

**Mémoire présenté devant le CNAM pour l'obtention du Master
Droit Economie Gestion, mention Actuariat
et l'admission à l'Institut des Actuares**

le 13 mai 2024

Par : Habib CHEBCHOUB

Titre: Généralisation du modèle de Mack à l'environnement 3D de la
Responsabilité Civile Décennale

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

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

Président du Jury :
M. Stéphane LOISEL

signatures

Entreprise :

Nom : PwC

Membres présents du jury de
l'Institut des Actuares :

M. Arnaud LACOUME

Directeur de mémoire en entreprise :

Nom : Alexandre YOU

Signature :

M. Florence PICARD

Invité :

Nom : Hugo CHAUMONT

Signature :

Membres présents du jury du
Cnam :

M. Nathanaël ABECERA

M. Olivier DESMETTRE

M. David FAURE

M. François WEISS

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

Secrétariat :

Signature du responsable entreprise

Bibliothèque :

Signature du candidat

Table des matières

Table des matières	2
Table des figures	5
Table des tableaux.....	6
Résumé.....	7
Abstract	8
Note de Synthèse.....	9
Synthesis note.....	14
Remerciements.....	18
Introduction.....	19
Chapitre I : Environnement juridique et économique de l'assurance construction.....	21
1.1 Cadre de l'assurance construction	21
1.1.1 Opération de construction	21
1.1.2 Intervenants	22
1.1.3 Loi Spinetta	24
1.1.4 Garanties de l'assurance construction	27
1.2 Deux types de provisions.....	30
1.2.1 Provision pour sinistres non encore manifestés	30
1.2.2 PSNEM réglementaire.....	31
1.2.3 PSNEM économique	31
1.3 Marché français de l'assurance construction.....	38
1.3.1 Historique du marché	38
1.3.2 Poids économique.....	38
1.3.3 Matière assurable.....	40
1.3.4 Indices	41
1.4 Synthèse.....	43
Chapitre II : Méthode d'estimation des provisions en assurance construction	45
2.1 Analyse préliminaire des données	45
2.1.1 Présentation et qualité des données.....	45

2.1.2	Vision Survenance × Développement (vision PSAP).....	46
2.1.3	Vision DOC × Survenance (vision PSNEM).....	48
2.2	Méthodes classiques de provisionnement Non-Vie.....	51
2.2.1	Méthodes déterministes.....	51
2.2.2	Méthodes stochastiques.....	62
2.3	Vers un Modèle de Mack 3D pour l'Assurance Construction.....	73
2.3.1	Environnement 3D de l'assurance construction.....	73
2.3.2	Modèle GLM « 3D » : modèle multiplicatif Poisson ODP	74
2.3.3	Modèle de Mack 3D.....	78
2.4	Synthèse.....	85
Chapitre III : Mise en pratique du modèle de Mack 3D.....		86
3.1	Calcul de la PSNEM.....	86
3.2	Résultat 2D	87
3.3	Résultat 3D	88
3.4	Résultat du modèle GLM.....	90
3.5	Comparaison des résultats	91
3.6	Synthèse.....	92
Conclusion.....		93
Bibliographie.....		95
1.	Ouvrage	95
2.	Articles.....	95
3.	Mémoires et Rapports.....	95
Annexes.....		97
Annexe 1 : Développement du calcul de la MSEP.....		97
Annexe 2 : Triangles de liquidation		99
Annexe 3 : Graphiques test de linéarité vision Survenance × Développement		101
Annexe 4 : Graphiques test d'indépendance vision Survenance × Développement.....		103
Annexe 5 : Graphiques test de variance vision Survenance × Développement.....		105

Annexe 6 : Graphiques test de linéarité vision DOC × Survenance	107
Annexe 7 : Graphiques test d'indépendance vision DOC × Survenance.....	109
Annexe 8 : Graphiques test de variance vision DOC × Survenance.....	111

Table des figures

Figure 1.1-1 Etapes clés de l'opération de construction	21
Figure 1.1-2 Illustration des différentes garanties de l'assurance construction	27
Figure 1.2-1 Provisions pour sinistres en assurance construction.....	30
Figure 1.2-2 Synthèse de la méthode « Nombre × Coût moyen »	34
Figure 1.2-3 Mécanisme de ventilation des IBNR par DOC	36
Figure 1.3-1 L'activité des assurances de biens et de responsabilité en 2020 (source : FFA).....	39
Figure 1.3-2 Cotisations selon le périmètre du marché (Sources : ACPR, FFA)	39
Figure 1.3-3 Evolution du nombre de logements mis en chantier	40
Figure 1.3-4 Evolution du ratio combiné construction (net de réassurance)	41
Figure 2.1-1 Cohorte des règlements (en k€) en vision Survenance x Développement	47
Figure 2.1-2 Cohorte des charges (en k€) en vision Survenance x Développement.....	47
Figure 2.1-3 Répartition des charges par année de survenance	48
Figure 2.1-4 Cohorte des règlements (en k€) en vision DOC x Délai de Survenance.....	49
Figure 2.1-5 Cohorte des charges (en k€) en vision DOC x Délai de Survenance	49
Figure 2.2-1 Triangle de liquidation Survenance x Développement	52
Figure 2.2-2 Méthode de liquidation Chain Ladder.....	55
Figure 2.2-3 Illustration de l'erreur de prédiction.....	63
Figure 2.2-4 Illustration de l'erreur d'estimation	64
Figure 2.2-5 Quantification de l'erreur de processus.....	64
Figure 2.3-1 Visualisation des données en pyramide.....	73
Figure 2.3-2 Visualisation des données en pyramide.....	74

Table des tableaux

Tableau 1.1-1 Résultats GLM 3D	12
Tableau 1.1-2 Indicateurs des modèles GLM 3D	12
Tableau 1.1-3 synthèses des résultats.....	12
Tableau 1.1-1 Résultats GLM 3D	16
Tableau 1.1-2 Indicateurs des modèles GLM 3D	17
Tableau 1.1-3 synthèses des résultats.....	17
Tableau 1.1-1 Récapitulatif des garanties en assurances construction	29
Tableau 2.2-1 Test de linéarité - Triangle Survenance x Développement.....	53
Tableau 2.3-1 Test de linéarité - Triangle DOC x Survenance.....	80
Tableau 3.1-1 Résultats des provisions pas année de DOC	87
Tableau 3.2-1 Résultats du modèle Mack 2D	88
Tableau 3.3-1 Résultats du modèle Mack 3D	89
Tableau 3.4-1 Résultats GLM 3D	90
Tableau 3.4-2 Indicateurs des modèles GLM 3D	91
Tableau 3.5-1 synthèses des résultats.....	91

Résumé

Dans un contexte de reprise économique du secteur du BTP et de l'évolution des matériaux et des techniques de construction (nouvelles normes environnementales et énergétiques, etc.), l'assurance construction en France ne semble toujours pas se diriger vers une sortie de la crise.

En assurance construction, les équilibres techniques demeurent fragiles et la mesure de la rentabilité par les assureurs est un exercice complexe. Cela s'explique notamment par la dimension temporelle des garanties décennales obligatoires (Responsabilité Civile Décennale, Dommages Ouvrage), caractéristiques de l'assurance construction, ainsi que par la projection sur des périodes longues de paramètres économiques tels que l'inflation et les produits financiers.

