. L'introduction de nouveaux critères est un moyen de rendre l'algorithme tarifaire plus adéquat ou du moins, d'apporter un tarif plus segmenté (et donc plus dissuasif et « moralisateur »).

Comme il l'a déjà été noté, l'aspect géolocalisation n'est paradoxalement pas pris en compte dans l'algorithme tarifaire aujourd'hui. Or, si une chose est certaine dans un risque Habitation, c'est sa localisation. La création d'un zonier propre à chaque risque couvert pourrait être mise en place.

A cela, il saurait s'ajouter tous les indicateurs « géolocalisés » susceptibles d'être pertinents dans l'appréciation du risque. Par exemple, le critère IRIS et les caractéristiques propres à cette variable, telles que les informations :

- socio-démographiques de la région,
- météorologiques,
- climatiques,
- de présence d'argile dans les sols.

L'ajout de critères spécifiques à l'habitation est également une piste intéressante, avec :

- la notion d'étage,
- le type d'habitat (principal ou secondaire),
- les matériaux utilisés dans la construction,
- la date de construction.

Les variations de la mesure du risque, quelle que soit la composante considérée (fréquence, coût moyen, et, par effet prime pure) pourraient être ainsi réduites.

Mais « jouer » uniquement sur les cotisations n'est pas l'unique solution dans une amélioration des résultats techniques et maîtrises des risques. Il est toujours possible d'agir en amont. La prévention est un de ces principes. Les actions de prévention mises en place par les Pouvoirs publics et par Groupama sont des exemples concrets de maîtrises « probables » des coûts (l'aspect gain étant relativement difficile à chiffrer).

Conclusion

L'impact des récents événements naturels dans les résultats du Groupe et la montée en puissance des modèles actuariels (en lien avec la nouvelle réglementation Solvabilité 2) ont permis d'engager une réflexion sur les produits Multi-Risques Habitation commercialisés à ce jour, notamment dans l'approche tarifaire de la garantie « événement climatique ».

Les incertitudes liées au changement climatique de la planète et le manque de profondeur d'historique rendent délicate l'approche tarifaire classique. Ainsi, une connaissance même approfondie du passé ne laisse pas supposer un comportement identique dans les années à venir.

Afin de répondre à cette problématique, la tarification proposée dans ce rapport a été élaborée en deux temps :

- des échelles de risques ont tout d'abord été modélisées, sur les bases de données historiques, via des techniques de régression utilisées classiquement en tarification non vie ;
- le niveau de la prime pure a ensuite été calibré à partir des modélisations « catastrophes » du modèle interne sur lesquelles s'appuie le besoin de fonds propres économiques au quantile 1/200.

Le chargement de réassurance est directement issu de ces modélisations en simulant la charge probable cédée en réassurance via des tirages de Monte-Carlo.

Aux vues des différents éléments apportés dans ce rapport, cette nouvelle approche tarifaire semble être une méthodologie adaptée à ce risque erratique.

De plus, l'intégration du coût du capital et les effets de diversifications des risques sont des notions qui ont également été traitées dans ce rapport. La performance des métiers ne se base plus uniquement sur les marges techniques nettes réalisées. Elle est parallèlement analysée au regard du besoin de fonds propres économiques.

En vision économique, des problématiques d'allocation de gains de diversification apparaissent : différents modèles d'allocations de capitaux s'immiscent donc progressivement dans les organismes assurantiels.

Ces questions sont essentielles et peuvent conduire à différentes allocations de fonds propres par métier. Ces modèles économiques de mesure de la rentabilité au regard des fonds propres économiques alloués déformeront probablement l'offre du marché dans le temps.

L'approche pluriannuelle demandée, dans le cadre de l'ORSA, constitue donc un axe de développement intéressant pour de futurs travaux.

Afin de bénéficier des effets de mutualisation des risques souscrits, il est également opportun d'étudier l'ajout de critères tarifaires à l'actuel algorithme. Ces travaux sont en cours de réalisation et constituent là aussi, un prolongement à cette étude.

« When written in Chinese the word crisis is composed to two characters: one represents danger, and the other represents opportunity », John Fitzgerald Kennedy.

Dans l'écriture chinoise, le mot « crise » est composé de deux caractères : l'un représente le danger, l'autre représente l'occasion.

Bibliographie

Articles

Association Française de l'Assurance (2012) *L'assurance des catastrophes naturelles en 2010 et premières tendances 2011*, Etude publiée en collaboration avec la FFSA et le GEMA en mai 2012

CADOUX David, LOIZEAU Jean-Marc (2004) «*Copules et dépendances : application pratique à la détermination du besoin en fonds propre d'un assureur non vie.*», Bulletin français d'actuariat n°12, volume 6, Janvier 2004 - Juin 2004, pp. 173-208

CHARPENTIER Arthur (2002), *Titrisation des risques catastrophes : les Cat Bonds.*, FFSA – Direction Etudes, Statistiques & Systèmes d'information

CONFAIS Josiane, GRELET Yvette, LE GUEN Monique (2005) *La procédure FREQ de SAS : Tests d'indépendance et mesures d'association dans un tableau de contingence*, Revue Modulad, pp. 209-212

DERIEN Anthony (2010), *L'horizon temporel dans Solvabilité 2*, Bulletin français d'actuariat n°19, volume 10, janvier - juin 2010, pp. 43-62

Fédération Française Des Sociétés d'Assurances (FFSA, 2009), *Synthèse de l'étude relative à l'impact du changement climatique et de l'aménagement du territoire sur la survenance d'événements naturels en France*, avril 2009, pp. 4-5

FOULQUIER Philippe (2009) *Solvabilité II : une opportunité de pilotage de la performance des sociétés d'assurance*, Publication EDHEC, pp. 22-27

GIEC (2007), *Changements Climatiques 2007, Rapport de synthèse*, publié par le Groupe d'experts intergouvernemental sur l'évolution du climat, pp. 2-22

KAAS Rob (2010), *Compound Poisson Distributions and GLM's, Tweedie's distribution*, Department of Quantitative Economics, Université d'Amsterdam

L'Argus de l'assurance (hebdomadaire n°7241 du 18 novembre 2011), Publication Groupe Industrie Services Info, pp. 46-47

Les Echos (20 & 21 mars 2012), *Réchauffement climatique : les élus savent désormais à quoi s'en tenir dans leurs territoires*, n° 21156, p. 7

MAGNET J. C. (2012), *Le sud-est touché par un séisme*, extrait du journal 20 Minutes, 28 février 2012, pp. 7-8

Mission Risques Naturels (MRN, 2005), *Mémento pratique du particulier Risque « mouvements de terrain »*, Mission des sociétés d'assurances pour la connaissance et la prévention des risques naturels, pp. 6-7

PERRUCHOT Christiane (2012), *Le moniteur des travaux publics et du bâtiment*, n°5652, 23 mars 2012, pp. 3-10

SENAT (enregistré à la Présidence du Sénat le 3 avril 2012), *Projet de Loi portant réforme du régime d'indemnisation des catastrophes naturelles*, n°491, pp. 3-40

SHONO Hiroshi (2010), *Application of the Tweedie distribution to zero-catch data in CPUE analysis*, SCIENTIFIC COMMITTEE SIXTH REGULAR SESSION

THERON Pierre (2008) *IFRS, Solvabilité 2, Embedded Value : Quel traitement du risque ?*, Bulletin Français d'Actuariat, Vol. 8, n°15, janvier-juin 2008, pp. 67-96

Polycopiés de Cours

BORCARD Daniel (2008), *Tests de comparaison des variances de plus de deux échantillons indépendants*, Université de Montréal

BORKOWSKI John (2009), *HOV Tests, transformations, weighted least squares*, Université du Montana

CHARPENTIER Arthur (2005), *Cours de séries temporelles théorie et applications*, Université de Beaulieu, Rennes 1

DELYON Bernard (2005), *Régression*, Université de Beaulieu, Rennes 1

DELYON Bernard (2005), *Simulation et modélisation*, Université de Beaulieu, Rennes 1

JAL Pierre (2010), *Réassurance Non-Vie*, ISFA, Université Claude Bernard, Lyon 2

LOISEL Stéphane (2009), *Cours de gestion des risques d'assurances et de théorie de la ruine*, ISFA, Université Claude Bernard, Lyon 2

MARCOTTE Denis (2011), *Cours de Géostatistique*, Ecole Polytechnique de Montréal

ROUVIERE Laurent (2009), *Régression sur variables catégorielles*, Université de Rennes 2

THEROND Pierre (2010), *Théorie de la Crédibilité*, ISFA, Université Claude Bernard, Lyon 2

Mémoires

ANDRIANTAVY C. (2008), *Tempêtes : prise en compte de la dépendance spatiale avec la théorie des couples*, GROUPAMA, formation Université Paris Dauphine

BELGUISE O. (2001), *Tempêtes : étude des dépendances entre les branches auto et incendie avec la théorie des copulas*, Guy CARPENTER, formation Université Louis Pasteur Strasbourg

DESCHEEMAEEKERE D., PERRON B. (1999), *Les événements naturels en France : étude du risque tempête pour une compagnie d'assurances*, UAP, formation CEA

FINAS Bernard, GILLES Sébastien (2011), *La gestion indicielle du risque climatique, Application à l'énergie éolienne*, AON, formation ENSAE

GUETTE V. (2010), *La prise en compte des catastrophes dans la modélisation de la mortalité*, WINTER & Associés, formation ISFA

MILOE Y. (2011), *Tarification d'un produit MRH à l'aide de la méthode des marges*, Axéria IARD, formation ISFA

SAUSER Carine, GROISNE Mélanie, MILHAUD Xavier (2007), *Projet Etienne Marceau : Méthodes statistiques en assurance non-vie*, formation ISFA.

SIBAI D. (2004), *Modélisation dynamique des catastrophes naturelles*, formation ENSAE

VEILLEUX Lucie (2005), *Modélisation de la trajectoire criminelle de jeunes contrevenants à l'aide de modèles linéaires généralisés mixtes*, formation Université de Laval, Québec

Ouvrages

AUGROS Jean-Claude (2010), *Gérer les risques autrement*. RB Edition

BROWN Lawrence (1986), *Fundamentals of statistical exponential families with applications in statistical decision theory*, Inst. of Math. Statist., Hayward, California

CEIOPS, COMMISSION EUROPEENNE (2010), *QIS5 Technical Specifications*, 5 juillet 2010, et, *Errata to the QIS5 Technical Specifications*, 27 septembre 2010

DENUIT Michel, CHARPENTIER Arthur (2004), *Mathématiques de l'assurance non-vie Tome I : Principes fondamentaux de théorie du risque et Tome II : Tarification et provisionnement*, Economica

DENUIT Michel, MARECHAL Xavier, PITREBOIS Sandra, WALHIN Jean-François (2007), *Actuarial Modelling of Claim Counts, Risk Classification, Credibility and Bonus-Malus Systems*, Edition John Wiley & Sons Ltd

HULL John (2007), *Options futures et autres actifs dérivés*, 6ième édition, Pearson Education

Lawrence Brown (1986), *Fundamentals of statistical exponential families with applications in statistical decision theory*, Edition Series, Shanti S. Gupta.,

LUZI Michel (2006), *Assurance IARD, Interprétation des chiffres*, Economica

PARTRAT Christian, BESSON Jean-Luc (2005), *Assurance Non-Vie Modélisation, Simulation*, Economica

ROBERT Christian (2006), *Le choix bayésien: principes et pratique*, Springer

SAPORTA Gilbert (2006), *Probabilité, Analyse des données et Statistique*, Edition Technip

Présentations

BLANCHARD Michelle, AJI Salaheddine (28 juin 2012), *L'exposition actuelle et future au péril sécheresse en France métropolitaine*, 3ème édition de la journée CCR

JOUBERT Jean (Conférence du 8 novembre 2011), *Les évènements climatiques*, organisé par « l'Argus de l'Assurance ».

NOUY Danièle, RICHARD Philippe, DREYFUSS Marie-Laure, FLICHE Olivier, DUSSUTOIR Jean-Jacques, REVELIN David, AURAN Thierry, ROUX Cyril (7 octobre 2011), *Solvabilité 2 et Pilier 3 vers une nouvelle information prudentielle*, Présentation ACP

Price Water Cooper (2011), *Weather Risk Derivative Survey Prepared for the Weather Risk Management Association May 2011*, Survey Results, PwC

RASPILLER Sébastien (28 juin 2012), *Présentation du projet de réforme du régime d'indemnisation des Catastrophes Naturelles*, Direction Générale du Trésor

VESIN Aurélien (2007), *La régression de Poisson, Théorie et Applications SAS*, INSERM

Sites Internet

- Association des Risques Naturels : <http://www.mrn.asso.fr/>
- Autorité de Contrôle des Assurances et des Mutuelles : http://www.acam-france.fr/modeles_internes/
- Centre Interprofessionnel Technique d'Etudes de la Pollution Atmosphérique : <http://www.citepa.org/>
- Cours tutoriaux de statistiques : http://www.aiaccess.net/y_tut_liste.htm
- Cours de SAS sous forme de polycopier : http://www.math.jussieu.fr/~goutte/index_fichiers/AutresEnseignement/Master_2_Info/
- Direction générale de la prévention des risques / Ministère du Développement Durable en partenariat avec « Réseau Idéal, Sciences Frontières Production » et l'Institut français des formateurs risques majeurs éducation : <http://www.Primnet.fr/>
- Fédération Française du Bâtiment : <http://www.ffbatiment.fr/>
- Institut National de la Statistique et des Etudes Economiques : <http://www.insee.fr/>
- Organisation Météorologique Mondiale : <http://www.wmo.int/>
- Service Public du Développement Durable : www.developpementdurable.gouv.fr
- Service Public de la diffusion du Droit : <http://www.legifrance.gouv.fr/>
- Service Public sur la veille hydrographique : <http://www.eaufrance.fr/>
- Service Public sur les séismes (BCSF) : <http://www.franceseisme.fr/>
- Weather Risk Management Association: http://www.wrma.org/members_survey.html

Annexes

Annexe [1.] : Les constats environnementaux au niveau mondial

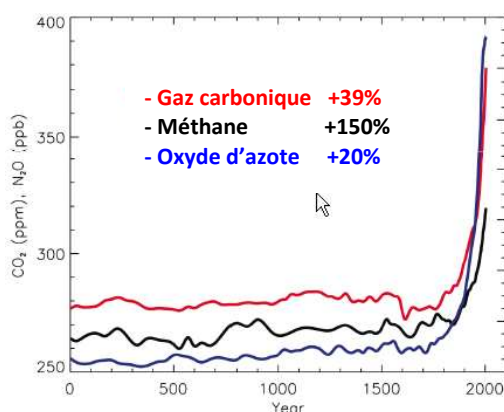
Le Groupe d'experts Intergouvernemental sur l'Evolution du Climat (GIEC) a été créé en novembre 1988, à la demande du G7, par deux organismes de l'ONU : l'Organisation Météorologique Mondiale (OMM) et le Programme des Nations Unies pour l'Environnement (PNUE).

Le GIEC est un organe intergouvernemental ouvert à tous les pays membres de l'ONU. Il « a pour mission d'évaluer, sans parti-pris et de façon méthodique, claire et objective, les informations d'ordre scientifique, technique et socio-économique qui sont nécessaires à une meilleure compréhension des risques liés au changement climatique d'origine humaine, de cerner plus précisément les conséquences possibles de ce changement et d'envisager d'éventuelles stratégies d'adaptation et d'atténuation »¹.

Lors d'une conférence introduite par Jean Joubert, 3 certitudes ont été révélées² :

- Premièrement, les activités humaines modifient la composition de l'atmosphère en gaz à effet de serre ;
- Deuxièmement, en 100 ans, il a été observé un réchauffement de près de 1°C ;
- Enfin, troisième et dernière certitude, le réchauffement climatique va se poursuivre.

Les pollutions atmosphériques



Constats :

- Avant la révolution industrielle, la concentration dans l'atmosphère de CO₂ était de 280 ppm.
- Aujourd'hui, la concentration dans l'atmosphère est de 390 ppm, soit une croissance de 2 ppm par an.

Pour les climatologues³, à ce rythme, dans 30 ans, une concentration de 450 ppm combinée aux autres gaz à effet de serre pourrait engendrer des modifications climatiques de grandes ampleurs.

Graphique A. 1. Evolution de la concentration de CO₂ en ppm (partie par million), du méthane (CH₄) et de l'oxyde d'azote (N₂O) en ppb (partie par billion ~ milliard, équivalent français)

Les gaz présentés ci-dessus sont les trois principaux gaz à effet de serre. Leur impact et leur durée de vie dans l'atmosphère sont relativement différents :

- Le dioxyde de carbone (CO₂) reste dans l'atmosphère près de 100 ans,
- Le méthane (CH₄) reste une douzaine d'années dans l'atmosphère,
- L'oxyde d'azote (N₂O) a une durée de vie de 150 ans.

Afin de savoir quel gaz a le plus d'impact sur le climat, les scientifiques raisonnent en « équivalent CO₂ » ; Ainsi, l'effet de chaque gaz est exprimé en fonction de l'effet du CO₂ (qui vaut 1 par définition) avec le calcul suivant :

$$\text{Tonne du gaz} \times \text{PRG du gaz} = \text{tonne d'équivalent CO}_2 \text{ d'un gaz.}$$

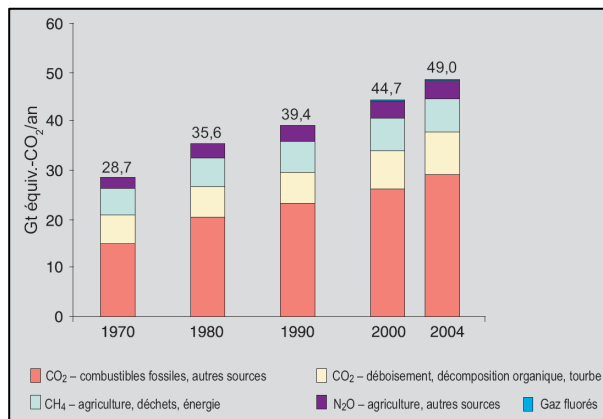
A titre d'exemple⁴, le potentiel de réchauffement global (PRG) du méthane (CH₄) est 21 fois plus puissant que celui du CO₂. L'oxyde d'azote (N₂O) a, quant à lui, un effet 310 fois plus puissant que celui du CO₂.

¹ Extrait du portail du GIEC : http://www.ipcc.ch/home_languages_main_french.shtml

² Conférence du 8 novembre 2011 sur les événements climatiques organisée par « l'Argus de l'Assurance ».

³ Nouveaux travaux ayant conduit à cette nouvelle estimation de concentration

⁴ Extrait du site <http://www.citepa.org/>, Centre Interprofessionnel Technique d'Etudes de la Pollution Atmosphérique



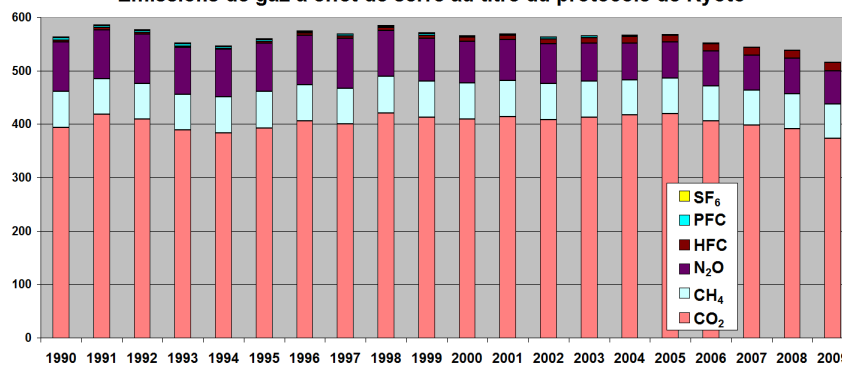
Graph A. 2. Emissions des gaz à effet de serre anthropiques dans le monde de 1970 à 2004 en équivalent CO₂ pondérés selon la CCNUCC⁵

Constat :

Le gaz carbonique (couleurs orange et jaune) a une influence plus importante sur le climat que le méthane (bleu clair) ou l'oxyde d'azote (violet), du point de vue des émissions actuelles.

Ce constat d'augmentation mondiale n'est pas le reflet de toutes les politiques mondiales. En effet, en France, par exemple, il est possible de constater que les rejets en gaz à effet de serre (selon les critères du protocole de Kyoto) ont été réduits entre 1990 et 2009 (une baisse de 8,1% des PRG a été constatée) avec notamment une inflexion observable à partir de 2006 (cf. Graph A.3. ci-dessous).

Emissions de gaz à effet de serre au titre du protocole de Kyoto



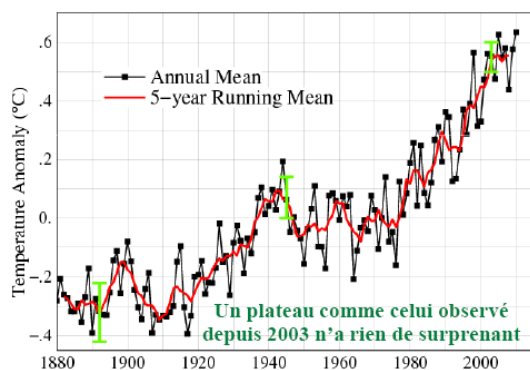
Graph A. 3. Emissions des gaz à effet de serre en France entre 1990 et 2009.

En aparté, dans le cadre d'un modèle prédictif sur l'évolution des températures, ces différentes concentrations de particules dans l'atmosphère pourraient être intégrées et considérées comme variables explicatives (d'un modèle prédictif au niveau mondial).

La corrélation entre l'augmentation des gaz à effet de serre et le réchauffement planétaire semble indéniable, même si certains continuent de réfuter ces faits.

Il existe malgré tout un constat indéniable : les températures actuelles sont les plus élevées jamais enregistrées. Ce constat est présenté dans le point suivant.