Dans l'optique d'établir une mesure fiable de la rentabilité, un enjeu clef pour les assureurs, ces derniers s'échinent à disposer de méthodes fiables d'évaluation des provisions techniques ainsi que de l'incertitude associée. Des méthodes actuarielles stochastiques sont donc nécessaires. Cependant, les approches existantes s'avèrent soit inadaptées (Mack, Merz Wüthrich, Bootstrap...), soit très coûteuses en temps de calcul (Nested Bootstrap).

Le présent mémoire s'articule autour de trois chapitres. La première partie présente l'environnement juridique et économique de l'assurance construction. Ensuite, un intérêt sera porté aux différentes méthodes d'estimation des provisions en Non-Vie en générale et en assurance construction en particulier et serviront de base pour le développement d'une généralisation du modèle stochastique de Mack à l'assurance construction. Enfin la dernière partie sera consacré à la mise en pratique de la méthode développée de Mack 3D sur un portefeuille réel d'assurance construction ainsi qu'à l'évaluation de sa performance, notamment en comparaison à d'autres approches.

Mots-clés : Assurance construction, DO, Dommages Ouvrages, RCD, Responsabilité Civile Décennale, Provisionnement, Chain Ladder, PSAP, PSNEM, Pyramide de liquidation, Provisionnement 3D.

Abstract

In a context of economic recovery of the construction sector and the evolution of materials and construction techniques (new environmental and energy standards, etc.), construction insurance in France still does not seem to be heading towards an exit from the crisis.

In construction insurance, technical balances remain fragile and the measurement of profitability by insurers is often complex. This can be explained by the two steps process of the ten-year mandatory guarantees (RCD, DO), as well as by the projection of economic parameters such as inflation and financial products over extended periods.

To establish a reliable measure of profitability, one of the key issues for insurers is to have reliable methods for assessing the technical reserves and the related uncertainty. Stochastic actuarial methods are therefore necessary. However, the existing approaches prove to be inadequate (Mack, Merz Wüthrich, Bootstrap ...), or very computing time-consuming (Nested Bootstrap).

This dissertation is structured around three chapters. The first part presents the legal and economic environment of construction insurance. Next, the various methods for estimating reserves in P&C in general, and in construction, will be discussed. These will serve as a basis for developing a generalization of Mack's stochastic model to construction insurance. Finally, the last part will be devoted to putting Mack 3D's developed method into practice on a real construction insurance portfolio, and assessing its performance, particularly in comparison with other approaches.

Keywords: Construction insurance, Property, Casualty, Public liability, Reserving, Chain Ladder, PSNEM, PSAP, Pyramid data, DO, RCD, 3D reserving.

Note de Synthèse

Domaine d'expert, l'assurance construction présente plusieurs spécificités tant à l'échelle juridique qu'à l'échelle technique. De ce fait, la mesure de la rentabilité des portefeuilles et l'évaluation de la précision de l'estimation du passif représentent des défis majeures pour les assureurs du BTP. Ces difficultés constituent un frein à la fiabilité de la mesure et du pilotage du niveau de prudence des calculs des provisions techniques.

Présentant une structure particulière et des mécanismes spécifiques, la configuration du secteur de l'assurance construction date de la promulgation de la loi Spinetta de 1978. Motivée par une sous-assurance et une insuffisance de la couverture du maître d'ouvrage, cette dernière a introduit le principe de l'obligation des assurances construction ainsi que le mécanisme à double détente. En effet, deux obligations de souscription d'assurance ont été définies : une assurance dommages ouvrage (DO) pour le maître d'ouvrage et une assurance responsabilité civile décennale (RCD) pour les constructeurs du chantier. Ces derniers étant définis, à la suite de la survenance d'un sinistre à l'édifice assuré, l'assureur dommages ouvrage intervient afin de dédommager immédiatement l'intégralité des préjudices subis. Après établissement des responsabilités, l'assureur DO entame des recours sur les garanties de responsabilité civile des constructeurs.

La gestion en capitalisation des garanties DO et RCD et le caractère décennal de la couverture obligent un suivi spécifique des sinistres. Les sinistres étant rattachés à la DOC (exercice de souscription), il faut considérer une nouvelle dimension et donc une représentation en 3D de type « pyramide » : Survenance × Développement × DOC. Pour une DOC donnée, on dispose, comme pour les branches « classiques », de données de survenance par développement.

En pratique deux types de provisions pour sinistres doivent être constituées.

- La provision pour sinistres à payer (PSAP) : valeur estimative des dépenses en principal et en frais, tant internes qu'externes, nécessaires au règlement de tous les sinistres survenus et non payés, y compris les capitaux constitutifs des rentes non encore mises à la charge de l'entreprise (Art. R343-7 du Code des Assurances).
- La provision pour sinistres non encore manifestés (PSNEM) : elle couvre les sinistres liés à des DOC antérieures à la date d'inventaire, mais non survenus.

Bien que la distinction soit faite en pratique pour des raisons de simplification de la lecture, d'un point de vue juridique et comptable, la PSNEM fait également partie des PSAP.

Les PSAP sont calculées par exercice de survenance et couvrent :

- Les sinistres connus évalués dossier par dossier (jugement d'expert, montant garanti, etc.), le coût d'un dossier comprenant toutes les charges externes individualisables (frais d'expertises notamment) ;
- Les sinistres survenus et non déclarés (IBNR) évalués au global selon des méthodes statistiques (pas de préconisation de méthodes).

La PSNEM peut être calculée par la méthode réglementaire, à l'arrêté de décembre, selon les modalités définies par l'article 143-14 du Règlement N°2015-11 de l'ANC. Le calcul s'effectue par DOC, séparément pour la DO et la RCD.

Le calcul de la PSNEM selon le modèle réglementaire de l'article 143-14 du Règlement N°2015-11 de l'ANC est une contrainte minimale. Une société peut par exemple construire son propre barème à partir des données de son portefeuille, en estimant ses propres cadences de manifestation. Si ce barème est plus prudent que le réglementaire, elle doit retenir le montant économique.

La méthode réglementaire ne permet pas de prendre en compte les caractéristiques propres au portefeuille de l'assureur concerné. Il est possible d'estimer la charge ultime des sinistres non manifestés, sur la base des statistiques de l'assureur : estimation d'une PSNEM économique.

La méthode « classique » de calcul de la PSNEM économique est une méthode qui s'articule en trois étapes :

- 1ère étape : Calcul des IBNR par projection du triangle Survenance × Développement ;
- 2ème étape : Ventilation des IBNR par DOC ;
- 3ème étape : Estimation de la PSNEM par projection du triangle DOC × Survenance ventilé (contenant les IBNR).

Les méthodes traditionnelles de provisionnement stochastique, qui ont été utilisées pendant de nombreuses années pour évaluer les provisions pour sinistres, ne sont pas assez flexibles pour s'adapter aux complexités de l'assurance construction.

Notamment, une limite du modèle de Mack classique, qu'on appellera Mack 2D, en assurance construction réside dans son traitement des IBNR. Dans le modèle de Mack 2D, le triangle considéré est supposé être déterministe tant dis que ce dernier contient les IBNR, résultat d'une estimation.

Le calcul de la variance doit ainsi être conditionné aux estimations IBNR pour tenir compte du biais introduit par leur ventilation. Cette hypothèse déterministe peut donc conduire à une sous-estimation de la variance des provisions.