Les conséquences environnementales



Constat :

En 100 ans, il a été observé un réchauffement de près de 1°C.

Graph A. 4. Evolution de la température mondiale entre 1880 et 2005.

⁵ CCNUCC : Convention Cadre des Nations Unies sur les Changements Climatiques

Une telle augmentation des températures conduit naturellement à se poser la question des effets produits. Plusieurs impacts ont déjà été observés :

- Une diminution des surfaces des glaciers et glaces de mer. Ces fontes ont conduit à une réévaluation du niveau de la mer qui n'est pas sans conséquence. Lors d'intempéries ou fontes des neiges, l'écoulement des eaux devient difficile. De ce fait, la probabilité d'inondation n'en est que plus forte ;
- Selon les régions, il a été observé de plus grandes périodes de sécheresses ;
- Une acidification des pluies avec des effets sur les matériaux, les écosystèmes forestiers et les écosystèmes d'eau douce (lacs, cours d'eau). De même, l'acidification des mers et océans a également entraîné des modifications (voire destructions) de la vie aquatique.

D'autres manifestations en termes d'évènement naturel existent, c'est ce qui va être approché dans le chapitre suivant.

Les différents types d'évènements naturels

A partir du site internet Prim.net⁶ et de l'application Gaspar⁷ proposée, différents types d'évènements sont énumérés. Seuls les risques naturels métropolitains sont décrits ci-dessous et étayés de quelques chiffres :

Aléa	Quelques chiffres										
Inondation	- 13 000 communes exposées au risque d'inondation - 27 000km ² de zones inondables pour 160 000km de cours d'eau - plus de 6,1 millions d'habitants sont placés en zone inondable sur le territoire Plus de 70 inondations de grandes ampleurs ont été recensées entre 1900 et 2010.										
	Année	1910	1930	1940	1957	1958	1965	1974	1977	1979	1980
	Nb d'inondation recensée	1	1	1	1	1	1	1	1	1	1
	Année	1983	1987	1988	1992	1993	1994	1995	1996	1997	1999
	Nb d'inondation recensée	1	1	1	2	1	2	1	2	2	2
	Année	2000	2001	2002	2003	2005	2006	2007	2008	2009	2010
	Nb d'inondation recensée	4	5	6	5	4	2	1	11	4	3
Mouvement de terrain	Année	Description et localisation									Nombre de décès
	1926	Eboulements et coulées boueuses - Roquebillière (Alpes-Maritimes)									28
	1930	glissement de la colline des Balmes - Lyon – Fourvière, quartier Saint Jean									40
	1932	Lyon – Cours d'Herbouville									30
	1961	Effondrement de carrière, plusieurs maisons englouties - 8ha (hectares) concernés - Clamart (Hauts-de-Seine)									21
	1970	coulée de débris rocheux Roc-des-Fiz - Plateau d'Assy (Haute-Savoie)									72
	1992	Indre									
	1992	Sécheresse géotechnique - plus de 2 milliards d'euro de dommages - Dordogne									
	1996	Yvelines									
	1996	Nord									
Séisme	Année	Localisation								Echelle de Richter	
	1905	Chamonix (Haute-Savoie)								7-8	
	1909	Lambesc, Saint-Cannat, Rognes (Bouches-du-Rhône)								6	
	1913	Volx (Alpes de Haute Provence)								7-8	
	1917	Guadeloupe								8	
	1951	Chasteuil (Alpes de Haute Provence)								7-8	
	1959	Saint-Paul (Alpes de Haute Provence)								7-8	
	1967	Pyrénées-Atlantiques								5,1	
	1980	Ossau, Arudy (Pyrénées-Orientales)								7-8	
	1996	Annecy (Haute-Savoie)								5,2	

⁶ Ce site Web est une initiative de la Direction générale de la prévention des risques / Ministère du Développement Durable en partenariat avec « Réseau Ideal, Sciences Frontières Production » et l'Institut français des formateurs risques majeurs et protection de l'environnement.

⁷ GASPARE : Gestion Assistée des Procédures Administratives relatives aux Risques naturels et technologiques

Aléa	Quelques chiffres ⁸		
Avalanche	Année	Localisation	Nombre de décès
	1917	Valmanya-Roquejalère (Pyrénées Orientales)	12
	1921	Savoie	13
	1923	Savoie	21
	1927	Corse	13
	1934	Ortiporio (Haute-Corse)	37
	1950	Huez (Isère)	10
	1970	Val d'Isère (Savoie)	39
	1999	Montroc - Chamonix (Haute-Savoie)	12
Feu de forêt	En 1949, Forêt landaise – Cestas (Gironde), 140 000 ha détruits En 2003, dans le massif des Maures, cible chronique des incendies, 20 000 hectares ont été livrés aux flammes, soit 20% du massif des Maures et de l'Esterel dévastés. De même, en Corse, 26 000 ha touchés.		
Phénomène lié à l'atmosphère	Année	Localisation	Commentaires
	1969	Bretagne	plus de 10 morts
	1982	Rhône-Alpes	deux tempêtes dans le centre et le sud de la France
	1987	Bretagne	des vents de 150 à 200 km/h, 1/4 de la forêt bretonne est dévastée
	1990	Ile-de-France	série de tempêtes et de pluies diluviennes dans le Nord et Ouest de la France : des vents atteignant parfois 150 km/h arrachent dans le pays l'équivalent de la forêt de Fontainebleau
	1999		2 tempêtes Lothar (26/12) et Martin (27 et 28/12) : 118 millions de m ³ de bois à terre, 92 décès et 15 milliards d'euros de dommages pour les deux tempêtes
	2001	Nord-Pas-de-Calais	Liaisons Eurostar et régionales interrompues, 37 500 foyers privés d'électricité, (31 personnes décédées en Europe)
	2001	Strasbourg	1 200 interventions en Alsace (67/68), dégâts sur voiture
	2003	Métropole	vague de chaleur en métropole : 15 000 victimes, près de 2 milliards d'euro de dommages assurés et près de 4 500 communes reconnues en état de catastrophe naturelle
	2006	Métropole	vague de chaleur en métropole : 2065 victimes

Tableau A. 1. Énumération d'évènements « Cat. Nat. » recueillis à partir de l'application GASPAP

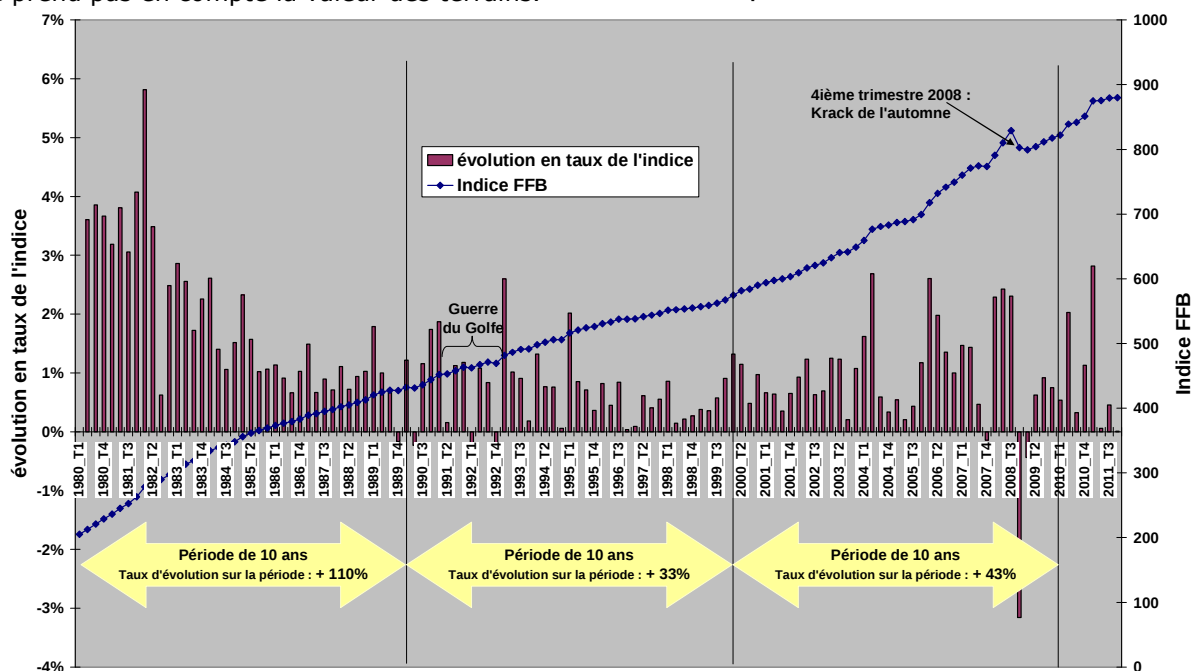
Constats :

- Il est possible de constater des événements naturels variés avec des conséquences plus ou moins désastreuses.
- La plupart des événements renseignés ici sont des événements majeurs. D'autres ont eu lieu mais n'ont pas été référencés car leurs ampleurs étaient moindres.

⁸ Sources de données : Application GASPAP tiré du site Web www.developpement-durable.gouv.fr

Annexe [2.] : Indice FFB de 1980 à fin 2011

L'indice FFB est calculé à partir du prix de revient d'un immeuble de rapport de type courant à Paris. Il enregistre les variations officielles du prix des différents éléments constitutifs des ouvrages (salaires et charges, prix des matériaux) constatées au dernier jour de chaque trimestre. Ce calcul ne prend pas en compte la valeur des terrains.



Graphique A. 5. Evolution trimestrielle de l'indice FFB de 1980 à fin 2011

Les valeurs de l'indice FFB ont été extraites du Site Internet de la Fédération Française du Bâtiment.

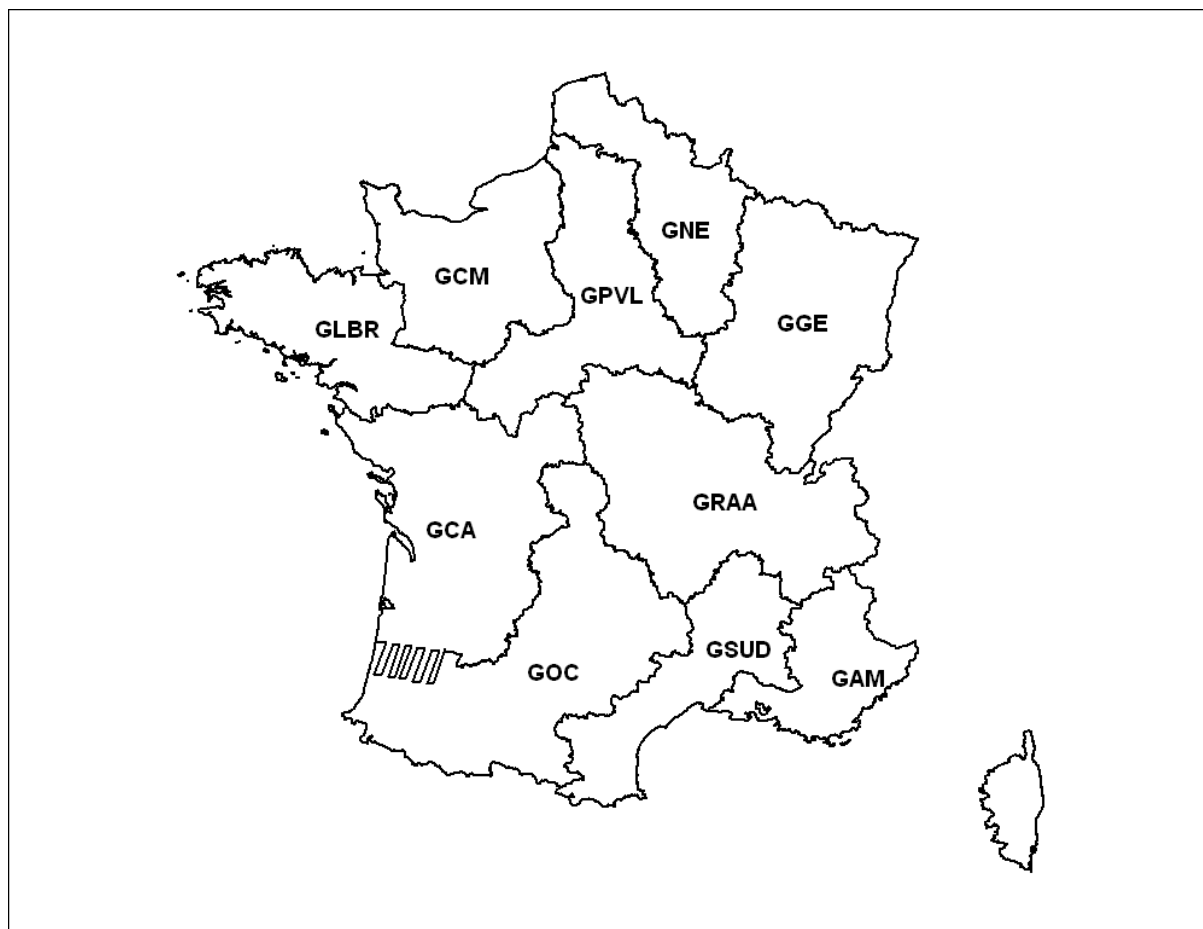
Panorama des différents types d'indice :

Indices et Index	Libellé	Origine	Périodicité	Méthodes de calcul	Année de création
FFB	Indice FFB du coût de la construction (ICC FFB)	Fédération Française du Bâtiment	Trimestriel	L'indice FFB du coût de la construction est calculé à partir du prix de revient d'un immeuble de rapport de type courant à Paris. Il enregistre les variations de coût des différents éléments qui entrent dans la composition de l'ouvrage. Ce calcul ne prend pas en compte la valeur des terrains.	1941
BT01	Index BT01 fait parti des index nationaux du bâtiment (BT)	Ministère de l'Équipement, des transports, du logement, du tourisme et de la mer	Mensuel	Ces index sont des indicateurs de coûts des facteurs de construction dans le secteur du bâtiment. Ils sont utilisés pour les révisions de prix calculés à partir de formules préétablies, intégrant un certain nombre de paramètres (salaire, matériaux, matériel, transport, énergie et divers) L'index du BT01 est un index agrégé qui reflète le coût des facteurs de construction pour l'ensemble du secteur du bâtiment.	1974
ICC	Indice du coût de la construction	INSEE et Ministère de l'Équipement, des transports, du logement, du tourisme et de la mer	Trimestriel	Concerne les bâtiments neufs à usage principal d'habitation non communautaire. Le calcul est effectué chaque trimestre à partir d'un échantillon représentatif retraçant l'évolution des prix de la construction des quelques 6 000 logements neufs au travers du suivi de 320 dossiers.	1953
IPEA	Indice des prix de l'entretien et de l'amélioration	Ministère de l'écologie, du développement et de l'aménagement durables	Trimestriel	Le principe du calcul de l'IPEA consiste à suivre un panel d'entreprises exerçant une activité d'entretien et d'amélioration de logements et à mesurer l'évolution de prix pour des prestations représentatives déterminées par ces entreprises. Les prix sont des prix de marché (hors sous-traitance) hors TVA, tenant compte des rabais et remises accordés aux clients.	1993

Tableau A. 2. Panorama des différents types d'indice d'inflation liés au coût du bâtiment

Annexe [3.] : Les Caisses Régionales Groupama en France métropolitaine

Afin d'avoir une vision plus précise en termes de localisation des Entités, il est proposé ci-dessous une cartographie des Caisses Régionales.



Graphe A. 6. Cartographie des Caisses Régionales à la fin de l'exercice 2010

Abréviation	Libellé
GAM	Groupama Alpes Méditerranées
GCA	Groupama Centre Atlantique
GCM	Groupama Centre Manche
GGE	Groupama Grand Est (récente fusion avec Groupama Alsace : GAL)
GLBR	Groupama Loire Bretagne
GNE	Groupama Nord Est
GOC	Groupama Oc
GPVL	Groupama Paris Val de Loire
GRAA	Groupama Rhône Alpes Auvergne
GSUD	Groupama SUD

Tableau A. 3. Abréviation des Caisses Régionales

À ces Caisses Régionales métropolitaines peuvent être ajoutées deux autres Entités :

- GAG (Groupama Antilles Guyanes) ;
- GOI (Groupama Océan Indien).

Il est précisé que cette cartographie a évolué sur l'exercice 2011. En effet, les Caisses Régionales GSUD et GAM ont fusionné afin de former l'Entité GMED (Groupama Méditerranée).

Annexe [4.] : Compléments sur les méthodes de régression

Dans cette annexe, il sera décrit dans un premier temps, ce que sont les principes du Maximum de Vraisemblance et Moindres Carrés Ordinaires. Dans un second temps, pour les lois étudiées dans ce rapport, il sera montré l'appartenance des lois de Poisson, Binomiale Négative et Gamma aux familles exponentielles et leur fonction de lien canonique. Enfin, il sera décrit ce que sont les familles de lois conjuguées et leurs intérêts dans la modélisation d'une prime pure.

Annexe [4.1.] : Moindres Carrés Ordinaires

Par définition, l'estimateur aux moindres carrés ordinaires correspond au minimum de la somme des carrés des erreurs de prédiction. Si $\hat{\beta}$ est l'estimateur, y la réponse et X le régresseur, alors l'estimateur est défini tel que $\hat{\beta} = \arg \min_{\beta} \|y - X\beta\|^2 = \arg \min_{\beta} \left(\sum_i (y_i - x_i \beta)^2 \right)$.

Annexe [4.2.] : Maximum de Vraisemblance

Soit $X = (x_1, \dots, x_n)$ une réalisation de variables aléatoires dont la densité de probabilité est définie par une fonction $f_{\theta}(x)$ de paramètre θ inconnu, alors la vraisemblance $L(\theta)$ et la log-vraisemblance $\ell(\theta)$ sont définies respectivement telles que $L(\theta) = \prod_{i=1}^n f(x_i, \theta)$ et $\ell(\theta) = \log L(\theta)$.

La vraisemblance est particulièrement utilisée :

- pour trouver des estimateurs dont la volatilité (variance) soit la plus petite possible, ou
- pour construire des tests puissants.

La variable aléatoire $\hat{\theta}$ qui satisfait l'équation $L(\hat{\theta}) \geq L(\theta)$ ou l'équation $\ell(\hat{\theta}) \geq \ell(\theta)$, pour tout θ s'appelle l'estimateur du maximum de vraisemblance. Cette seconde expression est utilisée beaucoup plus souvent dans les calculs, de par sa flexibilité.

Estimateur du Maximum de vraisemblance pour la loi de Poisson

Partant de la loi de distribution $P(y = n) = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$, avec $\lambda > 0$, pour un échantillon de n observations, il vient aisément $L(\lambda) = \prod_{i=1}^n P(y = k_i) = \exp(-n\lambda) \times \prod_{i=1}^n \frac{\lambda^{k_i}}{k_i!}$ où k_i désigne la valeur de la $i^{\text{ème}}$ observation. Par passage au logarithme, il ressort $\ell(\lambda) = -n\lambda + \sum_{i=1}^n k_i \cdot \log \lambda - \log \sum_{i=1}^n \log k_i$.

En dérivant $\ell(\lambda)$ par λ et en posant $\frac{\partial \ell(\lambda)}{\partial \lambda} = 0$, il vient $\frac{\partial \ell(\lambda)}{\partial \lambda} = 0 = -n + \frac{1}{\lambda} \sum_{i=1}^n k_i$. Ainsi, le maximum de vraisemblance $\hat{\lambda}$ est défini tel que $\hat{\lambda} = \frac{\sum_{i=1}^n k_i}{n}$.

Estimateur du Maximum de vraisemblance pour la loi Binomiale Négative

La loi Binomiale Négative de paramètres n et p a pour distribution $P(y = k) = C_{k-1}^{n-1} \cdot p^n \cdot (1-p)^{k-n}$. Pour un échantillon de m observations, il vient aisément

$L(n, p) = \prod_{i=1}^m P(y = k_i) = p^{nm} \times \prod_{i=1}^m \left(C_{k_i-1}^{n-1} \times (1-p)^{k_i-n} \right)$ où k_i désigne la valeur de la $i^{\text{ème}}$ observation.

Par passage au logarithme, il ressort $\ell(n, p) = nm \log p + \sum_{i=1}^m \log(C_{k_i-1}^{n-1}) + \sum_{i=1}^m (k_i - n) \log(1-p)$.

Les dérivées en n et en p ne donnent pas de solutions analytiques simples. Seuls des résultats informatiques peuvent permettre ces estimations. Néanmoins, l'estimateur $\hat{p}$ peut être expliqué brièvement en fonction de n . En effet, $\frac{\partial \ell(n, p)}{\partial p} = 0 = \frac{nm}{p} - \frac{1}{1-p} \sum_{i=1}^m (k_i - n)$, alors le maximum de vraisemblance $\hat{p}$ est défini tel que $\hat{p} = \frac{m}{m + \sum_{i=1}^m \left(\frac{k_i}{n} - 1 \right)}$ dépendant du paramètre n .

Estimateur du Maximum de vraisemblance pour la loi Exponentielle

La fonction de densité de la loi Gamma se note $f(x) = \theta e^{-\theta x}$. Il vient aisément pour un échantillon X de n observations $L(X, \theta) = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n \theta e^{-\theta x_i} = \theta^n e^{-\sum_{i=1}^n x_i}$ où x_i désigne la valeur de la $i^{\text{ème}}$ observation. Par passage au logarithme, il ressort $\ell(X, \theta) = n \log(\theta) - \theta \sum_{i=1}^n x_i$. En dérivant ℓ par θ et en posant $\frac{\partial \ell(X, \theta)}{\partial \theta} = 0$, il vient $\frac{\partial \ell(X, \theta)}{\partial \theta} = 0 = \frac{n}{\theta} - \sum_{i=1}^n x_i$. Ainsi, le maximum de vraisemblance $\hat{\theta}$ est défini tel que $\hat{\theta} = \frac{n}{\sum_{i=1}^n x_i}$.

Estimateur du Maximum de vraisemblance pour la loi Gamma

La fonction de densité de la loi Gamma se note $f_{(\alpha, \beta)}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ ou selon la littérature $f_{(\alpha, \lambda)}(x) = \frac{1}{\lambda^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\lambda}}$ avec $\lambda > 0$ et $\alpha > 0$. En prenant cette dernière équation, il vient aisément pour un échantillon de n observations $L(X, \alpha, \lambda) = \prod_{i=1}^n f_{(\alpha, \lambda)}(x_i) = \frac{1}{\lambda^{n\alpha} (\Gamma(\alpha))^n} \times \prod_{i=1}^n x_i^{\alpha-1} e^{-\frac{x_i}{\lambda}}$ où x_i désigne la valeur de la $i^{\text{ème}}$ observation. Par passage au logarithme, il ressort $\ell(X, \alpha, \lambda) = -n\alpha \log(\lambda) - n \log(\Gamma(\alpha)) + (\alpha-1) \sum_{i=1}^n \log(x_i) - \sum_{i=1}^n \frac{x_i}{\lambda}$. Il y a, ici aussi, deux paramètres à déterminer. Analytiquement, la résolution des dérivées de $\ell(X, \alpha, \lambda)$ en α et en λ n'a pas de solutions simples. Néanmoins, l'estimateur $\hat{\lambda}$ peut être expliqué brièvement en fonction de α . En

effet : $\frac{\partial \ell(\alpha, \lambda)}{\partial \lambda} = 0 = -\frac{n\alpha}{\lambda} + \frac{1}{\lambda^2} \sum_{i=1}^n x_i = -n\alpha + \frac{1}{\lambda} \sum_{i=1}^n x_i$ et il vient ainsi $\hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n\alpha}$

Concernant la résolution des équations, des calculs informatiques sont nécessaires.

Annexe [4.3.1] : Familles exponentielles et fonctions de lien

La définition des familles exponentielles donnée au chapitre II.1. est notée de nouveau ci-dessous :

Une distribution de probabilités appartient à la famille exponentielle si, pour une mesure de Lebesgue ou de comptage désigné par $v(dy)$ sur $\mathbb{R}$, sa loi f peut s'exprimer sous la forme

$$f(y|\theta, \phi) = \exp\left\{\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)\right\} \text{ telle que } f(y|\theta, \phi)v(dy) \text{ forme une distribution de probabilité}$$

pour tout ϕ variant dans un certain domaine de $\mathbb{R}_+$, avec θ désignant le profil de risque et ϕ le paramètre de nuisance.

Le terme $c(y, \phi)$ contient toute la partie non-linéaire en y non déjà présente dans $v(dy)$ permettant de définir la variable Φ de la paramétrisation, puis θ .

A noter que chacune des lois de la famille exponentielle possède une fonction de lien spécifique, dite fonction de lien canonique. Cette fonction de lien permet de relier l'espérance de la variable aléatoire y au paramètre naturel θ de la loi considérée.

Une terminologie spécifique est donnée pour certaines fonctions de lien classique :

Fonction identité	$g(z) = z$
Fonction logarithme	$g(z) = \text{Ln}(z)$
Fonction inverse	$g(z) = 1/z$
Fonction Logit (ou logistique)	$g(z) = \text{Ln}(z/(1-z))$
Fonction cloglog	$g(z) = \text{Ln}(-\text{Ln}(1-z))$
Fonction puissance	$g(z) = z^p, p \neq 0$
Fonction réciproque	$g(z) = -1/z$
Fonction Probit	$g(z) = \Phi(z)$

où ϕ désigne la fonction de répartition de la loi normale $N(0,1)$.

Le choix de la fonction de lien est dicté selon le type de variable à expliquer y :

- Si y est un nombre naturel, les liens couramment utilisés seront les liens logarithme, probit ou cloglog
- Si y est continu, un choix classique est le lien inverse ou réciproque, ...

Cette annexe va préciser l'appartenance des lois de Poisson, Binomiale Négative, Gamma et Inverse Gaussienne aux familles exponentielles.

Appartenance de la loi de Poisson aux familles exponentielles

Partant de la loi de distribution $P(y = n) = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$, avec $\lambda > 0$ il vient aisément, après passage au

logarithme, $P(y = n) = \exp(-\lambda + n \cdot \ln(\lambda) - \ln(n!))$.

Ainsi, par identification avec $\theta = \ln(\lambda)$, $\phi = 1$, $b(\theta) = \exp(\theta) = \lambda$ et $c(y, \phi) = -\ln(n!)$, il est possible de conclure que la loi de Poisson appartient à la famille exponentielle.

De par la définition donnée à la loi de Poisson, le lien canonique est défini tel que $g(\lambda) = \theta$. Or, $b'(\theta) = \lambda$. Il vient donc, $b'(\cdot) = g^{-1}(\cdot)$ ($= r(\cdot)$ la fonction décrite au chapitre II.1.).

Le lien canonique naturel de la loi de Poisson est donc la fonction logarithme. En effet, puisque $b(\theta) = \exp(\theta)$, il vient $b'(\theta) = \exp(\theta) = g^{-1}(\theta)$ ce qui implique $g(u) = \ln(u)$.

Ce type de modèle est appelé communément modèle log-linéaire.

Il est possible, pour certaines études, que le lien identité soit mieux adapté que le lien log-linéaire qui implique un effet multiplicatif des facteurs.

Des tests avec différentes fonctions de liens ont été effectués. La fonction de lien la plus adéquate était la fonction de lien logarithmique.

Appartenance de la loi Binomiale Négative aux familles exponentielles

Comme il a été dit dans le chapitre II.1., la loi Binomiale Négative est également utilisée pour la modélisation des fréquences de sinistres, notamment parce qu'elle permet d'avoir une dispersion⁹ plus importante que la loi de Poisson.

En partant, là aussi, de la loi de distribution $P(y = k) = C_{k-1}^{n-1} \cdot p^n \cdot (1-p)^{k-n}$, il vient, après passage au logarithme, $P(y = k) = \exp((n-1) \cdot \ln(p) + (k-n) \ln(1-p) + \ln(C_{k-1}^{n-1}))$

$$= \exp\left((n-1) \cdot \ln\left(\frac{p}{1-p}\right) + (k-1) \cdot \ln(1-p) + \ln(C_{k-1}^{n-1})\right)$$

En posant, $y = k-1$, $\phi = 1$, $\theta = \ln(1-p)$, $b(\theta) = (n-1) \cdot \ln(p/(1-p)) = (n-1)(\ln(p) - \theta)$ et $c(y, \phi) = \ln(C_{k-1}^{n-1})$, il est possible de conclure que la loi Binomiale Négative appartient à la famille exponentielle. Le lien canonique naturel de la loi Binomiale Négative est la fonction réciproque.

Appartenance de la loi Gamma aux familles exponentielles

La loi Gamma a pour fonction de densité $f_{(\alpha, \beta)}(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot y^{\alpha-1} e^{-\beta y}$ ou selon la littérature

$$f_{(\alpha, \lambda)}(y) = \frac{1}{\lambda^\alpha \Gamma(\alpha)} \cdot y^{\alpha-1} e^{-\frac{y}{\lambda}} \text{ ou bien encore } f_{(a, d)}(y) = \frac{1}{\Gamma(a)} \left(\frac{a}{d}\right)^a \cdot y^{a-1} e^{-\frac{ay}{d}}. \text{ En prenant cette}$$

dernière équation, il vient aisément, après passage au logarithme :

$$f_{(a, d)}(y) = \exp\left(\ln\left(\Gamma(a) \left(\frac{a}{d}\right)^a\right) + (a-1) \ln(y) - \frac{ay}{d}\right) = \exp\left(-\frac{ay}{d} + a \ln y - \ln y - \ln \Gamma(a) + a \ln\left(\frac{a}{d}\right)\right)$$
$$\frac{y\left(-\frac{1}{d}\right) - \left(-\ln\left(\frac{1}{d}\right)\right)}{1/a} - \ln(y\Gamma(a)) + a \ln(ay)$$

Il vient $f_{(a, d)}(y) = \exp$

Ainsi, en posant $\theta = -\frac{1}{d}$, $b(\theta) = \ln(-\theta) = \ln\left(\frac{1}{d}\right)$, $\phi = \frac{1}{a}$, et $c(y, \phi) = a \ln(ya) - \ln(y\Gamma(a))$, il est possible de conclure que la loi Gamma appartient à la famille exponentielle. Le lien canonique naturel de la loi Gamma est la fonction réciproque. En effet, puisque $b(\theta) = \ln(-\theta)$, il vient $b'(\theta) = -1/\theta = g^{-1}(\theta)$ impliquant $g(u) = -1/u$.

Il peut sembler intéressant de préciser que la fonction de lien logarithmique est la fonction de lien inverse la plus utilisée dans le cadre du modèle de régression Gamma. C'est également le cas dans ce rapport. En effet, le modèle tarifaire actuel est multiplicatif. Le modèle Gamma de fonction de lien logarithmique permet d'interpréter de manière plus simple les résultats.

Appartenance de la loi Inverse Gaussienne aux familles exponentielles

La loi Inverse Gaussienne a pour fonction de densité $f_{(a, d)}(y) = \frac{\sqrt{a}}{\sqrt{2\pi} y^3} e^{-\frac{a(y-d)^2}{2d^2 y}}$ avec $a > 0$ et $d > 0$.

Il vient après passage au logarithme :

$$f_{(a, d)}(y) = \exp\left(\frac{-ay^2 + 2ayd - ad^2}{2d^2 y} + \ln\left(\frac{\sqrt{a}}{\sqrt{2\pi y^3}}\right)\right) = \exp\left(\frac{\left(\frac{-1}{2d^2}\right)y - \left(\frac{-1}{d}\right)}{1/a} - \frac{a}{2y} + \frac{1}{2} \ln\left(\frac{a}{2\pi y^3}\right)\right)$$

Ainsi, en posant $\theta = -\frac{1}{2d^2}$, $b(\theta) = -\frac{1}{d} = (-2\theta)^{-\frac{1}{2}}$, $\phi = \frac{1}{a}$, et $c(y, \phi) = -\frac{1}{2y\phi} - \frac{1}{2} \ln(2\pi y^3 \phi)$, il est

possible de conclure que la loi Inverse Gaussienne appartient à la famille de loi exponentielle. Le lien canonique naturel de cette loi est la fonction puissance.

⁹ Selon le rapport moyenne/variance :

- si ce rapport est significativement supérieur à 1, la situation est qualifiée de sur-dispersée,
- si par contre ce rapport est inférieur à 1, la situation est qualifiée de sous-dispersée.

Annexe [4.4.] : Familles de lois conjuguées¹⁰

Lorsqu'un assureur dispose d'informations sur la structure collective des risques d'un portefeuille, l'information peut être résumée par une distribution de probabilité $U(\theta)$ sur Θ .

De manière plus précise, si les assurés ont des profils de risque θ_i ($i \in \{1, \dots, l\}$) inconnus de l'assureur, alors les θ_i prennent leur valeur dans Θ . Dans le cas d'un portefeuille réellement homogène, Θ est réduit à un singleton.

Par définition, la distribution $U(\theta)$ est alors appelée fonction de structure du portefeuille.

Soient,

F_θ la loi conditionnelle du nombre de sinistres pour $\Theta = \theta$ et

$U(\theta)$ la fonction de structure.

Respectivement,

$A = \{F_\theta : \theta \in \Theta\}$ la famille des distributions possibles et

$B = \{U_\gamma : \gamma \in \Gamma\}$ la famille des fonctions de structure.

La famille B est conjuguée à la famille A si, pour tout $\gamma \in \Gamma$ et pour toute réalisation x du vecteur des observations X , il existe $\gamma' \in \Gamma$ tel que $U_\gamma(\theta | X = x) = U_{\gamma'}(\theta)$, pour tout $\theta \in \Theta$.

Un des principaux intérêts d'utilisation des familles de distributions conjuguées est que la distribution a posteriori obtenue sur une période, peut être utilisée comme *a priori* sur la période suivante.

Ceci a notamment pour effet de simplifier les processus de mise à jour des primes : les méthodes mises en place peuvent être ainsi conservées pour les futurs exercices.

Si, en pratique, la spécification de A n'est pas évidente, celle de B l'est encore moins.

En règle générale, les conditions suivantes doivent être vérifiées :

La famille B doit être assez grande pour contenir les distributions qui peuvent décrire le collectif.

La famille B doit être aussi petite qu'il est possible de représenter la connaissance du collectif.

Idéalement, les familles B et A devraient être choisies de manière à ce que la prime de Bayes puisse s'exprimer analytiquement.

Il est rappelé ici les quelques familles exponentielles usuelles recensées avec leurs lois *a priori* conjuguées dans l'ouvrage de Christian Robert (2006) *Le choix bayésien: principes et pratique*.

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
Loi de Poisson $P(\theta)$	Loi Gamma $G(\alpha, \beta)$	$G(\alpha + x, \beta + 1)$
Loi Binomiale Négative $BN(n, \theta)$	Loi Bêta $Be(\alpha, \beta)$	$Be(\alpha + n, \beta + x)$
Loi Binomiale $B(n, \theta)$	Loi Bêta $Be(\alpha, \beta)$	$Be(\alpha + x, \beta + n - x)$
Loi Normale $N(\theta, \sigma^2)$	Loi Normale $N(\mu, \tau^2)$	$N(\frac{\sigma^2\mu + \tau^2x}{\sigma^2 + \tau^2}, \frac{\sigma^2\tau^2}{\sigma^2 + \tau^2})$
Loi Gamma $G(v, \theta)$	Loi Gamma $G(\alpha, \beta)$	$G(\alpha + v, \beta + x)$

¹⁰ Cette annexe est inspirée des cours dispensés par THEROND Pierre (2010), *Théorie de la Crédibilité, ISFA*

Annexe [5.] : Descriptif des tests d'indépendance et de significativité¹¹

Un test est un mécanisme (une règle de décision) qui permet de trancher entre deux hypothèses (H_0 et H_1) à l'aide de résultats observés sur un échantillon. Ces hypothèses peuvent, en autres, porter sur des estimations (valeur d'un moment, égalité de variances, nature d'une loi de probabilité, ...).

Soient H_0 et H_1 deux hypothèses, dont une et une seule est vraie. La décision aboutira à choisir H_0 ou H_1 . Il y a donc 4 possibilités dont les probabilités peuvent être résumées dans le tableau suivant :

	H_0 est vraie	H_1 est vraie
H_0 est privilégiée	$1 - \alpha$	β
H_1 est privilégiée	α	$1 - \beta$

α et β sont appelées respectivement erreurs de première espèce et de deuxième espèce.

- α est la probabilité de privilégier H_1 , alors que H_0 est vraie,
- β est la probabilité de privilégier H_0 , alors que H_1 est vraie.

Ces deux erreurs sont antagonistes, plus α sera grand (respectivement petit), plus β sera petit (respectivement grand). Le fait d'imposer un α faible conduit à une règle de décision plus stricte. Elle aboutira le plus souvent à abandonner l'hypothèse H_0 que dans des cas rarissimes. L'hypothèse sera donc conservée quelquefois à tort.

La quantité α est appelée « niveau du test ». La quantité $1 - \beta$ est appelée « puissance du test ».

Dans la pratique des tests statistiques, il est de règle de se fixer α comme donnée (les valeurs les plus courantes sont 0.05, 0.01 ou 0.1) de préférence en fonction du risque de première espèce. Les tests sont construits de telle sorte à avoir la plus grande puissance possible.

En général, H_0 joue le plus souvent un rôle prédominant par rapport à l'hypothèse H_1 (H_0 joue le rôle d'hypothèse de référence alors que H_1 est souvent limitée à l'hypothèse contraire). Par exemple, il est possible d'avoir H_0 : « $m = m_0$ » ce qui est relativement facile à tester. Dans ce cas, H_1 est tout simplement « $m \neq m_0$ ».

Bien souvent, α est choisi *a priori*, mais l'apparition de plus en plus fréquente de processus numériques d'approximation rapides et précis permet une autre approche. En l'occurrence, cette approche consiste à rechercher la plus petite valeur de α pour laquelle l'hypothèse H_0 reste vraie.

Pour conclure cette introduction, il existe deux types de tests :

- les tests paramétriques : issus d'une loi paramétrée ;
- les tests non paramétriques : les cas opposés.

Annexe [5.1.] : Test du Khi-Deux

La statistique du Khi-Deux, ou Chi-Deux, notée encore χ_n^2 a été inventée par le mathématicien britannique, Karl Pearson.

Les principales utilisations du test du Khi-Deux ont pour objectif de vérifier, sous H_0 , l'hypothèse :

- d'égalité des distributions observées (test d'homogénéité),
- d'indépendance entre deux caractères qualitatifs (test d'indépendance)
- de conformité à une loi de probabilité connue (test d'ajustement ou selon la littérature test d'adéquation).

Quel que soit le test du χ^2 réalisé, l'objectif est de déterminer si les écarts aléatoires entre la distribution des effectifs observés et la distribution des effectifs théoriques espérés sont conformes, et donc significatifs, ou si les écarts sont imputables uniquement aux fluctuations d'échantillonnage.

¹¹ Cette annexe est inspirée de l'ouvrage de SAPORTA Gilbert (2006) *Probabilité, Analyse des données et Statistique*. Edition Technip, du site Internet <http://spiral.univ-lyon1.fr/mathsv/> ainsi que des cours de *Régression* dispensés à l'université de Beaulieu, Rennes 1, par DELYON Bernard (2005)

Expression d'un test d'indépendance pour 2 variables :

Dans ce cadre, le test du Khi-Deux est un test non paramétrique.

Soient X et Y deux variables ayant respectivement a et b modalités.

Soit n_{ij} le nombre d'individus possédant simultanément la modalité i de la variable X et la modalité j de la variable Y , avec bien évidemment $1 \leq i \leq a$ et $1 \leq j \leq b$.

La probabilité théorique correspondante est p_{ij}

Soit le test suivant :

$$H_0 : p_{ij} = p_{i.} \times p_{.j} \quad \text{avec} \quad p_{i.} = \sum_{j=1}^b p_{ij} \quad \text{et} \quad p_{.j} = \sum_{i=1}^a p_{ij}$$

$$H_1 : p_{ij} \neq p_{i.} \times p_{.j}$$

La statistique suivante est alors calculée : $D_n = \sum_{i=1}^a \sum_{j=1}^b \frac{(n_{ij} - n_{i.} n_{.j} / n)^2}{n_{i.} n_{.j} / n}$ où $n_{i.}$ désigne la moyenne de la variable Y , $n_{.j}$ désigne la moyenne de la variable X et n la moyenne globale.

Sous H_0 , la loi approchée par D_n est la loi du Khi-Deux : $\chi^2_{(b-1, a-1)}$.

Par conséquent, la statistique D_n est comparée à une valeur critique V_{crit} (définie par α) spécifique à la loi du Khi-Deux. La comparaison s'effectue dans la table du Khi-Deux et permet de rejeter ou d'accepter l'hypothèse H_0 (si $D_n < V_{crit}$, alors H_0 est acceptée, rejetée sinon).

En aparté, le coefficient de contingence C de Pearson est défini tel que $C = \left(\frac{D_n}{n + D_n} \right)^{1/2}$.

Si la mesure vaut 0, alors il y a indépendance, si la mesure vaut 1, alors il y a une liaison fonctionnelle.

Rappel sur la loi du Khi-Deux

Soient k variables aléatoires $X_1, \dots, X_k$ indépendantes de même loi normale centrée et réduite, alors

par définition, la variable X , telle que $X = \sum_{i=1}^k X_i^2$ suit une loi du χ^2_k à k degrés de liberté.

La densité de X notée f sera : $f(x) = \frac{t^{\frac{k}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})}$ pour tout x positif et où Γ est la fonction gamma.

Annexe [5.1.bis] : Les limites du test d'indépendance du Khi-Deux

Rappel sur les propriétés du test d'indépendance du Khi-Deux

Le Khi-Deux de Pearson prend des valeurs positives. Il est nul sous H_0 en cas d'association « parfaite » entre 2 variables. Il prend une valeur qui dépend de N et du nombre de modalités : $N \times \text{minimum}(I-1, c-1)$ (où I désigne le nombre de lignes et c le nombre de colonnes).

Il permet de tester l'hypothèse d'indépendance à l'aide d'une statistique de test qui suit asymptotiquement (c'est-à-dire si N est grand) une loi du Khi-Deux à $(I-1)(c-1)$ degrés de liberté.

Pour remédier à l'influence de N dans le calcul, plusieurs tests ont été mis en place¹². Par exemple, Pearson a proposé le coefficient ϕ^2 .

$$\phi^2 : \phi^2 = \chi^2 / N$$

Cette statistique a la particularité de ne pas dépendre de la taille de l'échantillon. C'est une statistique descriptive. Elle prend, elle aussi, la valeur 0 sous indépendance. A l'inverse, elle vaut sous association parfaite, le minimum de $(I-1, c-1)$.

SAS donne $\phi = \sqrt{\phi^2}$ avec un signe positif si, dans la table, l'association se retrouve suivant la diagonale ou négatif, si celle-ci se retrouve sur « l'anti-diagonale ».

Cependant, il est délicat d'utiliser cette statistique pour tester l'indépendance, car la loi de ϕ n'est pas connue de façon théorique.

¹² Extrait de l'article : « La procédure FREQ de SAS : Tests d'indépendance et mesures d'association dans un tableau de contingence » de CONFAIS Josiane, GRELET Yvette, LE GUEN Monique (2005)

Pour obtenir une autre mesure qui ne dépende pas de l'effectif total N , et soit plus petite que 1, Pearson a proposé le coefficient de contingence C .

C (« Contingency coefficient ») :
$$C = \sqrt{\frac{\chi^2}{N + \chi^2}} = \sqrt{\frac{\phi^2}{1 + \phi^2}}$$

Cette statistique est comprise entre 0 et 1, sans que la valeur 1 soit atteinte. C vaut 0 sous indépendance, sa valeur sous association parfaite dépend, là aussi, de l et c (si $l = c$, la valeur est de $\sqrt{1-1/l}$). La valeur peut être très éloignée de 1. Il est, là aussi, délicat d'utiliser cette statistique pour tester l'indépendance. En effet, la loi de C n'est pas connue.