Le modèle de Mack classique, qui se concentre sur les provisions pour sinistres en deux dimensions, est adapté dans ce travail pour tenir compte des spécificités de l'assurance construction. Cette adaptation à trois dimensions, qu'on appellera Mack 3D, permet de mieux modéliser les risques associés.

Pour cela, le modèle de Mack 3D proposé dans ce mémoire est une généralisation d'un modèle de Mack, conditionnel à l'estimation des IBNR. Le résultat de cette généralisation est présenté ci-après.

$$\widehat{MSEP}(PSNEM_k) = \hat{E}_{k,n}^2 \sum_{l=n-k+1}^{n-1} \frac{\hat{\phi}_l^2}{\hat{\lambda}_l^2} \left(\frac{1}{\hat{E}_{k,l}} + \frac{1}{\sum_{s=1}^{n-l} C_{s,l}} \right) + \text{erreur « 3D »}^2 .$$

Dans le cadre de notre étude, nous introduisons également le Modèle « GLM : modèle multiplicatif Poisson ODP » utilisé comme point de comparaison pour évaluer les performances du modèle de Mack 3D. Cette comparaison nous permettra de mieux comprendre les avantages et les limites de chaque approche. Le résultat du modèle est présenté ci-après.

$$\overline{MSEP}(\hat{R}) = \hat{\phi} \sum_{i+(j-1)+(k-1) > N} \hat{C}_{ijk} + g^{-1}(M \cdot \hat{\beta}) \cdot M \cdot I^{-1} \cdot M' \cdot \left(g^{-1}(M \cdot \hat{\beta}) \right)'$$

Avant d'entamer la mise en œuvre pratique, une analyse des données est réalisée. La qualité des données étant un facteur non négligeable dans le processus de provisionnement, une étude de cette dernière sert, en outre, à expliquer les résultats trouvés.

Les données utilisées pour cette étude correspondent à des données réelles d'un portefeuille construction d'une mutuelle sur le marché français. Pour des considérations de confidentialité, les données ont été anonymisées.

Plusieurs contrôles ont été réalisés afin de fiabiliser les données. Des statistiques descriptives ont permis de mettre en évidence la cohérence de ces dernières et confirment que le portefeuille étudié présente des caractéristiques observées sur le marché de l'assurance construction en France.

Nous présentons ci-dessous les résultats obtenus grâce à l'application du modèle de Mack 3D, qui a été développé dans ce mémoire. Ce modèle, adapté aux spécificités de l'assurance construction, a permis d'obtenir des estimations plus précises de la volatilité de l'estimation des provisions des garanties construction. Une comparaison des résultats du modèle de Mack 3D avec le modèle de Mack 2D vise à démontrer l'impact de l'intégration de l'erreur de prédiction lié à la ventilation des IBNR dans l'estimation global de l'erreur.

Modèle	Provisions	RMSE	CV
Mack 2D	902 557 144	49 212 072	5%
Mack 3D	902 557 144	111 765 723	12%

En appliquant un modèle de Mack classique sur les données, nous obtenons une estimation de l'erreur quadratique (RMSE) de 50 m€, soit 5% des provisions. Cette estimation est le résultat du modèle de Mack sur le triangle de PSNEM sans tenir compte de la volatilité induite par l'estimation des IBNR ventilé dans le triangle.

Par ailleurs, bien que les années anciennes présentent des montant d'IBNR non nuls, l'estimation de l'erreur de Mack 2D sur celles-ci est nulle. Ce phénomène est corrigé par le modèle Mack 3D.

Comme expliqué précédemment, cette approche est critiquable du fait que l'application directe du modèle de Mack suppose que le triangle étudié est déterministe, ce qui n'est pas la cas. Ainsi, le modèle sous-estime fortement la volatilité des provisions en ne tenant pas compte de la volatilité induit par l'estimation des IBNR et leur ventilation. Ce qui est confirmé par les résultats.

L'application du modèle de Mack 3D introduit un terme supplémentaire dans le calcul de la RMSE. Ce terme qu'on appellera « erreur 3D », permet de capter la volatilité de l'estimation des IBNR, leur ventilation et leur projection. Ainsi, nous obtenons, une RMSE de 112 m€, soit 12% des provisions.

On note également que l'erreur 3D est plus importante sur les années récentes. Ceci est expliqué du fait que la projection est réalisée sur une assiette de charge y compris IBNR. Ainsi, une erreur d'estimation de cette dernière est répercutée à horizon de projection (caractère multiplicatif de la méthode Chain Ladder).

Afin de comparer les résultats obtenus par le modèle Mack 3D, nous mettons en place un modèle GLM 3D. Usuellement en provisionnement Non-Vie, le choix de la loi a priori se porte sur une loi de Poisson surdispersée (ODP). Nous avons cependant sélectionné trois lois candidates pour comparaison : Poisson ODP, Tweedie et Binomiale Négative.

Les résultats des différents modèles sont donnés dans le tableau ci-dessous.

Loi	Provisions	RMSE	CV
Poisson ODP	1 042 002 712	94 005 919	9%
Binomiale Négative	745 462 783	361 381 624	48%
Tweedie	899 170 234	147 806 550	16%

Tableau 1.1-1 Résultats GLM 3D

Dans le but de juger de la qualité des estimations des différents modèles GLM 3D, nous calculons différents indicateurs, présentés ci-dessous.

Loi	RMSE du modèle (€)	R ²	TopDecile Lift	Log Loss	Quotient
Poisson ODP	415 856	89%	6,293	4,026	1,000
Binomiale Négative	812 692	57%	6,318	4,026	1,245
Tweedie	482 993	85%	6,338	4,026	1,011

Tableau 1.1-2 Indicateurs des modèles GLM 3D

Les modèles avec les lois Poisson ODP et Tweedie présente à la fois une RMSE acceptable et un R² proche de 1. Leurs TopDecileLift, Log Loss et Quotient indiquent à la fois une bonne performance des modèles mais également une comparabilité des deux résultats. Se basant principalement sur la plus petite RMSE et le R² le plus proche de 1, le cas du Poisson ODP semble être le modèle le plus performant.

On note par ailleurs, que la performance du modèle Binomiale Négative est faible. L'erreur quadratique de ce dernier est nettement supérieure aux erreurs des autres modèles et son R² est très faible et inférieur à 60%. Ces constatation nous pousse à ne pas retenir ce dernier et l'exclure des comparaisons qui suivent.

Enfin, nous comparerons les résultats du Mack 3D obtenus avec les résultats d'un GLM 3D via hypothèses de distribution retenues.

Loi/Modèle	Provisions	RMSE	CV
Mack 3D	902 557 144	111 765 723	12%
Poisson ODP	1 042 002 712	94 005 919	9%
Tweedie	899 170 234	147 806 550	16%

Tableau 1.1-3 synthèses des résultats

On note que l'estimation des provisions est comparable entre les trois modèles retenus avec une légère surestimation dans le cas du modèle GLM 3D Poisson ODP.

On constate que les erreurs quadratiques calculées à la fois par le modèle Mack 3D et les modèles GLM 3D sélectionnés sont comparables et nettement supérieurs à l'erreur quadratique calculée par le modèle Mack 2D (50 m€). Ce constat permet de conforter le fait que l'application d'un modèle de Mack classique en assurance construction est erronée et de confirmer que le modèle de Mack 3D corrige cette erreur.

On note également qu'en pourcentage des provisions, l'erreur quadrique calculée par le modèle de Mack 3D est encadrée entre les deux erreurs calculées par les modèles GLM 3D sélectionnés. Ceci permet de situer un ordre de grandeur de la volatilité qu'on souhaite capter lors de l'estimation des provisions d'une part et d'autre part de confirmer la bonne estimation de l'erreur quadratique par le modèle de Mack 3D.

Synthesis note

As an expert field, construction insurance presents several specificities at both the legal and technical levels. As a result, measuring the profitability of portfolios and assessing the accuracy of liability estimates are major challenges for construction insurers. These difficulties are a hindrance to the reliability of the measurement and management of the level of prudence of the technical reserve calculations.

With its structure and specific mechanisms, the configuration of the construction insurance sector dates to the enactment of the 1978 Spinetta law. Motivated by an underinsurance and an insufficient coverage of the owner, this law introduced the principle of the obligation of construction insurance as well as the double trigger mechanism. Indeed, two obligations of insurance subscription were defined: a damage insurance (DO) for the owner and a ten-year third-party liability insurance (RCD) for the builders of the site. The latter being defined, following the occurrence of a disaster to the insured building, the damage to work insurer intervenes to immediately compensate the entirety of the prejudices suffered. Once the responsibilities have been established, the DO insurer initiates recourse on the civil liability guarantees of the builders.

The capitalization management of the DO and RCD guarantees, and the ten-year nature of the coverage require a specific handling of the claims. Since claims are linked to the DOC (underwriting year), a new dimension must be considered and therefore a 3D representation of the "pyramid" type: Occurrence x Development x DOC. For a given DOC, we have, as for the "classic" lines of business, occurrence data by development.

In practice, two types of claims reserves must be established:

- The reserve for claims payable (PSAP): the estimated value of the expenses in principal and costs, both internal and external, required to settle all claims that have occurred and not yet been paid, including the capital constituting annuities not yet payable by the company (Art. R343-7 of the French Insurance Code).
- The reserve for claims not yet incurred (PSNEM): this covers claims related to DOCs prior to the inventory date but not yet incurred.

The PSAP is calculated by year of occurrence and covers:

- Known claims assessed on a case-by-case basis (expert's judgment, guaranteed amount, etc.), the cost of a case including all expenses that can be individualized (notably expert fees).
- Claims incurred but not reported (IBNR) evaluated globally according to statistical methods (no recommended methods).

The PSNEM can be calculated by the regulatory method, at the end of December, in the manner defined by Section 143-14 of CNA Regulation No. 2015-11. The calculation is done by DOC, separately for DO and RCD.

The calculation of the PSNEM according to the regulatory model in section 143-14 of CNA Regulation No. 2015-11 is a minimum constraint. For example, a company can build its own rates from its portfolio data. If this estimate is more conservative than the regulatory one, it must retain the economic amount.

The regulatory method does not consider the specific characteristics of the insurer's portfolio. It is possible to estimate the ultimate cost of unmanifested claims, based on the insurer's own statistics: estimate an economic PSNEM.

The "classic" method of calculating the economic PSNEM is a three-step method:

- Step 1: Calculation of IBNR by projection of the Occurrence-Development triangle.
- Step 2: Breakdown of IBNR by DOC.
- Step 3: Estimation of the PSNEM by projection of the triangle DOC x Occurrence (containing the IBNR).

Traditional reserving methods, which have been used for many years to evaluate claims reserves, are not flexible enough to adapt to the complexities of construction insurance.

The classic Mack model, which focuses on claims reserves in two dimensions, is adapted in this work to take account of the specificities of construction insurance. This three-dimensional adaptation enables us to better model the associated risks.

One of the limitations of the Mack 2D model in construction insurance lies in its treatment of IBNR. In the Mack 2D model, the basic triangle is deterministic if it contains IBNR, the result of an estimation.

The calculation of variance, conducted using conventional stochastic methods, must therefore be conditional on the IBNR estimates to take account of the bias introduced by their breakdown. This deterministic assumption may therefore lead to an underestimation of the variance of provisions.

To this end, the Mack 3D model proposed in this thesis is a generalization of a Mack model, conditional on IBNR estimates. The result of this generalization is presented below.

$$\widehat{MSEP}(PSNEM_k) = \hat{E}_{k,n}^2 \sum_{l=n-k+1}^{n-1} \frac{\hat{\varphi}_l^2}{\hat{\lambda}_l^2} \left(\frac{1}{\hat{E}_{k,l}} + \frac{1}{\sum_{s=1}^{n-l} C_{s,l}} \right) + \text{erreur « 3D »}^2.$$

In the context of our study, we also introduce the "GLM model: Over-dispersed Poisson ODP model" used as a point of comparison to evaluate the performance of Mack's 3D model. This comparison will enable us to better understand the advantages and limitations of each approach. The result of the model is presented below.

$$\widehat{MSEP}(\hat{R}) = \hat{\phi} \sum_{i+(j-1)+(k-1) > N} \hat{C}_{ijk} + g^{-1}(M \cdot \hat{\beta}) \cdot M \cdot I^{-1} \cdot M' \cdot \left(g^{-1}(M \cdot \hat{\beta}) \right)'$$

Before starting to calculate provisions, a data analysis is conducted. As data quality is a key factor in the reserving process, a study of data quality also explains the results found.

The data used for this study corresponds to actual data from a mutual's construction portfolio on the French market. For reasons of confidentiality, the data has been anonymized.

Several checks were conducted to ensure the reliability of the data. Descriptive statistics were used to highlight the coherence of the data, confirming that the portfolio studied presents characteristics observed in the French construction insurance market.

Below, we present the results obtained by applying the Mack 3D model developed in this thesis. This model, adapted to the specificities of construction insurance, has enabled us to obtain more accurate estimates of the volatility of construction warranty reserve estimates. A comparison of the results of the Mack 3D model with the classic Mack model (2D) aims to demonstrate the impact of integrating the prediction error linked to the breakdown of IBNR into the overall error estimate.

Modèle	RMSE	CV
Mack 3D	49 212 072	5%
Mack 2D	111 765 723	12%

Applying a classic Mack model to the data, we obtain an estimate of the squared error (RMSE) of €50m, or 5% of the reserves. This estimate is the result of the Mack model on the PSNEM triangle, without considering the volatility induced by the IBNR estimate broken down in the triangle.

In addition, although older years have non-zero IBNR amounts, the estimate of the Mack 2D error on them is zero. The Mack 3D model corrects this phenomenon.

As explained above, this approach is open to criticism, as the direct application of the Mack model assumes that the triangle studied is deterministic, which is not the case. As a result, the model severely underestimates the volatility of provisions by failing to consider the volatility induced by the estimation of IBNR and their breakdown. This is confirmed by the results.

The application of Mack's 3D model introduces an additional term into the RMSE calculation. This term, called "3D error", captures the volatility of IBNR estimation, breakdown, and projection. This gives us an RMSE of €112m, or 12% of provisions.

We also note that the 3D error is greater in recent years. This is explained by the fact that the projection is based on a load base including IBNR. Thus, an error in estimating the latter is passed on to the projection horizon (multiplicative nature of the Chain Ladder method).

To compare the results obtained by the Mack 3D model developed in this thesis, we set up a GLM 3D model. In non-life provisioning, the a priori choice of law is usually an Over-dispersed Poisson distribution (ODP). However, we have selected three candidate distributions for comparison: Poisson ODP, Tweedie, and Negative Binomial.