Pour obtenir un coefficient qui puisse atteindre la valeur 1, Cramer a proposé le coefficient V .

V de Cramer :
$$V = \phi / \sqrt{\min(l-1, c-1)}$$

Les valeurs possibles de ce coefficient sont comprises entre -1 et +1. Il vaudra 0 sous indépendance et +1 ou -1 sous association parfaite.

C'est donc une mesure d'association ressemblant au coefficient de corrélation linéaire entre variables quantitatives. Ici, la loi de V n'est là aussi pas encore connue, il est donc délicat de l'utiliser pour tester l'indépendance de variables.

SAS donne, dans l'option CHISQ, deux autres mesures qui ont la propriété de suivre des lois du Khi-Deux, mais qui ne sont pas des dérivées du χ^2 d'indépendance.

Test du rapport de vraisemblance : G^2 « likelihood ratio » :

Il s'agit de la statistique du test d'indépendance construite à partir du Rapport du Maximum de

Vraisemblance (RMV) :
$$G^2 = -2\text{Log}(RMV) = 2\sqrt{n_{ij} \text{Log}\left(n_{ij} / (n_{i.} n_{.j} / N)\right)}$$
 littéralement $RMV = \frac{\sup_{\theta \in H_0} \ell(\theta)}{\sup_{\theta \in H_1} \ell(\theta)}$

(ℓ la log-vraisemblance définie à l'Annexe [4.2.]).

Ses valeurs sont positives. Il vaut 0 sous indépendance. Asymptotiquement (si N grand), il suit une loi du Khi-Deux à $(l-1)(c-1)$ degrés de libertés, et donc peut être utilisé pour tester l'indépendance.

Remarque : Si l'indépendance de H_0 est proche, ou si N est grand alors la valeur du G^2 est proche de celle du χ^2 .

Khi-Deux de Mantel-Haenszel appelé également Q_{mh} :

Cette statistique mesure l'association entre deux variables. Elle est calculée à partir du coefficient de corrélation linéaire ρ entre les variables dont les modalités sont codées numériquement (ce codage est défini par l'option SCORES de SAS) : il n'est donc à utiliser que si les variables sont ordonnées.

$$Q_{mh} = (N-1) \rho^2$$

Il vaut 0 sous indépendance et $((N-1)/N) \times \text{minimum}(l-1, c-1)$ sous association parfaite. Il a la propriété de suivre une loi du Khi-Deux à 1 degré de liberté quelle que soit la taille de la table.

Programme SAS [5.1.] : Test d'indépendance

```
*****;
** Test d'indépendance entre les variables : Test du khi-2 et autre ****;
*****;
%macro TESTIndependance(tabentree,Var1,Var2);
proc freq data=&tabentree. order=data;
  tables &Var1.*&Var2. / chisq relrisk expected CMH out=Khi2;
  exact pchi or;
  weight NbContrat;
  title 'Test d indépendance entre les variables';
run;
%mend;
%TESTIndependance(tabentree,Var1,Var2);
```

Liste des variables Var1 et Var2 :

NbPieces, TypeBat, LimiteCapitalMobilier, Franchise, Domotique, Dependance,...

Programme SAS [5.2.] : Extraction d'un échantillon aléatoire

```
*****;
*** Extraction d'un échantillon aléatoire dans une table donnée ***;
*****;

%macro extract(libref=,entree=,sortie=,nb=);
data &libref..&sortie (drop=i j count);
count=0;
array obsnum(&nb) _temporary_; * nombre d'observation dans la table entree;
do i=1 to &nb;
redo:
select=ceil(ranuni(11111)*n);
* création d'un nombre compris entre 1 et n (nombre de ligne d'entree);
set &libref..&entree point=select nobs=n;
do j=1 to count;
if obsnum(j)=select then goto redo; * récup. de la ligne select ds entree;
end;
position=select; count=count+1; obsnum(count)=select;
output; * affectation de la ligne dans la table sortie;
end;
stop;
set &libref..&entree;
run;
%mend;

%extract(libref=work,entree=tabentree,sortie=tabsortie,nb=30000);
```

Programme SAS [5.3.] : Constitution d'une base agrégée pour un gain de traitement

```
*****;
*** Constitution de la base agrégée pour estimation (gain de traitement);
*****;

proc means DATA=BaseMRH.BaseEvtCli noprint nway;
class Cais Exercice Qnnbp Ctnbi Limitecapitalmobilier Dependance fra_in
CDFORM CINBM CINBE CINBF CINBK Cqnbh RAN_lib;
var i_en ns_en s_en;
output out=Tabsortie(drop=_TYPE_ rename=(_FREQ_=N)) sum=;
run;

data Tab; set tab; logN=log(i_en);
NbContrat=round(i_en); TypBat_NbPiece=Ctnbi||Qnnbp; run;
```

Annexe [5.2.] : Test de Type I contre test de Type III sous SAS

Il convient ici de rappeler que le but premier des tests de Type I et de Type III de SAS est de montrer la significativité de certains facteurs.

Dans cette optique, le test de type III sera plus adéquat (car plus sévère et donc plus convaincant). En revanche, l'utilisation des tests pour sélectionner le modèle ne doit être vue que comme une application supplémentaire.

Variable significativement décorrélée.

Il peut arriver qu'une des variables explicatives ait une corrélation quasi-nulle avec les réponses (et soit même rejetée dans l'analyse de Type I mais qu'elle soit significative dans l'analyse de Type III. Ceci vient du fait qu'elle est implicitement présente dans d'autres régresseurs importants.

C'est le cas si le « vrai » modèle est $y_i = \beta z_i + u_i$ et si le modèle estimé est $y_i = \beta_1 x_i + \beta_2 x'_i + u_i$ basé sur les régresseurs $x_i = t_i + z_i$, $x'_i = t_i$ où t_i est une variable quelconque de variance assez grande.

La présentation habituelle de ces résultats d'analyse de variance consiste à donner sur chaque ligne du tableau :

- le facteur à tester,
- le « $RSS_0 - RSS^{13}$ » correspondant au test,
- les degrés de liberté,
- la statistique de Fisher, puis celle du Khi-deux,
- le niveau de signification obtenu pour ces 2 tests.

La somme des carrés ainsi présentée illustre la contribution de chaque facteur à la variance totale ; mais ces nombres n'ont toutefois pas grand intérêt, puisque les niveaux résument au mieux l'ensemble.

Annexe [5.3.] : Test de Fisher-Snedecor et Test de Student

Si, sur un premier échantillon, les réalisations d'une variable aléatoire X_1 suivent une fonction de répartition $F_1(x)$ et si, sur un deuxième échantillon, les réalisations d'une variable aléatoire X_2 suivent une fonction de répartition $F_2(x)$, alors il peut sembler légitime de tester les hypothèses :

$H_0 : F_1(x) = F_2(x)$ contre $H_1 : F_1(x) \neq F_2(x)$.

Dans la pratique, seule l'égalité des espérances et des variances de X_1 et X_2 est vérifiée.

Le test de Fisher-Snedecor, ou test de Fisher, ou test F , est un test d'hypothèse statistique qui permet de tester le rapport des variances de deux variables X et Y ou deux échantillons. Ce rapport ne doit pas dépasser une certaine valeur théorique dans la table de Fisher. Dans le cas contraire, l'hypothèse H_0 est rejetée.

Loi de Fisher

Soient X et Y des variables suivant indépendamment des lois χ_n^2 et χ_p^2 , alors $F(n; p) = \frac{X/n}{Y/p}$.

Si le test de Fisher-Snedecor aboutit à la conclusion d'égalité des variances, alors il faut tester l'égalité des espérances. Le test de Student est plutôt approprié pour ce cas. Il suppose implicitement l'égalité des variances.

Loi de Student

Soit une variable aléatoire U suivant une loi log-normale (0,1) et X une variable aléatoire suivant indépendamment de U une loi du χ_n^2 . La variable de Student T_n à n degrés de liberté est définie

comme étant : $T_n = \frac{U}{\sqrt{\frac{X}{n}}}$.

Une relation entre les variables de Student et de Fisher-Snedecor existe : $(T_n)^2 = F(1; n)$

¹³ RSS : « *Residual Sum of Squares* », Somme des carrés des résidus $RSS = \sum_i (y_i - x_i X . (X^t X)^{-1} X^t y)$ et RSS_0 la somme des carrés des résidus sous l'hypothèse d'égalité des facteurs.

Annexe [6.] : Technique de sélection de variables et choix d'un modèle¹⁴

Plutôt que de chercher à expliquer la réponse y par p critères explicatifs disponibles, il est possible de définir un nombre limité de variables parmi ces facteurs, pour rendre une reconstitution presque aussi satisfaisante de y .

Les objectifs d'une telle démarche sont multiples :

- économiser le nombre de prédicteurs,
- obtenir des formules stables et d'un bon pouvoir prédictif en éliminant des variables redondantes qui augmentent le facteur d'inflation de la variance,
- obtenir un modèle plus facile à interpréter.

Les critères de choix dépendent des usages souhaités pour la régression :

- reconstitution des y_i ,
- prévision de valeurs futures,
- ou estimation précise des paramètres d'un modèle.

Recherche exhaustive

Lorsque p n'est pas trop grand, il est possible d'étudier les C_p^k combinaisons possibles, où k désigne un certain nombre de variables. Ce qui représente $2^p - 1$ régressions.

A p fixé, le R^2 maximum sera privilégié (R^2 est défini ci-dessous). Si p n'est pas fixé, le choix sera plutôt porté sur la minimisation du $\hat{\sigma}^2$.

Les méthodes pas à pas

Ces méthodes pas à pas sont utilisées lorsque p est élevé et qu'il n'est pas possible de procéder à une recherche exhaustive. Il y a tout d'abord la méthode *backward* qui consiste à entrer toutes les variables retenues, puis à retirer une à une les variables non significatives du modèle, jusqu'à ce que les variables restantes soient toutes significatives. Il y a la méthode *forward* qui consiste à faire le « chemin » inverse. Et la méthode *stepwise* qui combine les deux précédentes méthodes.

La détermination d'un modèle se fait quant à elle par la vraisemblance pénalisée. Dans le cadre de modèles de prédiction, les critères qui suivront seront mieux adaptés.

Le coefficient de corrélation multiple R et le coefficient de détermination R^2

La corrélation empirique entre les données et les valeurs prédites est notée R et est définie telle

que :
$$R = \frac{\sum_i (y_i - \bar{y})(\hat{y}_i - \bar{y})}{\left(\sum_i (y_i - \bar{y})^2\right)^{1/2} \left(\sum_i (\hat{y}_i - \bar{y})^2\right)^{1/2}}, \text{ avec } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \text{ et } \hat{y} \text{ le vecteur des valeurs ajustées}$$

tel que $\hat{y} = X.(X^t X)^{-1} X^t y$.

L'interprétation la plus simple est de voir R comme une mesure de corrélation entre les variables explicatives (globalement) et les réponses. Plus R est proche de 1, plus le modèle représente bien les données.

R^2 est appelé *coefficient de détermination*, ou encore la *proportion de variance expliquée*.

Ce coefficient R^2 mesure l'adéquation géométrique du modèle aux données.

Ce critère R^2 est bien adapté dans le cadre d'une reconstitution des y_i . Il est d'ailleurs le critère le plus utilisé dans les programmes de régression pas à pas. Cependant, ce critère n'est pas à l'abri de critiques car il croît de façon monotone avec le nombre des variables : il ne peut qu'augmenter, même s'il y a un ajout d'un prédicteur peu corrélé avec y . Le R^2 n'est également pas adapté dans le cadre d'une minimisation de l'erreur de prévision.

Les critères d'information AIC, BIC

Soit p le nombre de paramètres à estimer. Il peut paraître évident que le modèle le plus compliqué (p grand) aura généralement l'erreur de prédiction la plus faible.

Pour comparer des modèles ayant des nombres de paramètres différents, il est coutumier d'utiliser les critères :

- D'Akaike : $AIC = -2\ell(y) + 2p$ AIC : Akaike Information Criterion
- De Schwartz : $BIC = -2\ell(y) + p \ln(n)$ BIC : Bayesian Information Criterion

¹⁴ Inspiré de l'ouvrage *Probabilité, Analyse des données et Statistique* de SAPORTA Gilbert, Edition Technip (2006) et des cours de Régression dispensés par Bernard Delyon à l'Université de Rennes 1.

avec notamment n le nombre d'observation et $\ell(y)$ la log-vraisemblance es observations (ℓ la log-vraisemblance définie à l'Annexe [4.2.]). En d'autres écritures, $\ell(y) = \ell(\hat{\beta}) = -\frac{n}{2} \ln(2e^1 \pi \sigma_{MV}^2)$. Le modèle qui sera privilégié sera le modèle pour lequel ces critères ont la plus faible valeur.

Lorsque n tend vers l'infini, la probabilité que le *BIC* choisisse le vrai modèle tend vers 1, ce qui est faux pour l'*AIC*.

Pour n fini, les résultats sont contradictoires. Le *BIC* ne choisit pas toujours le vrai modèle : il a tendance à choisir des modèles trop simples en raison de sa plus forte pénalisation.

L'*AIC* va choisir le modèle qui maximisera la vraisemblance de futures données et réalisera le meilleur compromis biais-variance.

L'*AIC* est un critère prédictif tandis que le *BIC* est un critère explicatif. (L'*AIC* est donc meilleur pour la prévision, le *BIC* choisit le « bon modèle »).

Méthode de validation croisée

La méthode qui suit est utilisée lorsqu'il a été fait le choix de séparer, de manière aléatoire, les données en deux parties distinctes *A* et *B*. Un ensemble de modèle peut alors être construit avec le jeu d'apprentissage *A*. Ensuite, tous ces modèles sont validés par le jeu de données *B*. Les coefficients trouvés sur le jeu de données *A* sont appliqués sur les observations du jeu de données *B*. Une comparaison des valeurs prédites et valeurs observées est alors réalisée.

Validation croisée : $CV = -2\ell(y_i / y^{(i)}) = \frac{1}{n} \sum_{i=1}^n \hat{u}_i^2 / (1 - h_i)^2$

Programme SAS [6] : Calculs des critères de tests et des paramètres des échelles

```
*****;
*** Récupération des OUTPUT souhaités + valeurs prédites;
*****;
%macro TESTPARAM(Tabentree,Var_etud,distrib,fct_lien,Var_pond,Tabsortie);

/* Critère d'information sur les modèles : AIC, BIC, ... */
data TEST1 TEST3 AJUST; set _null_; run;
ODS output type1=TEST1
            type3=TEST3
            modelfit=AJUST(where=(compress(criterion) ne "LogLikelihood"))
            parameterestimates=PARAM(keep=parameter level1 estimate probchisq
LowerCL UpperCL where=(compress(parameter) ne "Scale"))
            convergencestatus=infocvg;

/* Vérifier que la fréquence et le coût moyen des sinistres suivent une
certaine loi + test de significativité des variables + résidus */
proc genmod data=&Tabentree.;
class Var1 Var2 ... VarN;
model &Var_etud.= Var1 Var2 ... VarN /
dist=&distrib. link=&fct_lien. offset=&Var_pond. type1 type3 DSCALE;
Output out=&Tabsortie. Valpredite = Pred ResVraisembla = Reslik
                        ResBrut = Resraw
                        ResPearson = Reschi ResPearsonStd = Stdreschi
                        ResDevianc = Resdev ResDeviancStd = Stdresdev;

run;
%mend;

%TESTPARAM(Tabentree,Var_etud, distrib, fct_lien, Var_pond,Tabsortie);
```

Liste des variables (<i>Var_i</i>)	Loi de distribution (<i>distrib</i>)	Fonction de lien (<i>fct_lien</i>)	Variables étudiées (<i>Var_etud</i>)	Variables de pondération (<i>Var_pond</i>)
NbPièces, TypeBat, LimiteCapitalMobilier, Franchise, Domotique, Dependance, LocaAssureParAilleurs	Poisson, NEGBIN, BIN, gamma,...	Log, logit, probit, ...	ns_en (nombre de sinistre évènement. climatique), s_en (Coût des sinistres)	l_en (exposition), ns_en

Annexe [7.] : Descriptif méthodologique pour l'analyse des résidus

En régression, les résidus sont définis comme étant la différence entre la valeur observée y_i et la valeur prédite $\hat{y}_i$ déterminée par le modèle (les valeurs $\hat{u}_i = y_i - \hat{y}_i$ sont appelées résidus bruts).

L'étude des résidus est bien souvent la seule façon de vérifier empiriquement le bien-fondé des hypothèses posées pour le modèle. Cette analyse vise plusieurs objectifs, à savoir :

- vérifier l'hypothèse de normalité, d'indépendance et d'homoscédasticité des résidus,
- détecter des données aberrantes qui s'écarteraient considérablement du modèle jouant un rôle important dans la détermination de la régression,
- détecter des tendances particulières et des relations des résidus qui pourraient permettre d'affiner le modèle.

Les résidus bruts sont relativement difficiles à interpréter, car ils n'ont pas toujours la même variance. En effet, la loi de $\hat{u}_i$ suit, hypothétiquement, une $\mathcal{N}(0, (1-h_i)\hat{\sigma}^2)$. C'est pourquoi, selon le modèle étudié, les résidus de Pearson, les résidus de Student ou bien encore les résidus de déviance (appréciés pour les lois faisant parties de la famille exponentielle) leur seront généralement préférés :

- Pour $\hat{\sigma}^2$ désignant la variance des $\hat{y}_i$, les résidus de Pearson sont définis tels que $r_i^p = \frac{y_i - \hat{y}_i}{\hat{\sigma}}$.
- Les résidus de Student sont définis tels que $r_i^s = \frac{r_i^p}{\sqrt{1-h_i}}$ où h_i désigne le $i^{\text{ème}}$ terme diagonal du projecteur $X(X'X)^{-1}X'$ et X désigne la matrice des facteurs explicatifs du modèle de régression.
- Les résidus de déviance sont définis tels que $r_i^D = \text{signe}(y_i - \hat{y}_i) \sqrt{2\phi[\ell(y_i|Y;\phi) - \ell(\hat{y}_i|Y;\phi)]}$ où $\ell(\theta)$ décrit la log-vraisemblance (définie à l'Annexe [4.2.]), ϕ désigne le paramètre de nuisance (défini à l'Annexe [4.3.]) et où Y désigne les observations réelles.

Annexe [7.1.] : Tests de normalité

Dans le cadre de la normalité des résidus, plusieurs tests sont possibles :

- L'approche la plus simple consiste à tracer un histogramme des résidus. Les graphes des résidus en fonction des variables explicatives ne doivent laisser aucune tendance. Le cas contraire indiquerait que le modèle est mal adapté.
- La seconde consiste à tracer la droite de Henry, appelée également *QQ-plot*. Cette droite s'approxime en ordonnant les résidus bruts (ou autres), puis en les traçant en fonction de $Q(i/(n+1))$ (où Q désigne une fonction quantile, plus précisément, le quantile de la loi supposée) ; Dès lors, si les points sont approximativement alignés, alors la distribution des résidus correspond à la loi supposée pour le quantile. Toutefois, cette méthode est basée sur une linéarisation qui peut être très approximative pour des n petits.
- Enfin, des tests d'ajustement (Khi-Deux, Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises ou Anderson-Darling) peuvent être réalisés sur les résidus. Ces trois derniers tests sont décrits dans l'Annexe [7.1.bis].

Annexe [7.1.bis] : Complément sur les tests de normalité

Soit une variable aléatoire X de fonction de répartition F . Soit F_0 une fonction de répartition théorique continue. Les tests de normalité qui vont être exposés consistent à comparer F et F_0 . Les hypothèses suivantes sont alors posées : $\{H_0 : F = F_0\}$ contre $\{H_1 : F \neq F_0\}$.

Pour les 3 tests qui vont suivre, les variables (ou statistiques) Δ_n , I et A sont comparées à une valeur critique spécifique à chaque test (définie par α). Cette comparaison s'effectue dans une table et permet de rejeter ou d'accepter l'hypothèse H_0 .

Test d'ajustement de Kolmogorov-Smirnov

Le test d'ajustement de Kolmogorov-Smirnov est un test non paramétrique qui permet de tester l'hypothèse H_0 selon laquelle les données observées sont engendrées par une certaine loi de probabilité théorique, considérée comme étant une loi convenable au modèle.

Contrairement au test du Khi-Deux, la loi théorique doit ici être continue et entièrement spécifiée (tous les paramètres sont connus).

Dans ce test, les calculs se font sur les fonctions de répartition. Plus précisément, il est mesuré l'écart entre la fonction de répartition théorique et la fonction de répartition observée.

Si $X=(x_1, \dots, x_n)$ est un n -échantillon de X , la fonction de répartition empirique associée à cet échantillon est :

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n 1_{]-\infty, x_k]}(X_k)$$

$F_n(x)$ décrit la proportion des observations dont la valeur est inférieure ou égale à x .

L'écart entre les valeurs observées de F (i.e. F_n) et les valeurs théoriques de la fonction de répartition F_0 peut être mesuré par une variable aléatoire Δ_n .

Cette variable aléatoire est définie telle que $\Delta_n = \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|$. Ainsi posée, Δ_n sera la variable de décision, ou fonction discriminante, du test de Kolmogorov-Smirnov.

Test d'ajustement de Cramer-von Mises

Tout comme le test de Kolmogorov-Smirnov, le test d'ajustement de Cramer-von Mises est un test non paramétrique qui permet de tester l'hypothèse H_0 selon laquelle les données observées sont engendrées par une loi de probabilité théorique, considérée comme étant une loi convenable.

L'indicateur d'écart pour le test de Cramer-von-Mises est défini tel que $I = \int_{-\infty}^{+\infty} [F_0(x) - F(x)]^2 dF(x)$.