The results of the different models are shown in the table below.

Loi	Provisions	RMSE	CV
Poisson ODP	1 042 002 712	94 005 919	9%
Binomiale Négative	745 462 783	361 381 624	48%
Tweedie	899 170 234	147 806 550	16%

Tableau 1.1-1 Résultats GLM 3D

To assess the quality of the estimates of the various GLM 3D models, we calculate a number of indicators, presented below. To assess the quality of the estimates of the various GLM 3D models, we calculate a number of indicators, presented below.

Loi	RMSE du modèle (€)	R²	TopDecile Lift	Log Loss	Quotient
Poisson ODP	415 856	89%	6,293	4,026	1,000
Binomiale Négative	812 692	57%	6,318	4,026	1,245
Tweedie	482 993	85%	6,338	4,026	1,011

Tableau 1.1-2 Indicateurs des modèles GLM 3D

The Poisson ODP and Tweedie models show both acceptable RMSE and R² close to 1. Their TopDecileLift, Log Loss and Quotient indicate both good model performance and comparability of the two results. Based on the smallest RMSE and R² closest to one, the Poisson ODP case is the best-performing model.

On the other hand, the performance of the Negative Binomial model is weak. Its squared error is clearly greater than the errors of the other models, and its R² is extremely low, below 60%. These findings lead us to exclude this model from the following comparisons.

Finally, we will compare the results of the 3D Mack with those of a 3D GLM based on the distribution assumptions adopted.

Loi/Modèle	Provisions	RMSE	CV
Mack 3D	902 557 144	111 765 723	12%
Poisson ODP	1 042 002 712	94 005 919	9%
Tweedie	899 170 234	147 806 550	16%

Tableau 1.1-3 synthèses des résultats

We note that the estimate of provisions is comparable between the three selected models, with a slight overestimation in the case of the GLM 3D Poisson ODP model.

We note that the squared errors calculated by both the Mack 3D model and the selected GLM 3D models are comparable and significantly higher than the squared error calculated by the Mack 2D model (50 m€). This finding supports the view that the application of a conventional Mack model in construction insurance is erroneous and confirms that the Mack 3D model corrects this error.

We also note that, as a percentage of provisions, the quadratic error calculated by the Mack 3D model is framed between the two errors calculated by the selected GLM 3D models. This gives us an order of magnitude of the volatility we wish to capture when estimating reserves and confirms that the Mack 3D model's estimate of the squared error is correct.

Remerciements

Je tiens plus particulièrement à remercier Alexandre YOU pour son encadrement, sa disponibilité, ses explications, sa patience et les pistes de réflexion proposées tout au long de la rédaction de ce mémoire. Je souhaiterais également remercier Hugo CHAUMONT pour ses conseils avisés.

Je tiens à exprimer ma gratitude à mon encadrant académique, Stéphane Loisel, pour ses remarques pertinentes, ses conseils éclairés et sa bienveillance qui ont contribué à l'aboutissement de ce mémoire.

Introduction

Confronté à de nouvelles exigences tant réglementaires qu'environnementales, le secteur du BTP en France tente avec difficulté de sortir d'une crise qui dure maintenant depuis 10 ans. Les indicateurs économiques au vert, les acteurs de l'assurance misent désormais sur le secteur du bâtiment. Ils ont, cependant, plus que jamais besoin de mesurer leur rentabilité avec fiabilité et de pouvoir piloter leurs marges techniques.

Le poids important des provisions pour sinistres dans le passif d'un assureur Non-Vie rend le recours au provisionnement stochastique de plus en plus nécessaire. Cet enjeu est accru en assurance construction du fait du fonctionnement en capitalisation de la branche qui oblige les assureurs à provisionner les sinistres futurs au titre des exercices de souscription pour lesquels ils sont engagés, via la constitution de la provision pour sinistres non encore manifestés (PSNEM).

Sur le plan prudentiel, le recours au provisionnement stochastique apparaît également crucial pour le calcul du risque de réserve sous Solvabilité 2, notamment pour les acteurs ayant développé un modèle interne. Par ailleurs, la publication de la nouvelle norme IFRS 17 dont l'entrée en vigueur est prévue pour 2022 prévoit le calcul et la comptabilisation d'un ajustement au titre du risque non financier, ou Risk Adjustment (RA). Cet ajustement est destiné à compenser le coût pour l'assureur de l'incertitude associée à la liquidation de ses engagements. A la différence de la formule standard de Solvabilité 2 qui impose un mode de calcul très encadré de la Risk Margin, IFRS 17 ne prescrit pas de méthode particulière pour la détermination du Risk Adjustment, ce qui laisse la porte ouverte pour les approches stochastiques de provisionnement.

Bien que de nombreuses techniques de provisionnement stochastique en assurance Non-Vie existent, celles-ci ont été développées dans le cadre classique des triangles de sinistres bidimensionnels. Elles s'avèrent non adaptées aux spécificités du secteur de l'assurance construction, car souvent peu rigoureuses sur le plan mathématique ou consommatrices en temps de calcul (e.g. approches simulatoires basées sur le Bootstrap). Ceci motive le développement d'une formule fermée dans le cadre spécifique de l'assurance construction.

Ce mémoire a donc pour objet de proposer une méthodologie permettant de quantifier l'erreur de prédiction relatives à l'estimation des provisions techniques en assurance construction, en proposant notamment une approche par formule fermée comme extension de la formule de Mack adaptée à la nature tridimensionnelle des triangles de sinistres.

La première partie du mémoire fait, en deux chapitres, un état des lieux du secteur du BTP en général et de l'assurance construction en particulier. Le premier chapitre introduit les notions générales de l'assurance construction allant de la définition de l'opération de construction aux garanties résultant des lois en vigueur en passant par la présomption de responsabilité des intervenants sur un chantier. Le deuxième chapitre décrit le marché français de l'assurance construction, son historique et ses évolutions.

Une deuxième partie introduit à la fois les provisions réglementaires en assurance construction, à savoir les PSAP (Provisions pour Sinistres A Payer) et la PSNEM (Provision pour Sinistres Non Encore Manifestés), et les méthodes de provisionnement Non-Vie tant déterministes que stochastiques. Cette partie s'achève par une modélisation des deux méthodes de provisionnement stochastiques 3D sur lesquelles porte l'étude : les modèles GLM 3D et le modèle Mack 3D.

Une dernière partie est consacrée à l'illustration numérique des méthodes proposées sur un portefeuille d'assurance construction. Une description du portefeuille et des données utilisées sera proposée avant la présentation de la mise en œuvre des méthodes stochastiques et l'analyse des résultats.

Chapitre I : Environnement juridique et économique de l'assurance construction

1.1 Cadre de l'assurance construction

Ce chapitre a pour objectif d'introduire les notions générales de l'assurance construction. Ces dernières sont à même de présenter les particularités de ce secteur et les enjeux qui en découlent. Il sera suivi par une illustration des garanties relatives au secteur du BTP et d'une partie présentant l'évolution et les acteurs du marché français. Pour conclure, les détails techniques spécifiques au domaine de la construction seront développés.

1.1.1 Opération de construction

Une opération de construction est une opération visant à édifier un ouvrage dans un site selon le choix du maître d'ouvrage et conformément à la vision du concepteur. Le déroulement de l'opération de construction s'articule autour de quatre étapes clés, chronologiquement successives :

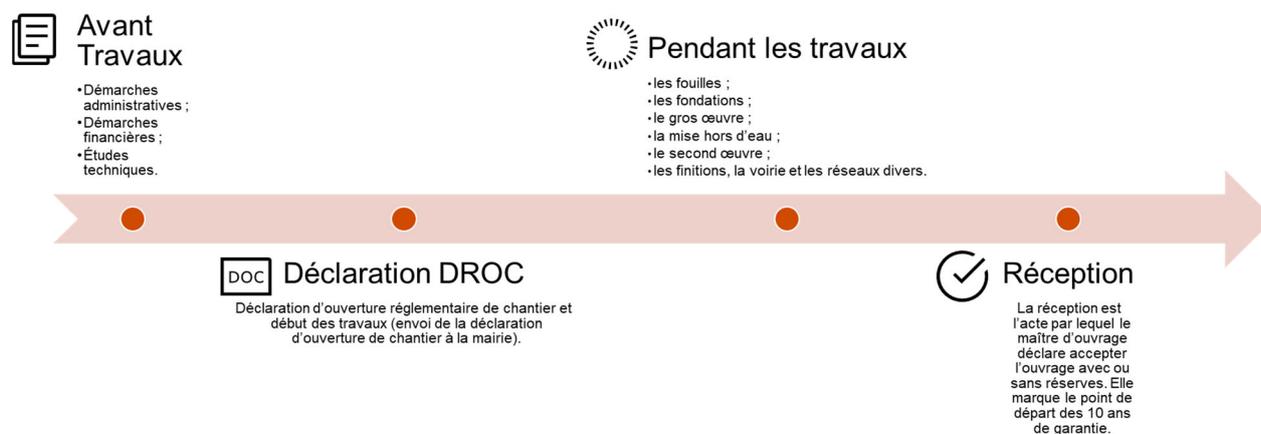


Figure 1.1-1 Etapes clés de l'opération de construction

Chacune des étapes illustrées ci-dessus voit intervenir différents acteurs. Ces derniers seront présentés dans la sous-section suivante.

1.1.2 Intervenants

Le maître d'ouvrage

Le maître d'ouvrage est une personne physique ou morale, publique ou privée, titulaire du droit de construire, et pour le compte de qui l'ouvrage est réalisé. Commanditaire du projet, il définit le cahier des charges ainsi que les besoins budgétaires et les exigences de qualité. Il est responsable du bon déroulement des travaux et procède à la réception de l'ouvrage, une fois achevé.

Constructeurs et assimilés

Afin de réaliser les travaux, le maître d'ouvrage (donneur d'ordres), se lie, au moyen d'un contrat de louage à un ensemble d'intervenants (locateurs d'ouvrage). Par ce contrat, le locateur d'ouvrage s'engage à réaliser un ouvrage déterminé, en totalité ou partie.

Sous la dénomination de constructeurs, l'Article 1792-1 du Code Civil désigne non seulement les entrepreneurs et la maîtrise d'œuvre mais également des intervenants assimilés tels que les vendeurs d'immeuble à construire ou après achèvement et les mandataires du maître d'ouvrage.

a) Un contrat de louage

Un contrat de louage est une convention par laquelle un maître d'ouvrage fait appel à une ou plusieurs personnes ou entreprises, en vue de leur faire réaliser un certain ouvrage (Article 1710 et 1779 du Code Civil). Dans le domaine de la Construction, ce contrat peut être passé entre le maître d'ouvrage et :

- Un entrepreneur, pour la construction ou la modification d'un ouvrage de bâtiment ou de génie civil.
- Un maître d'œuvre (architectes, ingénieur-conseil, bureaux d'études techniques), pour la conception de l'ouvrage, la direction et le contrôle de l'exécution des travaux.

b) Entrepreneurs

Les entrepreneurs sont des personnes morales ou physiques qui, dans le cadre de la maîtrise d'œuvre, peuvent intervenir, sous ordre du maître d'ouvrage, sur une ou plusieurs tâches, notamment la conception, l'édification ou la modification de l'ouvrage. Bien qu'engageant leur responsabilité sur une obligation de résultats, ils peuvent sous-traiter une partie ou la totalité des tâches qui leurs sont confiées.

c) Maître d'œuvre

Pour le compte du maître d'ouvrage, le maître d'œuvre agit en tant que fonction support pour diriger, contrôler et veiller à la bonne exécution des travaux et peut, dans certains cas, assister le

maître d'ouvrage dans la réception. Il est, en outre, responsable des aspects architecturaux, techniques et économiques du chantier. Cela étant, les maîtres d'œuvre peuvent être des architectes, des ingénieurs-conseils, ou encore des bureaux d'études techniques.

d) Les constructeurs de maisons individuelles (CMI)

Le constructeur de maisons individuelles est un entrepreneur soumis à une réglementation spécifique (Loi n° 90-1129 du 19 décembre 1990). C'est un maître d'œuvre dont le champ d'action se limite à la construction de maisons individuelles. Il est lié au maître d'ouvrage par l'intermédiaire d'un contrat de louage et est l'unique interlocuteur de ce dernier.

En engageant sa responsabilité juridique et technique, le CMI cumule les missions de modélisation, de construction, d'édition des plans et de supervision jusqu'à l'achèvement des travaux.

e) Vendeurs d'immeuble à construire ou après achèvement

La vente d'immeuble à construire, comme définie par l'Art. 262-1 du Code de la Construction et de l'Habitation, est un contrat par lequel un vendeur s'oblige à construire un immeuble dans un délai bien déterminé. Ce contrat peut être conclu à terme ou en l'état futur d'achèvement (VEFA). Dans ce dernier cas, le plus répandu, le vendeur est également assimilé au maître d'œuvre jusqu'à la fin des travaux.

Dans le cas de la vente après achèvement, on peut citer les exemples suivants :

- Vente de lots dans un programme immobilier à usage d'habitation : une partie des lots est vendue sur plan avant la construction (VEFA), l'autre partie est vendue après la livraison du programme.
- Vente ^{clé} en mains de bâtiments à usage industriel.
- Revente d'un immeuble par un marchand de biens qui a fait procéder à d'importants travaux de rénovation et de restauration.

f) Les mandataires du maître d'ouvrage

Les mandataires du maître d'ouvrage sont des personnes physiques ou morales mandatées par le maître d'ouvrage pour la réalisation d'une mission bien déterminée. En pratique, les deux cas les plus fréquents sont le promoteur immobilier et le maître d'ouvrage délégué.

Le contrat de promotion immobilière est défini par l'Art. 1831-1 du Code Civil comme un mandat d'intérêt commun par lequel une personne dite " promoteur immobilier " s'oblige envers le maître d'un ouvrage à faire procéder, pour un prix convenu, au moyen de contrats de louage d'ouvrage, à la réalisation d'un programme de construction d'un ou de plusieurs édifices ainsi qu'à procéder elle-même ou à faire procéder, moyennant une rémunération convenue, à tout ou partie des opérations juridiques, administratives et financières concourant au même objet. Ce

promoteur est garant de l'exécution des obligations mises à la charge des personnes avec lesquelles il a traité au nom du maître de l'ouvrage.

Dans le cas du maître d'ouvrage délégué, la réglementation en vigueur est celle des marchés publics (loi n° 85-704 du 12 juillet 1985 dite loi MOP).

Hormis le maître d'ouvrage et tous ceux qui lui sont liés par un contrat de louage (constructeurs et constructeurs assimilés), nous retrouvons divers autres intervenants prenant part à l'opération de construction.