De manière empirique, $I = \frac{1}{12n} + \sum_{i=1}^n \left[\frac{2i-1}{2n} - F_0(x_i) \right]^2$ où les valeurs de l'échantillon sont ordonnées en ordre croissant.

Le test de Cramer-Von Mises a les mêmes applications que le test de Kolmogorov, à ceci près :

- Pour le test de Kolmogorov-Smirnov, seul l'écart maximum entre la distribution empirique et la distribution d'ajustement est pris en considération.
- Pour le test de Cramer-von Mises, l'ensemble des données est mieux pris en compte, puisque le test est basé sur la somme des écarts de chaque point x_i .

Par conséquent, le test de Kolmogorov-Smirnov est donc beaucoup plus sensible à l'existence de points aberrants dans un échantillon que le test de Cramer-von Mises.

Néanmoins, une faiblesse est à noter pour le test de Cramer-von Mises. En effet, la différence entre la fonction de répartition empirique et la fonction de répartition de référence tend vers 0 lorsque $x \rightarrow -\infty$ ou $x \rightarrow +\infty$. De ce fait, la valeur de I est ainsi peu sensible aux observations « lointaines » dans les queues de distribution.

Ces observations lointaines (ou événements rares) ont cependant un caractère intéressant dans une analyse des distributions (cf. les diverses analyses des coûts moyens de ce rapport).

La statistique d'Anderson-Darling introduit une fonction de pondération qui va redonner de l'importance aux observations lointaines.

Test d'ajustement d'Anderson-Darling

La fonction de pondération la plus utilisée est la fonction $[F(x).(1 - F(x))]^{-1}$. Elle est minimale aux alentours de la médiane, et tend vers l'infini quand x tend vers $-\infty$ ou $+\infty$. En fait, la fonction $F(x).(1 - F(x))$ correspond à la variance de la fonction de répartition empirique en x .

La statistique A définie telle que $A = n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 \cdot [F(x)(1 - F(x))]^{-1} \cdot dx$ s'appelle la statistique d'Anderson-Darling.

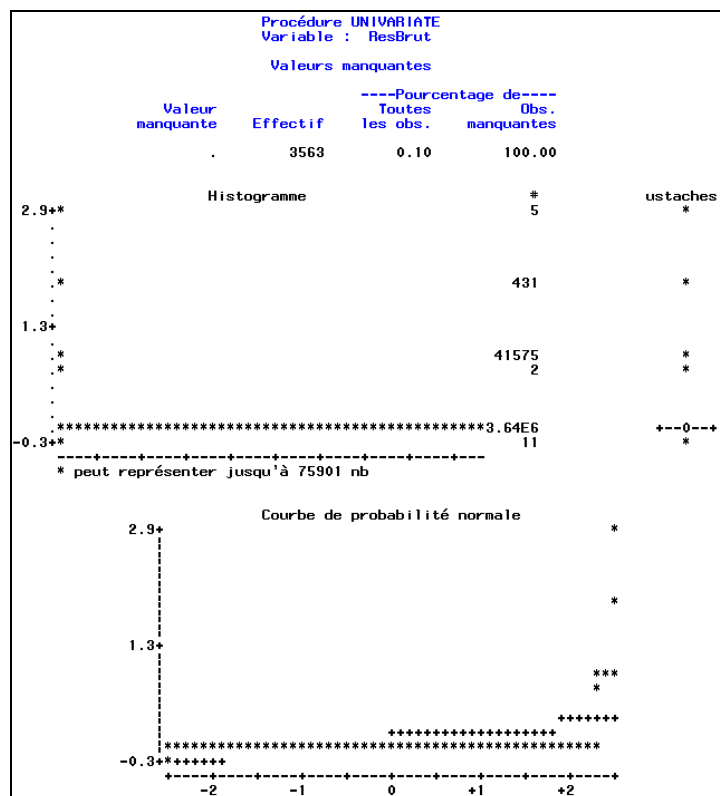
Cette statistique s'exprime empiriquement sous la forme d'une somme telle que $A = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln F(x_{(i)}) + \ln(1 - F(x_{(n+1-i)}))]$ où $x_{(k)}$ sont les observations ordonnées.

Remarques sur les tests d'adéquation basés sur la fonction de répartition empirique

- La distribution de la statistique de test ne dépend pas de la distribution candidate. Par exemple, la statistique de Kolmogorov, de Cramer-von Mises ou d'Anderson-Darling ont leur distribution qui dépend de n (la taille de l'échantillon), mais ne dépend pas de la distribution testée, lorsque celle-ci est complètement déterminée.
- Les tests d'adéquation basés sur la fonction de répartition empirique ont été adaptés, assez difficilement, aux distributions discrètes. Cependant, ces derniers ne sont pas fréquemment utilisés.
- La comparaison des puissances de ces tests est une question complexe. Néanmoins une certaine hiérarchie s'est introduite. Globalement, le test de Cramer-von Mises est plus puissant que le test de Kolmogorov-Smirnov. Les statisticiens semblent s'accorder sur le fait qu'en l'absence de paramètre à estimer, le test d'Anderson-Darling est de tous, le test le plus puissant.

Programme SAS [7.1.] : Tests de normalité des résidus

```
*****  
*** Test de normalité des résidus *****  
*****  
%macro TestNormalite(Table,Reponse);  
proc univariate DATA=&Table. normal /*(mu=0 sigma=1)*/ plot;  
    var &Reponse.;  
    *weight i_en; /*Pour les test de normalité ôter la pondération */  
run;  
%TestNormalite(Table,ResBrut);  
  
Reponse= ResBrut, ResPerson, ResStudent, ... ;  
  
***** Sorties du programme *****;
```



Graph A. 7. Sortie SAS dans le cadre de l'étude de la normalité des résidus de la fréquence

Constats :

- L'interprétation d'un tel graphique est relativement délicate. Cette simple sortie justifie, le choix de réaliser les graphiques sur Excel, via un tirage aléatoire des données.
- Sur cette sortie, 3536 lignes (et donc résidus) ne sont pas présentes. Il s'avère que des modalités étaient manquantes pour ces lignes. Aucune valeur prédite par le modèle ne pouvait, dans ce cas, apparaître.

Annexe [7.2.1] : Tests d'indépendance des résidus¹⁵

Pour vérifier l'indépendance des résidus, la méthode se décompose en deux temps.

Tout d'abord, il est nécessaire d'ordonner les résidus en fonction d'un critère. Ce critère peut être une variable interne au modèle (un des facteurs de X, la variable Y peut également être utilisée) ou une variable externe (exemple : le Code Socio-Professionnel ou le sexe du souscripteur, etc...).

Dans un second temps, il est nécessaire d'effectuer un test :

- soit des signes des résidus (c'est un test non-paramétrique qui examine si l'arrangement des signes des résidus dans la séquence est aléatoire, anormalement groupé ou, anormalement fluctuant),
- soit de corrélation entre résidus successifs dans la séquence ordonnée. Ce test consiste à calculer la corrélation entre les résidus et ces mêmes résidus décalés d'un pas dans la séquence. Si la corrélation est significative, alors les résidus ne sont pas indépendants.

Programme SAS [7.2.1] : Tests d'indépendance des résidus

```
*****  
*** Test de corrélation (<=> d'indépendance) des résidus ;  
*****  
%macro TestCorrel(tabentree,tabsortie,ListeVariables);  
  
/* Tout d'abord trier la table des résidus selon une variable quelconque */  
proc sort data=&tabentree. out=&tabsortie.; by &ListeVariables.; run;  
%mend;  
  
/* Création d'une nouvelle colonne des résidus décalées d'un pas */  
data =&tabsortie.; set =&tabsortie.;  
    retain Var_retain;  
    if _N_=1 then do; Var_retain=ResBrut; end;  
    else do; Var_new=var_retain; Var_retain=ResBrut; end;  
run;  
  
/* Test de corrélation entre les deux colonnes */  
Proc corr data=test_retain outp=Res_corr;  
*BY &ListeVariables.; VAR ResBrut var_new; WEIGHT i_en;  
run;  
  
%TestCorrel(tabentree, tabsortie, ListeVariables);  
ListeVariables = CDGEDL /* CSP */;
```

¹⁵ Cette annexe s'appuie sur le polycopié de cours de Géostatistique (2011) de MARCOTTE Denis, Ecole Polytechnique de Montréal

Annexe [7.3.1] : Tests d'homogénéité des variances¹⁶

Pour analyser l'homoscédasticité (ou l'homogénéité des variances) des résidus, il est nécessaire d'ordonner ces résidus selon un critère (comme énoncé en *Annexe [7.2.]*) et de vérifier que ces derniers montrent des variations de même amplitude pour toute la séquence ordonnée (si ce n'est pas le cas, il est possible de corriger la situation à l'aide de transformations telles le logarithme ou la racine carrée qui ont habituellement pour effet de stabiliser la variance).

La détection de données aberrantes s'effectue en considérant les résidus qui s'écartent le plus de zéro. Par principe, les résidus situés à plus de trois écarts-types sont suspects et doivent être examinés.

Si des erreurs sont responsables de ces valeurs élevées, il est nécessaire de les éliminer et reprendre la régression.

Si aucune erreur ne peut les expliquer, alors il faut, soit chercher à affiner le modèle pour mieux expliquer ces données, soit chercher de nouvelles observations avec les mêmes valeurs de X pour en vérifier la validité.

La détection de tendances particulières dans les données se fait en reportant, sur des diagrammes binaires, les résidus en fonction de chacune des variables X . Ces diagrammes, entre les résidus et des variables externes, peuvent suggérer l'inclusion de nouvelles variables ou la transformation de variables existantes dans le modèle, afin d'en améliorer la performance.

Certaines techniques majeures, comme ANOVA, supposent explicitement que les groupes d'observations considérés soient issus de populations normales indépendantes et de variances identiques. Afin que les données ne soient pas en contradiction avec cette hypothèse, il est utile de les soumettre à un test d'homogénéité des variances (comme par exemple, les tests de Bartlett, de Levene ou de Brown-Forsythe).

Un test d'homogénéité des variances met en concurrence :

- l'hypothèse nulle $H_0 : \sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_k (= \sigma^2)$,

et

- l'hypothèse alternative $H_1 : \sigma^2_i \neq \sigma^2_k$ pour au moins un i différent de k (au moins une des variances est différente des autres).

Test d'homogénéité des variances de Bartlett

Le test de Bartlett estime si les différentes sous-catégories, d'une variable de distribution normale, ont la même variance. Ce test est très sensible si la distribution n'est pas normale.

1) Il est possible de remarquer une similarité formelle entre ces hypothèses et celles d'ANOVA, le terme « moyennes » étant remplacé par le terme « variances ».

2) Quand $k = 2$, le test d'homogénéité des variances est simplement le test F .

Soient $n_1, n_2, \dots, n_k$ les tailles respectives de groupes. La statistique de Bartlett est définie telle que $Q_{Bart} = \frac{A}{B}$ avec

$$A = \left(\sum_{i=1}^n v_i \right) \log s^2 - \sum_{i=1}^n v_i \log s_i^2$$
$$B = 1 + \frac{1}{3(k-1)} \left[\sum_{i=1}^n \left(\frac{1}{v_i} \right) - \left(\frac{1}{\sum_{i=1}^n v_i} \right) \right]$$

$v_i = n_i - 1$, s_i^2 est l'estimateur sans biais classique de σ^2_i et s^2 est l'« estimateur agrégé » de la

variance commune σ^2 sous H_0 telle que $s^2 = \frac{v_1 s_1^2 + \dots + v_n s_n^2}{v_1 + \dots + v_n} = \frac{\sum_{i=1}^n v_i s_i^2}{\sum_{i=1}^n v_i}$

Ainsi, s^2 est le barycentre des variances des différents groupes pondérées par les tailles respectives des groupes moins 1.

La partie fonctionnelle de la statistique est le numérateur A .

Le dénominateur B ne contient pas de grandeur aléatoire. C'est un « facteur de correction » destiné à rapprocher la distribution de la statistique de Bartlett de sa forme asymptotique (voir ci-dessous). Son rôle est de rendre l'espérance de la statistique de test Q_{Bart} égale à la moyenne de cette distribution asymptotique. B tend vers 1 quand les tailles des groupes tendent vers l'infini.

¹⁶ Cette annexe est inspirée des cours de BORKOWSKI John (2009), *HOV Tests, transformations, weighted least squares*, Université du Montana et du polycopié de cours de BORCARD Daniel (2008), *Tests de comparaison des variances de plus de deux échantillons indépendants*, Université de Montréal

Distribution de la statistique de Bartlett

Le numérateur A est de la forme $-2\log(\Lambda)$, où Λ est la statistique du Test du Rapport de Vraisemblance construit dans le but de tester les deux hypothèses ci-dessus. Ainsi, il apparaît que le test de Bartlett est une version améliorée du Test du Rapport de Vraisemblance, destiné à tester l'hypothèse d'homogénéité des variances de distributions normales indépendantes.

Par conséquent, la distribution asymptotique de Q_{Bart} est une distribution du Khi-Deux, où le nombre de degrés de liberté est égal à $k - 1$. Comme la distribution exacte de la statistique de Bartlett n'est pas connue, le test utilise cette loi asymptotique comme approximation de sa distribution réelle.

Faiblesses du test de Bartlett :

- La statistique de Bartlett ne suivant qu'approximativement une distribution du Khi-Deux, le test de Bartlett a de mauvaises performances sur de petits échantillons.
- Le test de Bartlett est très sensible aux écarts à la normalité des distributions considérées.

Test d'homogénéité des variances de Levene et de Brown-Forsythe

Parmi les nombreux tests d'homogénéité des variances proposés dans la littérature, le test de Levene offre une alternative intéressante au test de Bartlett. La formule du test de Levene est la suivante :

$$W = \frac{(N - g) \sum_{j=1}^g n_j (z_{.j} - z_{..})^2}{(g - 1) \sum_{j=1}^g \sum_{i=1}^{n_j} (z_{ij} - z_{.j})^2}$$

avec pour une matrice X de facteurs :

- x_{ij} = une observation quelconque i du groupe j
- N = taille de l'échantillon et n_j = effectif du groupe j
- g = nombre de groupes
- $z_{ij} = |x_{ij} - x_{i.}|$ où $x_{i.}$ est la moyenne du groupe j
- $z_{.j}$ = moyenne des z_{ij} pour chaque groupe j
- $z_{..}$ = moyenne générale des z_{ij}

Ce test est toutefois moins puissant que le test de Bartlett, surtout lorsque les effectifs sont faibles ($n_j < 15$). De nombreuses variantes existent à ce test de Levene. Le meilleur compromis entre puissance et robustesse face aux écarts à la normalité semble être le test W de Brown-Forsythe. Ce test est basé sur les écarts absolus à la médiane de chaque groupe. Pour la formule ci-dessus, le symbolisme de ce test est, cette fois-ci, modifié avec :

$$z_{ij} = |x_{ij} - \tilde{x}_{i.}| \text{ où } \tilde{x}_{i.} \text{ est la médiane du groupe } j$$

Si H_0 est vraie et que les conditions d'application sont réunies (variable y distribuée normalement et observations indépendantes) alors la statistique W est distribuée comme un F de Fisher-Snedecor à $g - 1$ et $N - g$ degrés de liberté.

De manière plus informelle, le test de Levene ou celui de Brown-Forsythe peut être envisagé comme une ANOVA sur les écarts absolus des valeurs y_{ij} à la moyenne ou la médiane de chaque groupe.

Lorsque les écarts sont plus ou moins les mêmes dans chaque groupe, les variances intragroupes sont considérées comme homogènes. Le fait d'utiliser la médiane comme paramètre de position, plutôt que la moyenne, rend le test plus robuste, dans le cas où la distribution des données est asymétrique. Ainsi, le test tolère de plus grands écarts à la normalité.

Programme SAS [7.3.] : Tests d'homogénéité des variances

```
*****;
*** Test d'homogénéité des variances (homosédasticité) ***;
*****;
%macro TestHomosedasticite(TableEntree,Reponse,VarExpli,Test);
PROC ANOVA data=&TableEntree.;
CLASS &VarExpli.;
MODEL &Reponse.=&VarExpli.;
MEANS &VarExpli. / HOVTEST=&Test.;
%mend;

%TestHomosedasticite(EtudeRes_freq_trie,ResBrut,VarExpli,Test);
Test= Bartlett, BF, Levene, Obrien;
Residu= ResBrut, ResStudent;
VarExpli= CDGEDL ;
```

Annexe [8.] : Analyses complémentaires permettant la validation de différentes étapes dans l'élaboration de la prime pure proposée

Annexe [8.1.] : Validation de la stabilité des échelles d'un exercice sur l'autre

À titre d'exemple, l'hypothèse de stabilité des échelles de risques, par exercice, est montrée sur l'échelle de fréquence et coût moyen pour le critère « nombre de pièces ».

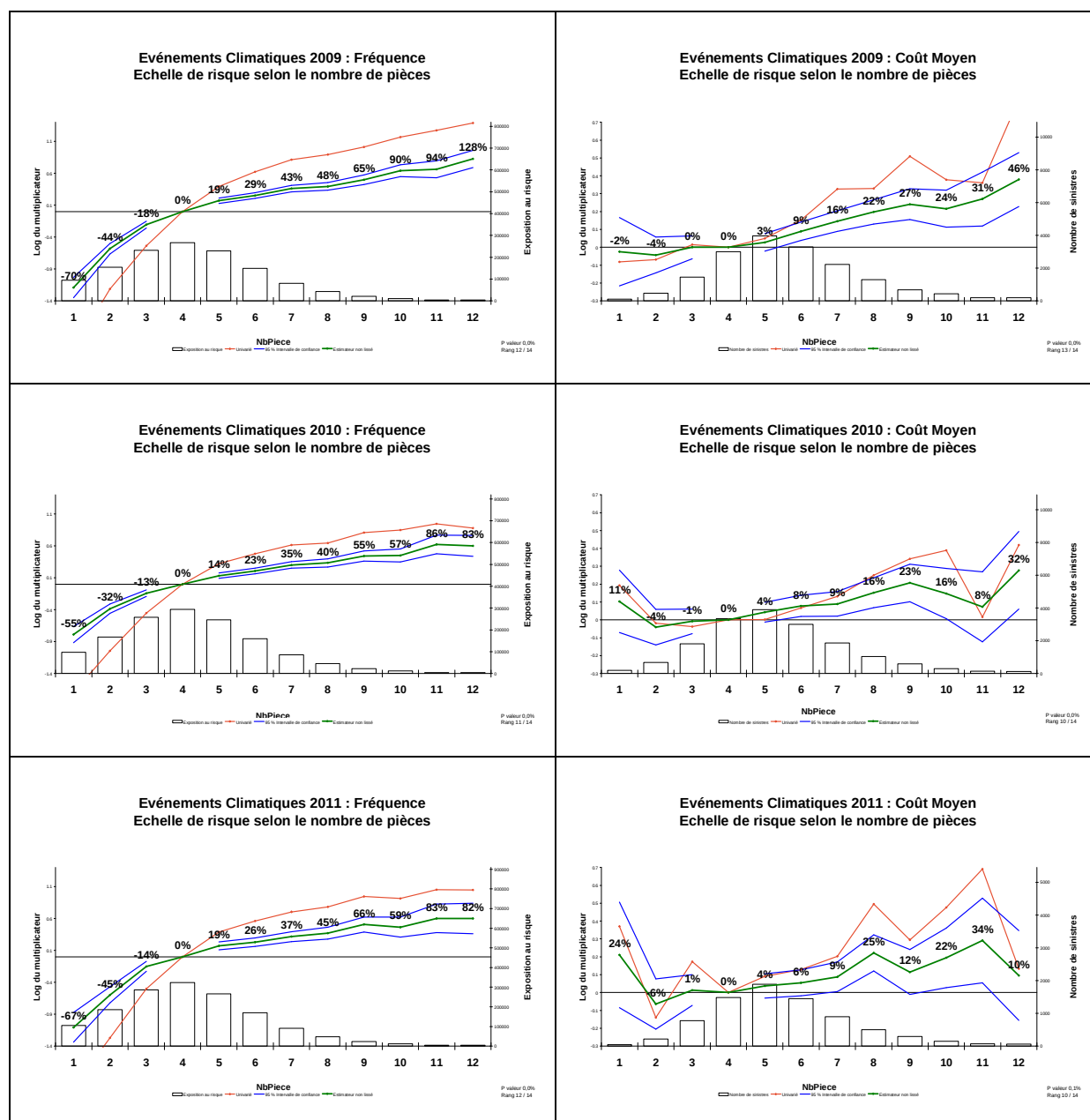


Tableau A. 4. Analyse de la stabilité des échelles de risques par exercice

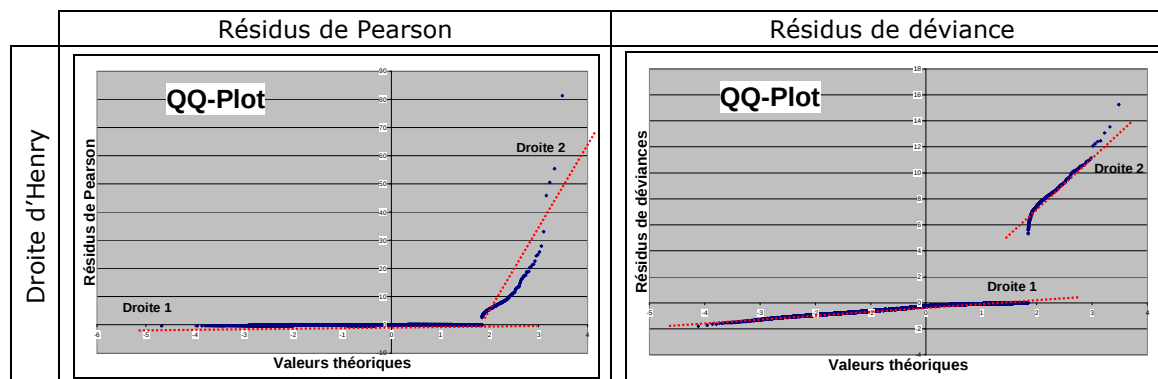
Constats :

- Dans la colonne de gauche, la modélisation de la fréquence est relativement identique d'un exercice à l'autre.
- Concernant la colonne de droite, la modélisation des coûts moyens est plus erratique (la volatilité est plus importante), même si les sinistres supérieurs à 5700€ sont exclus. Néanmoins, une certaine tendance croissante se dégage.
- Le rang de significativité du critère reste dans le même ordre de grandeur d'un exercice à l'autre.