Les sous-traitants

Ayant le choix de sous-traiter, les entrepreneurs mandatés par le maître d'ouvrage peuvent choisir d'engager des artisans du bâtiment pour la réalisation de lots bien déterminés. Ce faisant, les sous-traitants ont une obligation de résultats envers leurs contractants. Cependant, en cas de préjudice, ce sont les entrepreneurs qui sont mis en cause par le maître d'ouvrage.

Les négociants fabricants

Ce sont les fournisseurs de matériaux, composants ou éléments d'équipement nécessaires à la réalisation de l'ouvrage. Sous certaines conditions, ils sont solidairement responsables en cas de sinistre.

Le contrôleur technique

C'est une personne agréée par le ministre en charge de la construction (Art. L111-25 et R111-29 du CCH) dont le rôle est d'intervenir à la demande du maître d'ouvrage pour effectuer une mission de contrôle technique des études et des travaux de réalisation de la construction (Art. L111-23, L111-26, R111-38 et R111-39 du CCH).

Le coordonnateur SPS

Le coordonnateur SPS (coordonnateur sécurité et protection de la santé) intervient dans une opération de construction lorsque plusieurs entreprises travaillent sur le même chantier. Son rôle est de veiller au respect des règles de prévention édictées par le code du travail.

1.1.3 Loi Spinetta

Introduisant l'obligation en matière d'assurance construction pour le particulier et le constructeur, la loi Spinetta, encadre la responsabilité de l'assurance dans le secteur du bâtiment.

La réforme du 4 janvier 1978 a permis d'instituer une présomption de responsabilité pesant sur tous les intervenants de l'opération de construction.

En d'autres termes, la loi Spinetta rattache la responsabilité aux dommages ou aux malfaçons résultant d'un acte de construction et survenant dans les dix années suivant la réception des travaux et en conséquence impose la souscription d'une assurance.

L'instauration de cette loi a été motivée à la fois par un manque notoire d'assurance chez les entrepreneurs et les maître d'ouvrage, par des délais de règlements longs de sinistres, par des recours difficiles en cas de disparition du constructeur (courant dans le domaine de la construction) et par des couvertures d'assurance insuffisantes (plafonds de garanties insuffisants, nombreuses exclusions techniques).

Responsabilité des constructeurs

S'étant engagé sur une présomption de responsabilité, le constructeur est responsable des dommages causés à l'ouvrage. Ceci concerne en premier lieu les dommages qui compromettent la solidité de l'ouvrage ou le rendent impropre à sa destination (notion subjective qui s'interprète au cas par cas).

En effet, l'article 1792 du Code Civil corrigé par la loi Spinetta du 4 janvier 1978 stipule que :

« Tout constructeur d'un ouvrage est responsable de plein droit, envers le maître ou l'acquéreur de l'ouvrage, des dommages, même résultant d'un vice du sol, qui compromettent la solidité de l'ouvrage ou qui, l'affectant dans l'un de ses éléments constitutifs ou l'un de ses éléments d'équipement, le rendent impropre à sa destination. Une telle responsabilité n'a point lieu si le constructeur prouve que les dommages proviennent d'une cause étrangère. »

La présomption de responsabilité des constructeurs s'étend aux dommages aux éléments d'équipement « indissociables » (Art. 1792-2 du Code Civil) :

« La présomption de responsabilité établie par l'article 1792 s'étend également aux dommages qui affectent la solidité des éléments d'équipement d'un ouvrage, mais seulement lorsque ceux-ci font indissociablement corps avec les ouvrages de viabilité, de fondation, d'ossature, de clos ou de couvert. Un élément d'équipement est considéré comme formant indissociablement corps avec l'un des ouvrages de viabilité, de fondation, d'ossature, de clos ou de couvert lorsque sa dépose, son démontage ou son remplacement ne peut s'effectuer sans détérioration ou enlèvement de matière de cet ouvrage. »

L'intitulé de la responsabilité décennale découle de la présomption de responsabilité engagée par les constructeurs. Cette dernière s'étale sur 10 ans à compter de la réception de l'édifice. (Art. 1792-4-1 à 1792-4-3 du Code Civil).

Obligation d'assurance

Dans une logique de protection efficace du maître d'ouvrage, la responsabilité du constructeur, comme introduite par la loi ^{Spinetta}, s'inscrit dans un mécanisme d'assurance dit à double détente. Ce mécanisme repose sur la souscription obligatoire de deux assurances complémentaires avant le début des travaux :

- L'assurance Dommages Ouvrage (DO), souscrite obligatoirement par le maître d'ouvrage au profit des propriétaires successifs (Art. L242-1 du Code des Assurances),
- L'assurance de Responsabilité Civile Décennale (RCD), souscrite obligatoirement par chaque constructeur (Art. L241-1 et L241-2 du Code des Assurances).

Ce mécanisme de double détente agit, dans un premier temps, en faisant appel à l'assurance Dommages Ouvrage comme moyen de préfinancement sans établissement préalable de responsabilité. Il permet d'engager rapidement les travaux de réparation, payés en totalité par l'assureur de la garantie DO.

Dans un deuxième temps, l'assureur dommages ouvrage engage un recours envers les responsables des dommages subis et leurs assureurs. Ces derniers sont chargés de supporter la totalité des charges définitives du sinistre.

Exceptions à l'obligation d'assurance

Il est à noter que les obligations d'assurance ne s'appliquent pas à l'Etat lorsqu'il construit pour son compte (Art. L243-1 du Code des Assurances).

De plus, les catégories d'ouvrage suivantes ne sont pas soumises aux obligations d'assurance (Art. L243-1-1 du Code des Assurances)

- Les ouvrages de génie civil (ouvrages maritimes, lacustres, fluviaux, infrastructures routières, portuaires, aéroportuaires, héliportuaires, ferroviaires, de traitement de résidus urbains, de déchets industriels et d'effluents);
- Les voiries, les ouvrages piétonniers, les parcs de stationnement;
- Les réseaux divers, les canalisations, les lignes ou câbles et leurs supports, les ouvrages de transport, de production, de stockage et de distribution d'énergie, les ouvrages de stockage et de traitement de solides en vrac, de fluides et liquides, les ouvrages de télécommunications;
- Les ouvrages sportifs non couverts;
- Les ouvrages existant avant l'ouverture du chantier, à l'exception de ceux qui, totalement incorporés dans l'ouvrage neuf, en deviennent techniquement indivisibles.

Prise d'effet des garanties

La réception de l'ouvrage est le seul déclencheur des garanties. En effet, les garanties DO et RCD ne prennent effet que par l'acte par lequel le maître d'ouvrage déclare accepter l'ouvrage avec ou sans réserve (Art. 1792-6 du Code Civil).

Toutes réserves doivent être spécifiées sur le procès-verbal de réception et doivent être levées dans l'année qui suit.

Dans le cas de l'inexécution des engagements prévus au contrat de louage d'ouvrage par un intervenant, la prise d'effet de la garantie dommage ouvrage intervient entre la date d'ouverture du chantier (DOC) et la date de réception.

1.1.4 Garanties de l'assurance construction

Le présent paragraphe est dédié aux différentes garanties relatives à l'assurance construction. En règle générale, ces garanties sont réparties en deux groupes, celles qui sont obligatoires et celles qui sont facultatives.