La nécessité d'agréger et consolider plusieurs exercices afin d'obtenir une fiabilité dans le calibrage des échelles de risques, notamment au niveau de la modélisation des coûts moyens, est donc un concept important.

Annexe [8.2.1] : Analyse complémentaire sur les résidus du modèle de fréquence

Comme il l'a été signalé dans le rapport, la vérification de la normalité des résidus du modèle de fréquence n'est pas plus aisée, même avec l'utilisation des résidus de Pearson ou de déviance. Ce constat est proposé ci-dessous et focalisé essentiellement sur l'analyse de la droite de Henry.

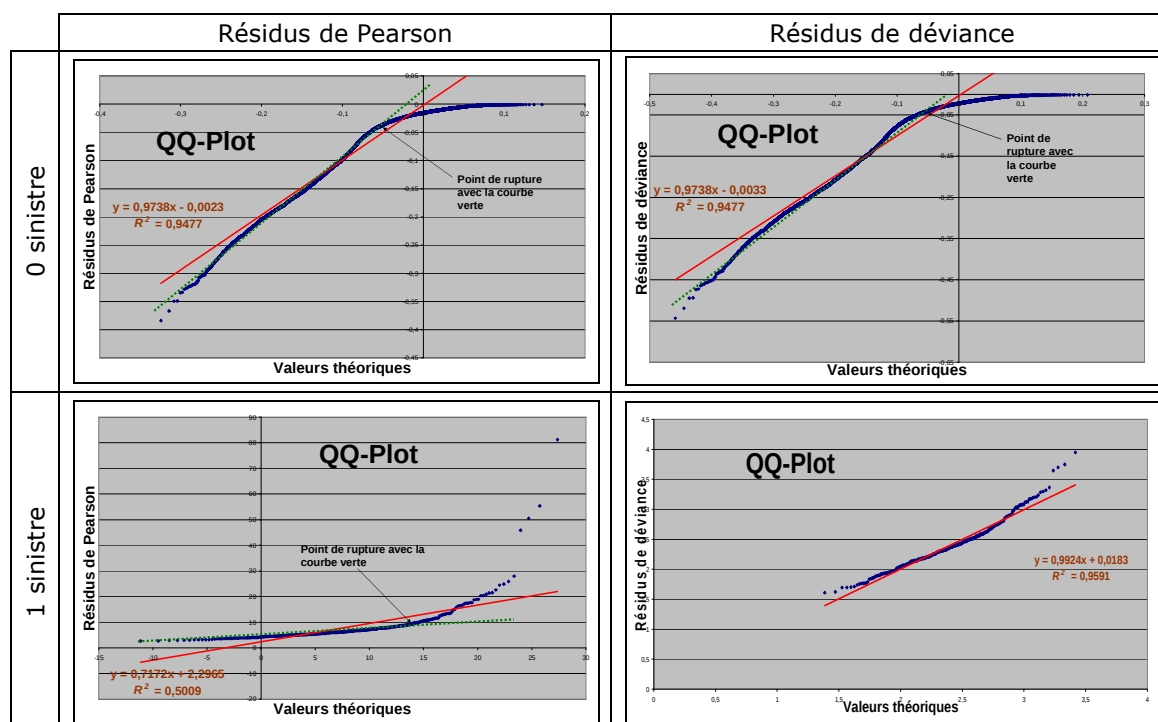


Constats :

- Pour les résidus de Pearson, la droite de Henry est scindée en 2 parties. La « Droite 1 » est relativement bien linéaire. Ce n'est pas le cas pour la « Droite 2 ».
- Pour les résidus de déviance, la droite de Henry est également scindée en 2 parties. La « Droite 1 » est là aussi relativement bien linéaire. Pour la « Droite 2 », la linéarité est plus « exacte ».

Comme pour les résidus bruts, avec de telles représentations graphiques, il est difficile d'interpréter une quelconque linéarité (et donc normalité des résidus).

Cependant, suite à ces courbes, il vient assez rapidement l'idée d'isoler les contrats ayant 0, 1 ou 2 sinistres et d'en regarder les QQ-Plot propres à chaque cas (l'hypothèse étant que la « Droite 1 » correspond au contrat sans sinistre, la « Droite 2 » à celle des contrats ayant un sinistre).



Interprétation graphique :

- Les points bleus représentent la « Droite de Henry »
- La droite rouge correspond à la droite de tendance de l'outil Excel. Il est ainsi possible de voir le coefficient de détermination R^2 .
- La droite verte est une droite tracée manuellement sur la partie de courbe la plus linéaire.

Constats :

- Sur les contrats non sinistrés, que ce soit pour les résidus de Pearson ou de déviance, la droite de tendance à un coefficient de détermination identique. Cependant, visuellement, la représentation graphique semble plus appropriée pour les résidus de déviance.
- Sur les contrats ayant eu un sinistre, les résidus de déviance ont une meilleure tendance affine que les résidus de Pearson. Pour ces derniers, le coefficient de détermination atteint la valeur de 0,9591. Dans le même temps, le coefficient de détermination de la droite de tendance des résidus de Pearson atteint la valeur de 0,5009 (valeur qui semble s'expliquer par les points supérieurs droites sortant de la courbe linéaire).

L'idée d'isoler les contrats ayant 0, 1 ou 2 sinistres et d'en regarder les *QQ-Plot* spécifiques apportent une certaine « linéarité » (et par effet « normalité »), notamment sur les résidus de déviance.

Pour aller encore un plus loin dans l'étude, il a été soulevé l'idée d'étudier le principe de loi composée (modifiée en zéro). Sans rédiger un nouveau rapport, il est proposé un rapide descriptif de ce qu'est ce principe. Les propos suivants s'inspirent du « *Projet Etienne Marceau, Méthodes statistiques en assurance non vie* » (2007) de Sauser Carine, Groisne Mélanie et Milhaud Xavier.

La modélisation du nombre de sinistres pour la garantie « événements climatiques » a une densité très forte en zéro.

Une loi de comptage modifiée en zéro consiste à définir une probabilité P_m telle que :

- pour $k = 0$, la probabilité modifiée $P_m(X = 0) = \alpha, 0 < \alpha < 1$
- pour $k \geq 1$, la probabilité modifiée $P_m(X = k) = \frac{1-\alpha}{1-P(X=0)} P(X = k), 0 < \alpha < 1$

Dans le cadre de cette étude, avec la loi de Poisson de paramètre λ , il vient

$$P_m(X = k) = \frac{1-\alpha}{1-e^{-\lambda}} \frac{\lambda^k e^{-\lambda}}{k!} \quad (\text{si dans l'étude } \lambda = 1,668\%, \text{ cette valeur doit être de nouveau réévaluée de telle sorte que } \sum_{k \geq 0} P_m(X = k) = 1).$$

Ces deux nouveaux paramètres peuvent être déterminés, soit à l'aide des méthodes des moments, soit à l'aide de la méthode du maximum de vraisemblance.

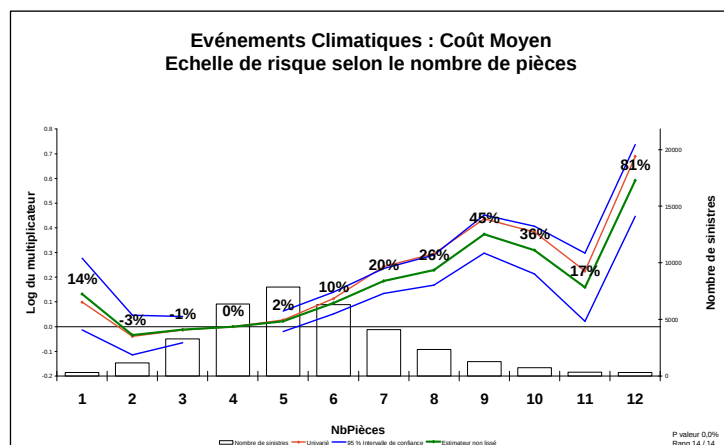
Selon ces nouveaux calibrages, l'adéquation aux données serait *a priori* meilleure et donc l'observation d'escaliers dans l'étude des résidus serait également *a priori* « atténuée ».

Une autre méthode souvent décrite dans la littérature consiste à étudier la matrice de covariance et corrélation.

De nombreux principes sont énumérés et l'étude des résidus peut relever d'un mémoire complet.

Annexe [8.3.] : Analyse des échelles et résidus des coûts moyens sans seuil d'écèlement, ni exclusion de données

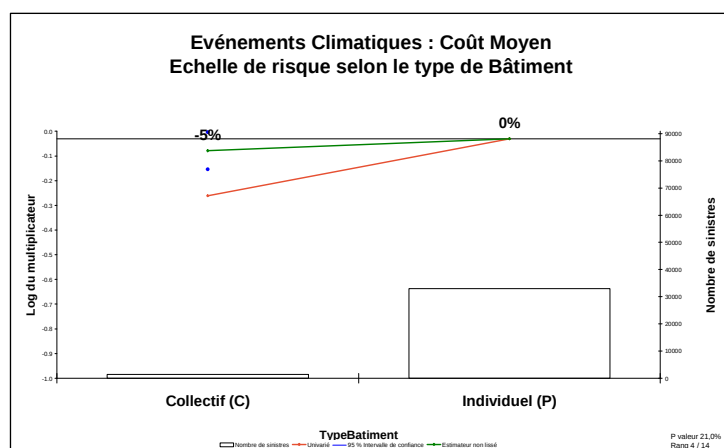
Les graphiques ci-dessous sont les résultats de la modélisation des coûts moyens, sans écèlement, ni exclusion sur la base de données composée des exercices 2009 et 2010 :



Graph A. 8. Echelle de risque des coûts moyens selon le nombre de pièces sur les données brutes

Constats :

Comme il est possible de le constater sur ce graphique, l'allure des courbes du GLM et de l'univarié restent relativement proches, voire même identiques.

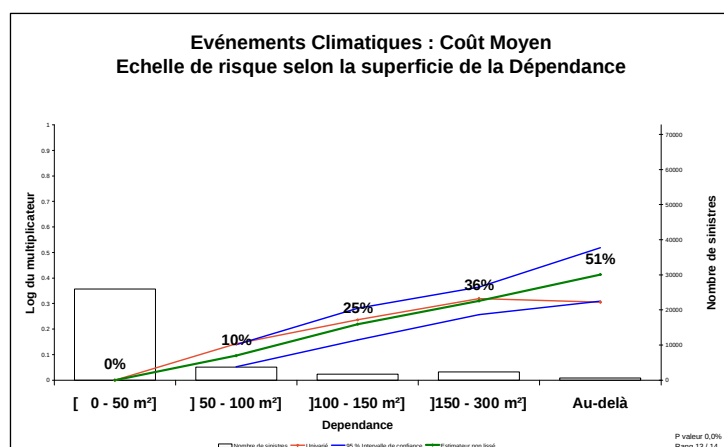


Graph A. 9. Echelle de risque des coûts moyens selon le type de bâtiment sur les données brutes

Constats :

Deux points peuvent être mis en évidence avec ce graphique :

- les habitations collectives (appartements) ont un coût moyen 6% plus faible que les habitations individuelles (maisons).
- la p-value est élevée et atteint les 21%. De ce fait, le critère n'est plus significatif.



Graph A. 10. Echelle de risque des coûts moyens selon la superficie de la Dépendance sur les données brutes

Constats :

Graphiquement, il est aisé de remarquer un coût moyen croissant selon la superficie des dépendances.

Rappel objectif	Points retenus	Etape suivante
Détermination de l'échelle de risque des coûts moyens selon les critères tarifaires sans écèlement des sinistres, ni exclusion	- échelles des coûts moyens sur le nombre de pièces relativement proche du constat en univarié - les variables ne sont plus toutes significatives et ont des caractères prédictifs différents de la fréquence	Vérifier la pertinence du modèle

Suite à cette modélisation, il est présenté ci-dessous le tableau des résultats de l'analyse des résidus :

	Sur résidus bruts	Sur résidus de Pearson																																																																																								
Hypothèse d'indépendance	Coefficients de corrélation de Pearson Proba > r sous H0: Rho=0 Nombre d'observations <table><tr><td></td><td>ResBrut</td><td>var_new ResBrut</td></tr><tr><td>ResBrut</td><td>1.00000 33567</td><td>0.04501 <.0001 33433</td></tr><tr><td>var_newResBrut</td><td>0.04501 <.0001 33433</td><td>1.00000 33566</td></tr></table>		ResBrut	var_new ResBrut	ResBrut	1.00000 33567	0.04501 <.0001 33433	var_newResBrut	0.04501 <.0001 33433	1.00000 33566	Coefficients de corrélation de Pearson Proba > r sous H0: Rho=0 Nombre d'observations <table><tr><td></td><td>Res Pearson</td><td>var_new Res Pearson</td></tr><tr><td>ResPearson</td><td>1.00000 33566</td><td>0.01669 0.0023 33431</td></tr><tr><td>var_newResPearson</td><td>0.01669 0.0023 33431</td><td>1.00000 33565</td></tr></table>		Res Pearson	var_new Res Pearson	ResPearson	1.00000 33566	0.01669 0.0023 33431	var_newResPearson	0.01669 0.0023 33431	1.00000 33565																																																																						
		ResBrut	var_new ResBrut																																																																																							
	ResBrut	1.00000 33567	0.04501 <.0001 33433																																																																																							
	var_newResBrut	0.04501 <.0001 33433	1.00000 33566																																																																																							
		Res Pearson	var_new Res Pearson																																																																																							
ResPearson	1.00000 33566	0.01669 0.0023 33431																																																																																								
var_newResPearson	0.01669 0.0023 33431	1.00000 33565																																																																																								
Hypothèse d'homoscédasticité	The ANOVA Procedure Bartlett's Test for Homogeneity of ResBrut Variance <table><tr><td>Source</td><td>DDL</td><td>Khi-2</td><td>Pr > Khi-2</td></tr><tr><td>CDGEDL</td><td>126</td><td>4267.5</td><td><.0001</td></tr></table> The ANOVA Procedure Brown and Forsythe's Test for Homogeneity of ResBrut Variance ANOVA of Absolute Deviations from Group Medians <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>123</td><td>2.8708E9</td><td>23340214</td><td>1.92</td><td><.0001</td></tr><tr><td>Error</td><td>33414</td><td>4.055E11</td><td>12135249</td><td></td><td></td></tr></table> The ANOVA Procedure Levene's Test for Homogeneity of ResBrut Variance ANOVA of Squared Deviations from Group Means <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>123</td><td>2.105E18</td><td>1.711E16</td><td>0.75</td><td>0.9842</td></tr><tr><td>Error</td><td>33414</td><td>7.67E20</td><td>2.295E16</td><td></td><td></td></tr></table>	Source	DDL	Khi-2	Pr > Khi-2	CDGEDL	126	4267.5	<.0001	Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	123	2.8708E9	23340214	1.92	<.0001	Error	33414	4.055E11	12135249			Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	123	2.105E18	1.711E16	0.75	0.9842	Error	33414	7.67E20	2.295E16			The ANOVA Procedure Bartlett's Test for Homogeneity of ResPearson Variance <table><tr><td>Source</td><td>DDL</td><td>Khi-2</td><td>Pr > Khi-2</td></tr><tr><td>CDGEDL</td><td>126</td><td>23007.5</td><td><.0001</td></tr></table> The ANOVA Procedure Brown and Forsythe's Test for Homogeneity of ResPearson Variance ANOVA of Absolute Deviations from Group Medians <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>123</td><td>966.3</td><td>7.8558</td><td>1.29</td><td>0.0172</td></tr><tr><td>Error</td><td>33413</td><td>203660</td><td>6.0952</td><td></td><td></td></tr></table> The ANOVA Procedure Levene's Test for Homogeneity of ResPearson Variance ANOVA of Squared Deviations from Group Means <table><tr><td>Source</td><td>DDL</td><td>Somme des carrés</td><td>Moyenne quadratique</td><td>Valeur F</td><td>Pr > F</td></tr><tr><td>CDGEDL</td><td>123</td><td>5016699</td><td>40786.2</td><td>0.67</td><td>0.9984</td></tr><tr><td>Error</td><td>33413</td><td>2.0477E9</td><td>61283.6</td><td></td><td></td></tr></table>	Source	DDL	Khi-2	Pr > Khi-2	CDGEDL	126	23007.5	<.0001	Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	123	966.3	7.8558	1.29	0.0172	Error	33413	203660	6.0952			Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F	CDGEDL	123	5016699	40786.2	0.67	0.9984	Error	33413	2.0477E9	61283.6		
	Source	DDL	Khi-2	Pr > Khi-2																																																																																						
	CDGEDL	126	4267.5	<.0001																																																																																						
	Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																																																																				
	CDGEDL	123	2.8708E9	23340214	1.92	<.0001																																																																																				
Error	33414	4.055E11	12135249																																																																																							
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																																																																					
CDGEDL	123	2.105E18	1.711E16	0.75	0.9842																																																																																					
Error	33414	7.67E20	2.295E16																																																																																							
Source	DDL	Khi-2	Pr > Khi-2																																																																																							
CDGEDL	126	23007.5	<.0001																																																																																							
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																																																																					
CDGEDL	123	966.3	7.8558	1.29	0.0172																																																																																					
Error	33413	203660	6.0952																																																																																							
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F																																																																																					
CDGEDL	123	5016699	40786.2	0.67	0.9984																																																																																					
Error	33413	2.0477E9	61283.6																																																																																							

Tableau A. 5. Analyse des résidus du modèle de coûts moyens

Constat sur l'hypothèse d'indépendance :

Pour les deux types de résidus, les tests d'indépendance sont significatifs.

N.B. : Le test de corrélation s'effectue entre les résidus ordonnés (noté ResBrut) et eux même décalés d'un pas dans la séquence (noté var_newResBrut). Idem pour les résidus de Pearson.

Constats sur l'hypothèse d'homoscédasticité :

- Le test de Bartlett et le test de Brown & Forsythe sont significatifs pour ces deux types de résidus.

- Le test de Levene, quant à lui, rejette très fortement l'hypothèse d'homoscédasticité aussi bien pour les résidus bruts, que les résidus de Pearson (ce qui n'était pas le cas dans l'analyse du point (d) du chapitre II.1.3.).

Pour rappel, ce test utilise non pas la moyenne, mais la médiane comme paramètre de position. Il semblerait donc que pour les résidus, la médiane ne convienne pas.

Et effectivement, cette analyse est confirmée par le constat réalisé sur l'analyse de normalité des résidus.

		Sur résidus bruts	Sur résidus de Pearson
Hypothèse de normalité	Test de normalité	Mesures statistiques de base Tendance centrale Variabilité Moyenne -841.83 Ecart-type 3933 Médiane -1450.83 Variance 15468160 Mode -2225.66 Intervalle 104200 Ecart interquartile 2202 Tests de tendance centrale : $\mu_0=0$ Test -Statistique- -p-Value- t de Student t -39.2157 Pr > t <.0001 Signe M -9250.5 Pr >= M <.0001 Rang signé S -1.639E8 Pr >= S <.0001 Tests de normalité Test -Statistique- -p-Value- Kolmogorov-Smirnov D 0.197121 Pr > D <0.0100 Cramer-von Mises W-Sq 521.287 Pr > W-Sq <0.0050 Anderson-Darling A-Sq 2815.24 Pr > A-Sq <0.0050	Mesures statistiques de base Tendance centrale Variabilité Moyenne 0.00001 Ecart-type 2.50242 Médiane -0.41987 Variance 6.26210 Mode -0.66159 Intervalle 164.32017 Ecart interquartile 0.52214 Tests de tendance centrale : $\mu_0=0$ Test -Statistique- -p-Value- t de Student t 0.000373 Pr > t 0.9997 Signe M -9251 Pr >= M <.0001 Rang signé S -1.407E8 Pr >= S <.0001 Tests de normalité Test -Statistique- -p-Value- Kolmogorov-Smirnov D 0.388963 Pr > D <0.0100 Cramer-von Mises W-Sq 1533.28 Pr > W-Sq <0.0050 Anderson-Darling A-Sq 7526.961 Pr > A-Sq <0.0050
	Histogramme des résidus		
	Droite de Henry		

Tableau A. 6. Analyse des résidus du modèle de coûts moyens (suite)

Constats sur le test de normalité :

Que ce soit pour les résidus bruts ou de Pearson, les tests de Kolmogorov-Smirnov (K-S), Cramer-von Mises et Anderson-Darling sont significatifs.

Constats sur les histogrammes des résidus :

- Sur les résidus bruts, en « ôtant » l'effectif (la « barre ») de droite, l'allure de l'histogramme pourrait s'approcher à une distribution normale. Les résidus de Pearson donnent une toute autre interprétation. La réduction de la variance a un effet direct sur la représentation des résidus.
- Sur ces deux histogrammes, la barre de droite représente 152 sinistres.
- De façon rétroactive, cela correspond à un quantile de 24K€. L'écèlement devait donc être au minimum à ce niveau. Avec le tableau de répartition, l'écèlement envisagé serait, soit aux alentours de 10K€, soit des 5,7K€. Les 2 valeurs ont été testées. La significativité était plus importante au seuil de 5,7K€.

Constats sur la Droite de Henry :

- Ni l'une, ni l'autre des deux courbes, ne représentent visuellement des droites.
- Les droites de tendances (en rouge), calculées par Excel, permettent d'obtenir un R^2 de 0,86 pour les résidus bruts et de 0,79 pour les résidus de Pearson.
- Avec de telles valeurs, ce sont les points extrêmes qui posent problème.

Rappel objectif	Points retenus
Vérifier la pertinence du modèle	- L'exclusion des valeurs supérieures à une certaine normalité doit être réalisée pour avoir une meilleure appréciation et modélisation du risque - Le seuil des 5700€ a été défini suite à un ensemble d'études et a permis d'établir les échelles de primes pures modélisées dans ce rapport.

Annexe [9.] : Rappel des différents types de traités de réassurance

Cette annexe est loin d'être exhaustive. Elle s'appuie essentiellement sur le polycopié¹⁷ de cours dispensés au sein de l'ISFA par Pierre JAL.

Il existe nombre de traités en réassurance. Dans cette annexe, il est uniquement présenté les 4 principaux. A partir de ceux-ci, de nombreux programmes sont construits. Ils peuvent être scindés en 2 produits : les traités dits proportionnels et ceux dits non proportionnels.

Traités proportionnels

Les deux formes classiques de traités de réassurance proportionnelle sont la « Quote-Part » et l'« Excédent de Plein » :

- Dans un traité en Quote-Part, un taux de cession $x\%$ est défini. En contre partie d'une cession de $x\%$ des cotisations ou primes perçues par la cédante, le réassureur s'engage à verser $x\%$ des sinistres à la cédante.
Cette forme de traité est particulièrement intéressante pour les petites structures ou dans le cadre d'un lancement d'une nouvelle branche. Néanmoins, elle ne permet pas un lissage des résultats dans le temps. En effet, le ratio S/C (ou S/P) demeure le même quelle que soit la valeur de $x\%$.
- Dans un traité en Excédent de Plein, un taux de cession de $x_i \%$ est défini pour chaque risque i assuré par la cédante selon le calcul suivant :

$$x_i = \min\left(\frac{C}{K_i}; \max\left(\frac{K_i - R}{K_i}; 0\right)\right)$$

R : Le plein de rétention défini comme étant le montant que la Cédante souhaite payer par sinistre.
 K_i : Le capital assuré i (ou SMP : Sinistre Maximum Probable)
 C : La capacité de l'excédent de plein (un multiple de R)

L'Excédent de Plein est souvent déséquilibré pour le réassureur qui doit faire face à un engagement important pour un niveau de cotisation relativement faible. En revanche, la Cédante, protégée par un excédent de plein, bénéficie d'un portefeuille plus équilibré.

Traités non-proportionnels

Comme le nom l'indique, il n'y a aucune relation de proportionnalité entre la prime cédée et la participation du réassureur dans les sinistres. Ce type de réassurance a pour but de protéger la cédante de la survenance d'événements catastrophiques. Il existe là encore deux types de contrats en réassurance non proportionnelle : l'« Excédent de sinistres » et l'« Excédent de perte annuelle ».

- L'Excédent de sinistres (abrégié « XS », dit également « *excess of loss* ») peut fonctionner par risque, par événement ou par risque et événement. Ce traité fonctionne comme un contrat d'assurance avec franchise déductible. Ici, la franchise est appelée « priorité ». Le réassureur s'engage donc à payer pour tous les sinistres dépassant cette franchise et uniquement pour le montant de ce dépassement. De plus, ces traités font également mention d'une « portée » qui correspond à l'engagement maximum du réassureur sur un sinistre. Si S est la charge de sinistre de l'assureur, alors la part de sinistre restant à la cédante est exprimée comme suit :

$$S = \min(\text{portée}; \max(X_i - \text{priorité}; 0)) \quad \text{avec } X_i \text{ la charge sinistre à l'origine}$$

- L'Excédent de perte annuelle (« *stop loss* ») a un principe identique à celui de l'XS, à cela près que ce type de réassurance permet de protéger le S/C (ou S/P) de l'année.
Ce genre de protection est utilisé classiquement dans le cadre des risques de grêles, les dégâts des eaux ou le gel.
Si S est la charge de sinistre de l'assureur, alors la part de sinistre restant à la cédante est exprimée comme suit :

$$S = P * \min\left(\text{portée}; \max\left(\frac{\sum X_i}{P} - \text{priorité}; 0\right)\right) \quad \text{avec } X_i \text{ la charge sinistre à l'origine}$$

¹⁷ JAL Pierre (année universitaire 2010-2011) *Réassurance Non Vie*, Cours ISFA 3^{ème} année

Annexe [10.] : Mécanisme des primes de reconstitution au sein du Groupe

Les primes versées aux différents réassureurs ont été fixées à 10M€ pour des raisons de confidentialité.

Prime de reconstitution	0%	0%	25%	100%	100%	100%	100%	100%	100%	100%
Nb de reconstitution	0	0	0	2	2	2	1	1	0	1

Extrait du doc RMS

Vision Groupe		Avant réassurance		Après réassurance	
probabilité	période de retour	AEP	OEP	AEP	OEP
0,02%	5000	6023,25	5912,41	3993,25	3882,41
0,05%	2000	4774,75	4504,76	2615,56	2474,76
0,10%	1000	3958,85	3789,78	1830,91	1759,78
0,20%	500	3304,61	3158,93	1203,84	1125,64
0,40%	250	2567,1	2399,85	492,04	369,85
0,50%	200	2301,78	2159,08	425,9	150
0,57%	175	2177,67	2034,8	404,78	150
0,67%	150	2036,22	1887,95	386,88	150
0,80%	125	1850,42	1680,7	370,71	150
1,00%	100	1618,01	1488,83	351,49	150
1,33%	75	1359,69	1220,82	329,73	150
2,00%	50	1063,11	937,05	304,37	150
4,00%	25	659,54	553,76	255,78	150
10,00%	10	306,15	223,45	232,7	150

Coût fixe de la réass en million d'euro

10	10	10	10	10	10	10	10	10	10
XS GA	XS GSJC	XS GSJ	Traité 1	Traité 2	Traité 3	Traité 4	Traité 5	Traité 6	Traité 7

Coût des reconstitutions selon l'évènement (hors AAD de 50M€, inclus dans l'OEP)

0	0	2,5	20,0	20,0	20,0	10,0	10,0	0,0	10,0
0	0	2,5	20,0	20,0	20,0	10,0	10,0	0,0	10,0
0	0	2,5	20,0	20,0	20,0	10,0	10,0	0,0	10,0
0	0	2,5	20,0	20,0	20,0	10,0	10,0	0,0	10,0
0	0	2,5	20,0	20,0	20,0	10,0	10,0	0,0	10,0
0	0	2,5	20,0	20,0	20,0	10,0	10,0	0,0	9,0
0	0	2,5	20,0	20,0	20,0	10,0	10,0	0,0	2,7
0	0	2,5	20,0	20,0	18,2	10,0	10,0	0,0	0,0
0	0	2,5	20,0	20,0	12,3	10,0	4,0	0,0	0,0
0	0	2,5	20,0	20,0	10,0	7,9	0,0	0,0	0,0
0	0	2,5	20,0	17,4	10,0	1,6	0,0	0,0	0,0
0	0	2,5	20,0	10,0	3,9	0,0	0,0	0,0	0,0
0	0	2,5	12,8	1,8	0,0	0,0	0,0	0,0	0,0
0	0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Coût moyen en M€ 0,32 2,04 1,36 0,90 0,43 0,28 0,00 0,14

Interprétation :

- Le tableau de gauche représente les simulations extraites de RMS.
- Le tableau de droite représente, sur chaque ligne, l'impact de ces événements selon les traités du programme « Force de la Nature ».
- La ligne « coût moyen » représente l'espérance de coût des primes de reconstitutions selon la probabilité d'occurrence de chaque événement et est pris en compte dans le solde de réassurance.

Annexe [11.] : Programmes d'aide à la décision pour les différentes Entités

Programme SAS [11.1.] : Somme des échelles de primes pures

```
*****;
*** Somme des échelles de prime pure pour chaque CR *****;
*****;

proc sql;
create table SommeEchelle as
select Cais, Exercice, sum(i_en) as i_en, sum(pp_estime) as pp_estime
from Validres.Etuderres_pp_evtcli
group by Cais, Exercice;
quit;
```

Programme SAS [11.2.] : Elaboration pour la présentation des résultats

```
*****;
*** Elaboration du tableau de résultat final *****;
*****;

data Resultat_pp_evtcli; set Validres.Etuderres_pp_evtcli;
*** Application des coef. multiplicateurs à la prime pure unitaire *****;
if Cais="GCA" then do; pp_brute=pp_estime*17.14558;
pp_nette=pp_estime*13.27150; end;

if Cais="SUD" then do; pp_brute=pp_estime*12.99316;
pp_nette=pp_estime*12.97354; end;
keep Cais Exercice Qnnbp Ctnbi Limitecapitalmobilier pni_en pp_brute
pp_nette i_en ns_en s_en; run;

*** Aide à la décision pour les Entités *****;

proc sql;
create table resume as
select Cais, Exercice, Qnnbp, Ctnbi, Limitecapitalmobilier,
sum(i_en) as i_en, sum(ns_en) as ns_en,
sum(pp_brute) as pp_brute, sum(pp_nette) as pp_nette,
sum(pni_en) as pni_en
from Resultat_pp_evtcli
group by Cais, Exercice, Qnnbp, Ctnbi, Limitecapitalmobilier;
quit;
```

Annexe [12.] : Comparaison entre le nouveau calibrage de la prime pure et les primes pures brutes observées en 2011 (« Back-Testing »)

Comme il l'a été mentionné, il est présenté ci-dessous une comparaison entre le nouveau calibrage de la prime pure et les primes pures brutes observées en 2011 :

Nombre de pièces	Collectif	Individuel	Total	poids
1	-57%	-77%	-76%	6.9%
2	-75%	-76%	-76%	12.2%
3	-45%	-57%	-57%	18.8%
4	-56%	-64%	-64%	21.4%
5	-39%	-56%	-56%	17.5%
6	-82%	-56%	-57%	11.2%
7	-65%	-56%	-56%	6.0%
8	-30%	-48%	-47%	3.2%
9	-50%	-59%	-59%	1.5%
10	-100%	-52%	-52%	0.8%
11	-55%	-19%	-20%	0.3%
12	-100%	-66%	-66%	0.3%
Total	-59%	-58%	-58%	
poids	26.4%	73.6%		

Tableau A. 7. Comparaison entre le nouveau calibrage et la prime pure réelle 2011

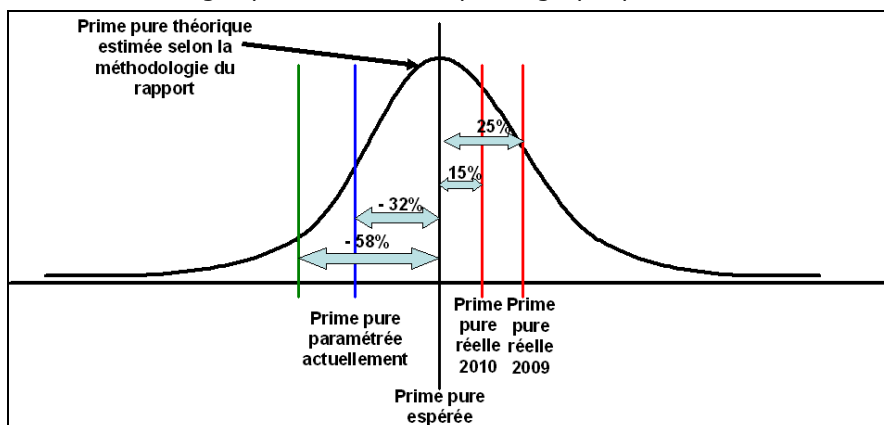
Interprétations du graphique :

- Les valeurs en rouge indiquent que les primes pures observées sont supérieures au nouveau calibrage de la prime pure (les valeurs sont supérieures à 0%, c'est le cas d'une « sous-estimation »).
- Les valeurs en noir indiquent que les primes pures observées sont en phase avec le profil de risque à -5% ou +5%.
- Les valeurs en vert indiquent que les primes pures observées sont inférieures au nouveau calibrage de la prime pure (les valeurs sont inférieures à 0%, c'est le cas d'une « sur-estimation »).

Constat lié au Tableau A.7 :

Ce « Back Testing » démontre que le nouveau calibrage de la garantie « événements climatiques » réalisé dans ce rapport est 58% supérieur aux primes pures brutes réelles 2011. C'est un point également rassurant, puisque cet exercice (au même titre que 2007 et 2008) n'a pas été marqué par des événements majeurs.

L'ensemble des « Back Testing » peut être traduit par le graphique suivant :

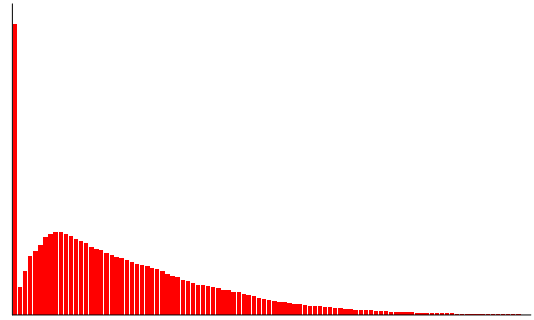


Graphique A. 11. Interprétation des événements climatiques 2009, 2010 et 2011 par rapport au nouveau calibrage de la prime pure

Annexe [13.] : La distribution Tweedie

La loi Tweedie est considérée comme une distribution appropriée lorsque la répartition des montants de sinistres a une masse en zéro, puis une distribution continue (cf. *Graphe A.11.*).

Pour une variable aléatoire $y > 0$ suivant une loi Tweedie, alors il existe des caractéristiques de masse et des paramètres qui permettent de retrouver des lois Poisson et Gamma pour des valeurs $y > 0$.



Graphe A. 12. Distribution d'une loi Tweedie

Avec λ paramètre d'une loi de Poisson et α, β paramètres d'une loi Gamma, alors la loi Tweedie peut être décrite telle que :

$$f_{(\theta, \lambda, \alpha)}(y) = \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \cdot \frac{\beta^{n\alpha} y^{n\alpha-1} e^{-\beta y}}{\Gamma(n\alpha)} \quad \text{pour } y > 0$$

La famille des distributions Tweedie est une sous classe des familles exponentielles avec une fonction de variance donnée par μ^p (où μ représente la moyenne et p est supérieur ou égal à 0).

Plus précisément, la distribution aura pour moyenne $\mu = \frac{\lambda\alpha}{\beta}$ et $\psi\mu^p = \frac{\lambda\alpha(\alpha+1)}{\beta^2}$

Ainsi, les distributions Normale ($p = 0$), Poisson ($p = 1$ avec $\varphi = 1$), Gamma ($p = 2$) et Gaussienne inverse ($p = 3$) sont des cas spéciaux de la distribution Tweedie.

Pour d'autres valeurs de puissance non remarquables, les distributions sont toujours définies mais ne peuvent pas être écrites dans une forme finie, et, par conséquent, il est très difficile de les estimer. Cependant :

- quand $1 < p < 2$ et $Y > 0$, les distributions sont continues ;
- lorsque $1 < p < 2$ et $Y = 0$, la distribution est discrète ;
- pour $p > 2$ et $Y > 0$, les distributions sont continues.

Le choix de p peut se faire simplement en analysant les résidus.

La variable dépendante doit être numérique, avec des données supérieures ou égales à zéro. Si une donnée est inférieure à zéro ou manquante, alors l'observation correspondante n'est pas utilisée dans l'analyse. La valeur fixe du paramètre de la distribution Tweedie peut être n'importe quelle valeur supérieure à 1 et inférieure à 2.

Cette loi est utile lorsqu'il est nécessaire d'ajuster un modèle sur le coût total.

Les résultats sont similaires à l'approche traditionnelle, à condition toutefois que les effets des modèles de fréquences et de coûts moyens sont, soit très significatifs, soit très faibles.

Le cas contraire, il existe un biais si les effets de fréquences et montants ne sont pas significatifs.

Avec ce type de modélisation, il est difficile de savoir si les effets résultent de la fréquence ou des coûts moyens. Le lissage en est donc plus délicat.

Annexe [14.] : Les provisions pour égalisation, spécificité de la réglementation française

Dans le cadre du caractère prudentiel énoncé au chapitre III.3.1., la France a mis en place une dotation spécifique appelée « Provision pour égalisation ». Cette provision est intéressante car elle permet de lisser les résultats sur plusieurs exercices. Elle est définie dans l'Article R.331-6 du Code des Assurances et s'exprime comme suit :

La provision pour égalisation est une franchise d'impôt qui peut être constituée aussi bien par les entreprises d'assurances que de réassurances selon trois cas de figure.

Cette provision est destinée à :

- faire face à des charges exceptionnelles afférentes aux opérations garantissant :
 - les risques dus à des éléments naturels,
 - les risques de responsabilité civile dus à la pollution,
 - le risque atomique,
 - les risques spatiaux,
 - les risques liés au transport aérien, et
 - les risques liés aux attentats et au terrorisme ;
- compenser, en assurance-crédit, la perte technique éventuelle apparaissant en fin d'exercice ;
- réduire les fluctuations de sinistralité, afférentes aux opérations d'assurance de groupe, contre les risques de dommages corporels.

Ces provisions étant des franchises d'impôt, elles sont calculées et fixées par divers articles dont l'« Article 39 quinquies G » du Code général des impôts.

Les limites dans lesquelles ces dotations annuelles peuvent être retranchées des bénéfices et celles du montant global de chaque provision sont fixées par décret (respectivement en fonction de l'importance des bénéfices techniques et du montant des primes ou cotisations, nettes de réassurances, de la catégorie de risques concernée).

Chaque provision est affectée, dans l'ordre d'ancienneté des dotations annuelles, à la compensation des résultats techniques déficitaires de l'exercice, par catégorie de risques correspondante.

Les dotations annuelles qui, dans un délai de 10 ans n'ont pu être utilisées, sont rapportées au bénéfice imposable de la 11^{ème} année.

Pour les risques « attentat » et « terrorisme », le délai est de 12 ans et les dotations sont rapportées au bénéfice imposable de la 13^{ème} année.

Pour les risques « transport aérien », le délai est de 15 ans, et le report s'effectue la 16^{ème} année.

A ce jour, ce caractère prudentiel peut encore être constitué dans le droit et normes comptables français. Cette souplesse permet de lisser les résultats au cours du temps, notamment lors d'événements importants, tels que Xynthia ou Klaus.

A l'inverse, les normes IFRS mises en place depuis 2005 ne permettent pas ce type de provisions. C'est le principe de la juste valeur. En conséquence, les résultats peuvent fluctuer dans des proportions importantes. Il en sera de même avec Solvabilité 2.

Annexe [15.] : Triangle de liquidation

Les triangles de liquidation (ou « cadencement », selon la littérature) permettent, pour différents métiers, d'obtenir une vision à l'ultime des coûts globaux d'indemnisations.

Ces triangles sont notamment utilisés dans le calcul des « Provisions pour Sinistres A Payer » (PSAP). D'un point de vue explicatif, cette provision est calculée dans le cadre des sinistres survenus à la date d'inventaire, mais qui n'ont pas encore été réglés. Ces sinistres se regroupent en deux catégories :

- Les sinistres déjà déclarés à l'assureur, mais dont les règlements n'ont pas commencé ou ne sont pas terminés,
- Les sinistres déjà survenus, mais qui n'ont pas encore été déclarés (appelé « IBNR » « *Incurred But Not Reported* »).

Les PSAP se calculent exercice par exercice comme le stipule le Code des assurances et sont généralement estimés au moyen de méthodes :

- déterministes : « Chain-Ladder », « London-Chain », « London Pivot », « Bornhuetter-Fergusson » (avec l'utilisation d'information externe au triangle) ;
- stochastiques : Méthode Mack, GLM, Bootstrap, séries chronologiques, ...

Ces triangles de liquidation sont la base des modèles déterministes et stochastiques.

A noter que la « vie » d'un sinistre est constituée de diverses dates, dont les dates de :

- survenance : date à laquelle s'est produit l'évènement ;
- déclaration : date à laquelle l'assuré informe la compagnie d'assurance ;
- paiement : date(s) de règlement(s) de diverses prestations ;
- clôture : date à laquelle l'assureur est censé avoir définitivement réglé les prestations.

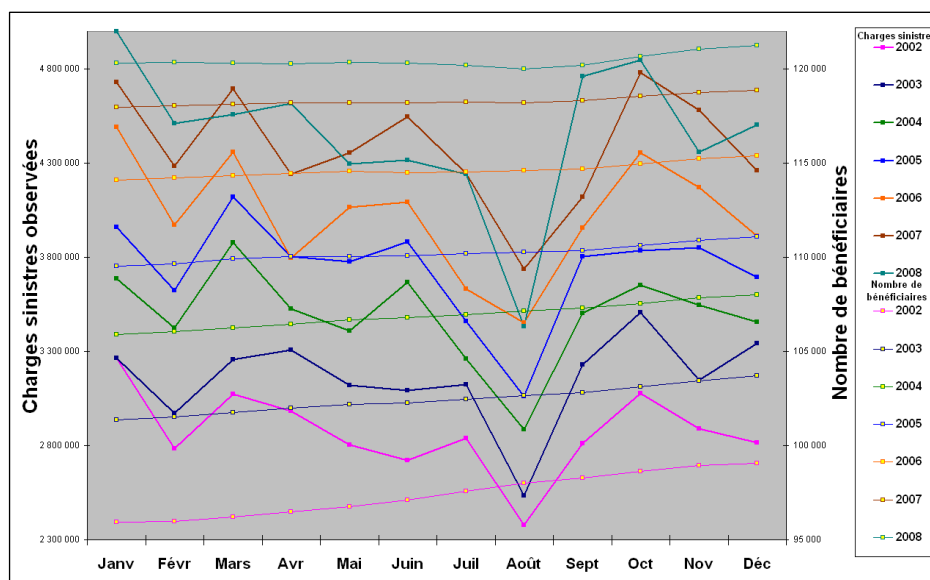
Dans un triangle de cadencement, les données sont généralement présentées sous cette forme :

		Date de développement (Année, AnMois)								
		Date 1	Date 2	Date 3	Date 4	Date 5	Date 6	Date 7	Date 8	Date 9
Date survenance (Année, AnMois)	Date 1	C _{1,1}	C _{1,2}	C _{1,3}	C _{1,4}	C _{1,5}	C _{1,6}	C _{1,7}	C _{1,8}	C _{1,9}
	Date 2	C _{2,1}	C _{2,2}	C _{2,3}	C _{2,4}	C _{2,5}	C _{2,6}	C _{2,7}	C _{2,8}	
	Date 3	C _{3,1}	C _{3,2}	C _{3,3}	C _{3,4}	C _{3,5}	C _{3,6}	C _{3,7}		
	Date 4	C _{4,1}	C _{4,2}	C _{4,3}	C _{4,4}	C _{4,5}	C _{4,6}			
	Date 5	C _{5,1}	C _{5,2}	C _{5,3}	C _{5,4}	C _{5,5}				
	Date 6	C _{6,1}	C _{6,2}	C _{6,3}	C _{6,4}					
	Date 7	C _{7,1}	C _{7,2}	C _{7,3}						
	Date 8	C _{8,1}	C _{8,2}							
	Date 9	C _{9,1}								

Tableau A. 8. Représentation d'un triangle de liquidation

Les dates de survenance et de développement ne sont pas toujours vues en « Année » (ou exercice d'inventaire). Elles peuvent également être décomposées en mensualités (nommées dans le Tableau A. 8. « AnMois »), ce qui permet dans ce cas, de produire des arrêtés trimestriels et d'observer des « comportements » dans les dérives de charges sinistres.

Ces propos peuvent être illustrés par une des premières analyses réalisées au sein de GGE, notamment sur les indemnités du risque Santé :



Graphique A. 13. Vision des indemnités Santé pour l'Entité GGE

Constats :

Un comportement cycle dans la « consommation » (les dépenses de soins) est observable, avec une charge relativement faible en août (congé médicaux et peu de maladie en saison estivale) et une charge sinistre plus importante en janvier (propagation de virus en période de fête).

En 2008, malgré la croissance du portefeuille en nombre de bénéficiaires, la charge sinistre globale de l'année s'est retrouvée à un niveau presque identique à celui de l'exercice 2007. Ce « comportement » s'explique par les points suivants :

- Groupama commercialise des contrats Santé « responsables », dans le sens où l'Entité accepte de rembourser les bénéficiaires assurés, si la Sécurité Sociale rembourse dans un premier temps, une partie des soins. En 2008, la Sécurité Sociale s'est désengagée de certains médicaments dont la performance médicale avait été remise en cause par la Haute Autorité de Santé (HAS). Là où la Sécurité Sociale remboursait 15% des frais et la complémentaire mutuelle les 85% restant, le désengagement de la Sécurité Sociale a entraîné de facto un non remboursement de ses médicaments par l'Entité ;
- Les différentes réformes des gouvernements, ainsi que la prise de conscience des patients/médecins sont peut en être la cause (la prise de génériques se démocratise, par exemple) ;

- La structure du portefeuille qui subit un rajeunissement ;
- La crise (avec une frilosité des personnes à se déclarer en arrêt maladie).

Dans le cadre de ce rapport, il n'a pas été nécessaire de prendre en compte ces triangles de liquidation. La charge d'indemnisation a été identifiée grâce aux calculs réalisés par la DAG.

Annexe [16.] : Modèles économiques pour 4 lignes de métiers distincts

Pour éventuellement mieux percevoir les concepts proposés liés à l'allocation et coût du capital, il est proposé ci-dessous les mêmes modèles présentés au chapitre III.4. mais avec cette fois-ci, 4 lignes métiers.

Hypothèses de départ pour les deux modèles :

La société commercialise 4 produits, correspondant à 4 métiers au sens Solvabilité 2. À partir du modèle standard, l'entreprise a été en mesure de déterminer pour « LoB », le besoin en fonds propres nécessaires. Les hypothèses sont les suivantes :

	Métier 1 (LoB 1)	Métier 2 (LoB 2)	Métier 3 (LoB 2)	Métier 4 (LoB 4)	SCR total individuel
Hypothèses SCR	2 300	3 500	2 500	3 100	11 400
Répartition	20.2%	30.7%	21.9%	27.2%	

Tableau A. 9. Hypothèses de départ pour les modèles économiques

De plus, de par la formule standard, une diversification est à intégrer. Ainsi, le SCR total n'est pas de 11 400, mais de 8 027 (en millions d'euros).

De manière détaillée, ce calcul est établi grâce à la matrice de corrélation présentée au chapitre III.II. (Tableau 3.2.) et l'équation $SCR_{nl} = \sqrt{\sum CorrNL_{r,c} \cdot NL_r \cdot NL_c} = 8\,027$.

Par étape, il vient :

métier		SCR métier		(SCR_{nl}) ²	Règle de comptage	« Impact »
M_r	M_c	NL_r	NL_c	$CorrNL_{r,c}$		
M1	M2	2 300	3 500	4 025 000	Compter 2 fois	8 050 000
M1	M3	2 300	2 500	2 875 000		5 750 000
M1	M4	2 300	3 100	1 752 500		3 565 000
M2	M3	3 500	2 500	2 187 500		4 375 000
M2	M4	3 500	3 100	2 712 500		5 425 000
M3	M4	2 500	3 100	1 937 500		3 875 000
M1	M1	2 300	2 300	5 290 000	Compter 1 fois	5 290 000
M2	M2	3 500	3 500	12 250 000		12 250 000
M3	M3	2 500	2 500	6 250 000		6 250 000
M4	M4	3 100	3 100	9 610 000		9 610 000
$\sum CorrNL_{r,c}.NL_r.NL_c =$						64 440 000

Tableau A. 10. Etape de calcul dans le SCR final réglementaire

Enfin, dernière hypothèse à prendre en compte, l'entreprise souhaite une rentabilité économique de 10% (ROE=10%). Le résultat cible est alors de 803 (=8 027x10%).

Premier modèle économique :

Le modèle proposé ici est le plus simple et se contente de répartir proportionnellement les résultats cibles et les fonds propres réglementaires entre les 4 métiers du dessus au prorata du ratio $\frac{SCR \text{ de la formule standard}}{SCR \text{ total individuel}} = \frac{8\,027}{11\,400}$.

$$\frac{SCR \text{ de la formule standard}}{SCR \text{ total individuel}} = \frac{8\,027}{11\,400}$$

Ces propos peuvent être traduits par le tableau suivant :

SCR déduit selon formule standard	8 027			
Hypothèse ROE	10%			
Résultat cible	803			
	Métier M1	Métier M2	Métier M3	Métier M4
Hypothèses SCR	2 300	3 500	2 500	3 100
Clé de répartition des contributions	20,2%	30,7%	21,9%	27,2%
SCR réalloué selon la clé	1 620	2 465	1 760	2 183
Résultat cible réalloué selon la clé	162	246	176	218
Contribution a posteriori	20,2%	30,7%	21,9%	27,2%
ROE a posteriori	7,0%	7,0%	7,0%	7,0%

Tableau A. 11. Résultats du premier modèle de capital économique

Constats :

- La répartition du SCR standard et du résultat économique cible est répartie de manière identique selon la clé de répartition des contributions.
- Le ROE *a posteriori* est identique à chaque métier. Celui-ci n'est pas de 10% par métier, mais seulement de 7,0% (rapport entre le résultat cible et l'hypothèse de SCR initial par métier).

Deuxième modèle économique :

Le deuxième modèle économique prend en compte les effets de diversification de chaque métier dans l'exigence de rentabilité. Les résultats sont les suivants :

Métier M_i M_j M_k			Contributions ajustées		Métier concerné	Rappel SCR initial	Baisse du SCR due à la diversification
			SCR réglementaire	Contrib SCR All - x			
M1	M2	M3	6 478	1 549	M4	3 100	-50,0%
M1	M2	M4	5 880	2 147	M3	2 500	-44,8%
M1	M3	M4	6 603	1 424	M2	3 500	-59,3%
M2	M3	M4	6 464	1 563	M1	2 300	-32,0%

Tableau A. 12. Calculs des contributions ajustées

Ainsi, le modèle économique est le suivant :

	Métier M1	Métier M2	Métier M3	Métier M4	
Contribution initiale	20,2%	30,7%	21,9%	27,2%	Somme des contributions ajustées
SCR "ajustée" (effet de diversification)	1 563	1 424	1 380	1 549	5 917
Effet de diversification (baisse du SCR initial)	-32,0%	-59,3%	-44,8%	-50,0%	
Répartition des contributions ajustées	26,4%	24,1%	23,3%	26,2%	
					Somme des valeurs
SCR initial	2 300	3 500	2 500	3 100	11 400
SCR formule standard réalloué selon la clé	2 121	1 932	1 872	2 102	8 027
Résultat cible (10% sur le SCR réalloué)	212	193	187	210	803
Gains de SCR après application de la clé de répartition	-7,8%	-44,8%	-25,1%	-32,2%	
ROE <i>a posteriori</i> (calculé sur SCR initial)	9,2%	5,5%	7,5%	6,8%	
Vérification de la contribution <i>a posteriori</i>	26,4%	24,1%	23,3%	26,2%	

Tableau A. 13. Résultats du second modèle de capital économique

Constats :

- L'activité M1 occupe un poids plus important que ne le laisse croire son SCR initial. Ce métier peut être considéré, dans ce modèle, comme l'activité la plus risquée. En effet, l'effet de diversification (rapport du SCR ajusté sur le SCR initial - 1) n'est que de 32,0%. Avec ce modèle, la rentabilité exigée pour ce métier est de 9,2%, soit un taux bien supérieur à ce qui a été constaté dans le premier modèle économique.
- L'activité M2 peut être considérée comme l'activité la plus rentable. L'exigence de résultat est la moins élevée et les effets de diversification sont importants (baisse du SCR initial de plus de 59,3%).
- Plus une compagnie est diversifiée, moins l'exigence réglementaire est élevée.

Tableaux

Tableau 1. 1. Classement des assureurs en 2010 selon leurs chiffres d'affaires	9
Tableau 1. 2. Classement des assureurs en 2010 selon le nombre de contrats en portefeuille	9
Tableau 2. 1. Méthodologie de détermination des échelles de primes pures	17
Tableau 2. 2. Répartition du nombre de sinistres référencés par contrat sur la garantie "événements climatiques"	18
Tableau 2. 3. Analyse de la stabilité dans la répartition des sinistres	19
Tableau 2. 4. Taux de sinistres à montant nul sur les exercices 2009 et 2010	19
Tableau 2. 5. Statistiques d'indépendance sur la base	20
Tableau 2. 6. Statistiques d'indépendance sur la base totale	21
Tableau 2. 7. Comparaison effectif théorique / effectif empirique sur la base totale	21
Tableau 2. 8. Statistiques d'indépendance sur un des échantillons test	21
Tableau 2. 9. Comparaison effectif théorique / effectif empirique sur un des échantillons test	21
Tableau 2. 10. Point d'étape dans l'analyse et la détermination de la prime pure	22
Tableau 2. 11. Matrice des critères tarifaires du modèle actuel	23
Tableau 2. 12. Répartition des sinistres selon les probabilités empiriques, de Poisson et Binomiale Négative	25
Tableau 2. 13. Test de significativité des variables pour le modèle poissonien	26
Tableau 2. 14. Test de significativité des variables pour le modèle Binomiale Négatif	26
Tableau 2. 15. Test d'ajustement pour le modèle de régression Poisson	27
Tableau 2. 16. Test d'ajustement pour le modèle de régression Binomiale Négatif	27
Tableau 2. 17. Échelles de risques de la fréquence	28
Tableau 2. 18. Analyse des résidus pour le modèle de fréquence	30
Tableau 2. 19. Analyse des résidus pour le modèle de fréquence (suite)	31
Tableau 2. 20. Point d'étape dans l'analyse et la détermination de la prime pure	32
Tableau 2. 21. Répartition des sinistres « événements climatiques » sur les exercices 2009 et 2010	33
Tableau 2. 22. Table de correspondance de lois a priori conjuguées	35
Tableau 2. 23. Analyse des résidus pour le modèle de coût moyen	39
Tableau 2. 24. Analyse des résidus pour le modèle de coût moyen (suite)	40
Tableau 2. 25. Récapitulatif de la méthodologie utilisée dans la détermination de l'échelle de primes pures	41
Tableau 2. 26. Exemple de sortie du logiciel RMS pour deux Entités	46
Tableau 2. 27. Résultat des modélisations RMS pour GCA	47
Tableau 2. 28. Résultat des modélisations RMS pour GSUD	47
Tableau 2. 29. Résultat des modélisations RMS pour l'ensemble des Caisses Régionales (hors GLBR)	47
Tableau 2. 30. Résultat des modélisations après application des programmes de réassurance pour GCA	51
Tableau 2. 31. Résultat des modélisations après application des programmes de réassurance pour GSUD	51
Tableau 2. 32. Résultat des modélisations après application des programmes de réassurance pour les CR (Hors GLBR)	51
Tableau 2. 33. Répartition des charges sinistres espérées selon le péril	53
Tableau 2. 34. Répartition des charges sinistres espérées selon le produit MRH	53
Tableau 2. 35. Mécanisme de calcul des coefficients multiplicateurs pour chaque Entité	54
Tableau 2. 36. Comparaison entre le nouveau calibrage et la prime pure réelle 2009	55
Tableau 2. 37. Comparaison entre le nouveau calibrage et la prime pure réelle 2010	55
Tableau 2. 38. Comparaison entre le nouveau calibrage de la prime pure et l'ancien paramétrage	57
Tableau 3. 1. Matrice de corrélation extraite du QIS5	65
Tableau 3. 2. Matrice de corrélation du risque Non Vie extraite du QIS5	66
Tableau 3. 3. Table des volatilités de primes et de réserves dans le cadre du modèle standard	66
Tableau 3. 4. Matrice de corrélation extraite du QIS5 pour les métiers non-vie	68
Tableau 3. 5. Table de comparaison entre le modèle standard et le modèle interne	70
Tableau 3. 6. Evaluation de la rentabilité d'un produit non-vie	71
Tableau 3. 7. Evolution des indicateurs de performances selon des modifications du solde de réassurance	72
Tableau 3. 8. Scénarios complémentaires selon les modifications du solde de réassurance	73
Tableau 3. 9. Hypothèses de départ pour les modèles économiques	75
Tableau 3. 10. Etape de calcul dans le SCR final réglementaire	75
Tableau 3. 11. Résultats du premier modèle de capital économique	76
Tableau 3. 12. Calculs des contributions ajustées	76
Tableau 3. 13. Résultats selon le second modèle de capital économique	76
Tableau 4. 1. Classement des mouvements de terrain selon leurs caractéristiques dynamiques	89
Tableau 4. 2. Echelle d'intensité macrosismique définie par le Bureau Central Sismologique Français	90

Tableau A. 1. Énumération d'évènements « Cat. Nat. » recueillis à partir de l'application GASPAR	103
Tableau A. 2. Panorama des différents types d'indice d'inflation liés au coût du bâtiment.....	104
Tableau A. 3. Abréviation des Caisses Régionales.....	105
Tableau A. 4. Analyse de la stabilité des échelles de risques par exercice.....	124
Tableau A. 5. Analyse des résidus du modèle de coûts moyens.....	128
Tableau A. 6. Analyse des résidus du modèle de coûts moyens (suite).....	129
Tableau A. 7. Comparaison entre le nouveau calibrage et la prime pure réelle 2011	132
Tableau A. 8. Représentation d'un triangle de liquidation	135
Tableau A. 9. Hypothèses de départ pour les modèles économiques.....	136
Tableau A. 10. Etape de calcul dans le SCR final réglementaire	136
Tableau A. 11. Résultats du premier modèle de capital économique.....	136
Tableau A. 12. Calculs des contributions ajustées	137
Tableau A. 13. Résultats du second modèle de capital économique	137

Graphes

Graphe 1. 1. Evolution des charges sinistres par garantie sur les Caisses Régionales (hors GLBR) en million d'euros.	10
Graphe 1. 2. Répartition de la charge sinistre par garantie des exercices 2007 à 2010.....	10
Graphe 1. 3. Nombre de sinistres recensés par IRIS* pour la garantie "Evènements Climatiques" sur l'exercice 2009.	12
Graphe 1. 4. Nombre de sinistres recensés par IRIS pour la garantie "Evènements Climatiques" sur l'exercice 2010.	12
Graphe 1. 5. Graphe de synthèse des différentes définitions	14
Graphe 2. 1. Représentation graphique des sinistres selon les probabilités empiriques, de Poisson et Binomiale Négative	25
Graphe 2. 2. Echelle de risque de la fréquence selon le nombre de pièces.....	29
Graphe 2. 3. Echelle de risque de la fréquence selon le type de bâtiment.....	29
Graphe 2. 4. Echelle de risque de la fréquence selon la superficie de la Dépendance	29
Graphe 2. 5. Histogramme de répartition des sinistres de la garantie « évènements climatiques » par tranche de 50 € (sinistres à montant nul et > 5700€ exclus).....	33
Graphe 2. 6. Comparaison des distributions théoriques et de la distribution empirique sur les sinistres appartenant à l'intervalle]0 ; 5700€[.....	37
Graphe 2. 7. Echelle de risque du coût moyen selon le nombre de pièces	38
Graphe 2. 8. Echelle de risque du coût moyen selon le type de bâtiment.....	38
Graphe 2. 9. Echelle de risque du coût moyen selon la superficie de la Dépendance	38
Graphe 2. 10. Echelle de risque de la prime pure selon le nombre de pièces.....	42
Graphe 2. 11. Echelle de risque de la prime pure selon le type de bâtiment.....	42
Graphe 2. 12. Echelle de risque de la prime pure selon la superficie de la Dépendance	42
Graphe 2. 13. Structure de calcul selon l'implémentation de la Commission Européenne.....	43
Graphe 2. 14. Mécanisme de modélisation des forces de la nature	45
Graphe 2. 15. Description du traité de réassurance interne des Caisses Régionales.....	48
Graphe 2. 16. Description du traité externe « Force de la Nature »	48
Graphe 2. 17. Interprétation des évènements climatiques 2009 et 2010 par rapport au nouveau calibrage de la prime pure	56
Graphe 2. 18. Interprétation du nouveau calibrage de la prime pure, par rapport à l'ancien paramétrage	57
Graphe 3. 1. Organigramme simplifié de Groupama SA	60
Graphe 3. 2. Passage entre la vision comptable française et la vision du bilan en norme IFRS (juste valeur)	63
Graphe 3. 3. Représentation graphique de l'exigence de besoin en capital sous Solvabilité 2	64
Graphe 3. 4. Représentation graphique de la sensibilité liée au risque de prime et de réserve	68
Graphe 3. 5. Représentation graphique de l'effet de diversification	69
Graphe 3. 6. Évolution du parc immobilier en France métropolitaine (Source SOeS – CGDD, Comptes du logement)	69
Graphe 3. 7. Représentation graphique des stratégies de rentabilité pour un produit Non-Vie.	73
Graphe 3. 8. Représentation graphique des effets de la réassurance dans les besoins en fonds propres	74

Graphe 4. 1. Répartition des différents types de contrats climatiques négociés de gré à gré au cours de la période 2000/2011.	82
Graphe 4. 2. Cartographie des IRIS selon leur taux de surface inondable.....	85
Graphe 4. 3. Cartographie des cours d'eau en France selon les données de MAP info	85
Graphe 4. 4. Cartographie des IRIS selon le nombre d'arrêtés catastrophes naturelles recensé jusqu'au 21/12/2011.	86
Graphe 4. 5. Répartition des arrêtés par nature de périls (cumul 1982-2010).....	86
Graphe 4. 6. Carte de cumul des demandes d'arrêtés de catastrophes naturelles « sécheresse » en France de 1982 à 2011	87
Graphe 4. 7. Prévisions sur le changement climatique en France sur les années 2021-2050	88
Graphe 4. 8. Carte de cumul d'arrêtés de catastrophe naturelle « mouvement de terrain » en France de 1982 à 2005.....	89
Graphe 4. 9. Localisation du séisme du 26 février 2012.....	90
Graphe 4. 10. Cartographie des départements selon leur aléa sismique.....	90
Graphe 4. 11. Réduction de la variance des résidus par l'ajout de critères supplémentaires.....	91
Graphe 4. 12. Portal d'accueil de Groupama Predict	94
 Graphe A. 1. Evolution de la concentration de CO ₂ en ppm (partie par million), du méthane (CH ₄) et de l'oxyde d'azote (N ₂ O) en ppb (partie par billion ~ milliard, équivalent français).....	100
Graphe A. 2. Emissions des gaz à effet de serre anthropiques dans le monde de 1970 à 2004 en équivalent CO ₂ pondérés selon la CCNUCC	101
Graphe A. 3. Emissions des gaz à effet de serre en France entre 1990 et 2009.	101
Graphe A. 4. Evolution de la température mondiale entre 1880 et 2005.....	101
Graphe A. 5. Evolution trimestrielle de l'indice FFB de 1980 à fin 2011	104
Graphe A. 6. Cartographie des Caisses Régionales à la fin de l'exercice 2010.....	105
Graphe A. 7. Sortie SAS dans le cadre de l'étude de la normalité des résidus de la fréquence	120
Graphe A. 8. Echelle de risque des coûts moyens selon le nombre de pièces sur les données brutes	127
Graphe A. 9. Echelle de risque des coûts moyens selon le type de bâtiment sur les données brutes	127
Graphe A. 10. Echelle de risque des coûts moyens selon la superficie de la Dépendance sur les données brutes	127
Graphe A. 11. Interprétation des événements climatiques 2009, 2010 et 2011 par rapport au nouveau calibrage de la prime pure	132
Graphe A. 12. Distribution d'une loi Tweedie	133
Graphe A. 13. Vision des indemnisations Santé pour l'Entité GGE.....	135