Garanties obligatoires

Introduite par la loi Spinetta, l'obligation d'assurance concerne la réparation des dommages portant atteinte à la solidité de l'ouvrage, l'impropriété à destination, la solidité des éléments d'équipements indissociables ou encore les effondrements résultant d'un vice de construction.

D'après le code des assurances, trois types de garanties sont définies :

- La garantie de parfait achèvement (GPA) qui couvre sur un an les désordres relevés par le maître d'ouvrage à la réception et ne relève pas de l'assurance.
- La garantie de bon fonctionnement (GBF) qui couvre sur deux ans les désordres relatifs aux éléments d'équipement dissociables et peut être couverte à titre facultatif.
- Les garanties décennales qui peuvent être souscrites soit par le maître d'ouvrage au profit des propriétaires successifs (DO) soit par les constructeurs (RCD).

L'ensemble de ces garanties sont illustrées dans la figure 1.2.

La présente étude se concentre sur les garanties décennales à travers ses deux variantes, à savoir la garantie DO et la garantie RCD.

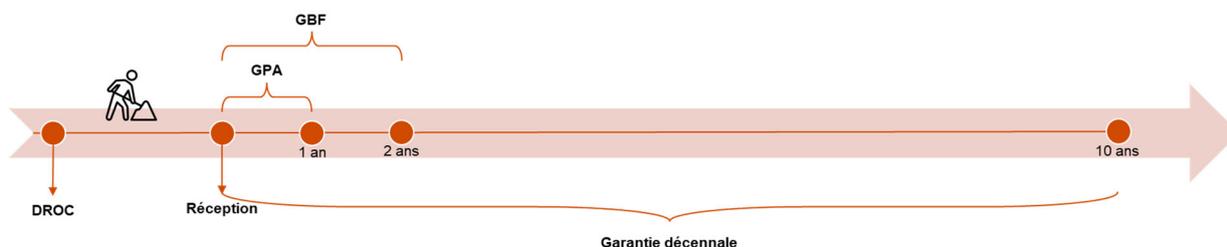


Figure 1.1-2 Illustration des différentes garanties de l'assurance construction

Dommages Ouvrage (DO)

Souscrite par le maître d'ouvrage avant l'ouverture du chantier, la garantie Dommages Ouvrage est une garantie obligatoire qui prend effet à partir de la réception de l'ouvrage et s'étale sur dix ans. Cette garantie concerne tous les dommages dont sont responsables les constructeurs au sens de l'Art. 1792 et l'Art. 1792-2 du Code Civil.

La garantie s'active dès la déclaration d'un sinistre et joue le rôle d'un préfinancement pour la couverture des frais de réparation des éventuels vices. Après établissement des responsabilités, l'assureur DO procède à un recours envers l'assureur RCD du constructeur pour son indemnisation.

Responsabilité Civile Décennale (RCD)

La Responsabilité Civile Décennale est une assurance obligatoire souscrite par tous les acteurs du chantier sur la base de la présomption de responsabilité de ces derniers. Elle permet de garantir une couverture à tous les intervenants de l'opération de construction devant assurer leur responsabilité propre.

Cette garantie est définie par l'Art. 1792 du Code Civil :

« Tout constructeur d'un ouvrage est responsable de plein droit, envers le maître ou l'acquéreur de l'ouvrage, des dommages, même résultant d'un vice du sol, qui compromettent la solidité de l'ouvrage ou qui l'affectant dans l'un de ses éléments constitutifs ou l'un de ses éléments d'équipement, le rendent impropre à sa destination. Une telle responsabilité n'a point lieu si le constructeur prouve que les dommages proviennent d'une cause étrangère. »

Garanties facultatives

En plus des garanties légalement obligatoires, il existe plusieurs autres garanties propres à l'assurance construction. Dans cette sous-section, nous détaillerons ces garanties, leur champ d'action, leurs caractéristiques et le risque qu'elles couvrent.

Deux types de garanties facultatives sont généralement à distinguer. Les garanties adossées à une garantie obligatoire et les autres types de garanties.

a) Garanties facultatives adossées à une garantie obligatoire :

Aux garanties obligatoires RCD et DO peuvent être ajoutées un certain nombre de garanties facultatives. Ces garanties proposées ont pour vocation une différenciation sur le marché des assurances construction.

La liste ci-dessous en présente les principaux exemples :

- Garantie Dommages (effondrement) avant réception : A travers cette garantie, l'assureur s'engage à prendre en charge tous les frais de réparation ou de remplacement des éléments constitutifs de l'ouvrage ayant subi un préjudice, à la suite d'un dommage accidentel consistant en un effondrement ou en résultant, avant la réception de l'édifice. Le dommage accidentel est défini comme un dommage présentant un caractère soudain et fortuit.
- Garantie sur les dommages aux éléments d'équipement dissociables (Garantie de bon fonctionnement) : Elle concerne les désordres atteignant les éléments d'équipement dissociables de l'ouvrage et s'étale sur 2 ans.
- Garantie sur les dommages intermédiaires : Elle couvre les vices de construction qui portent atteinte à la solidité de l'ouvrage et le rendent impropre à sa destination.
- Garantie sur les dommages matériels aux existants : Elle couvre tout dommage causé par les travaux sur les parties existantes sur le site de construction, avant l'ouverture du chantier.

- Garantie sur les dommages immatériels consécutifs : Elle offre un dédommagement au maître d’ouvrage en cas de privation d’un droit ou d’un service rendu par un immeuble ou la perte d’un bénéfice consécutif aux dommages matériels.

b) Autres types de garanties facultatives :

- Tous Risques Chantier (TRC) : Couvrant la partie pré-réception, la police tous risques chantier souscrite par le maître d’ouvrage, fait office d’une garantie globale pour le chantier. Elle couvre un large spectre d’événements source de préjudices importants sur le chantier. Ce spectre s’étale des catastrophes naturelles (glissement de terrain, inondation, tempête, etc.) à l’incendie en passant par une fausse manœuvre d’un engin de chantier.
- Police Unique Chantier (PUC) : La police unique chantier est une police d’assurance qui permet de regrouper à la fois la garantie dommages ouvrage et la garantie responsabilité civile décennale. Souscrite par l’un des intervenants de l’opération de construction, cette garantie couvre ce dernier ainsi que toute partie prenante du chantier.

c) Résumé des garanties usuelles

Le tableau suivant récapitule un panel de garanties d’assurance construction :

Nature du désordre	Police/Garantie	Nature de la garantie	Durée
Ouvrage non exécuté	Multirisque avant réception TRC/TRM	Facultatives	Selon la durée du chantier et la nature de la police souscrite
Effondrement avant réception			
Désordre à la réception	Garantie de parfait achèvement	Obligatoires selon la loi Spinetta comptabilisé en garanties facultatives	1 an
Désordres relatifs aux éléments d’équipements dissociables	Biennale de bon fonctionnement		2 ans
Désordres intermédiaires	RCG ou police « d’abonnement »	Facultatives	resp. 10 ans
Désordres immatériels			
Désordres matériels autres que à l’ouvrage			
Désordres décennaux	DO/RCD	Obligatoires	10 ans
	PUC & RD CNR		

Tableau 1.1-1 Récapitulatif des garanties en assurances construction

1.2 Deux types de provisions

En plus des provisions techniques classiques des branches non-vie, l'assurance construction présente des spécificités techniques. En effet, tous les sinistres liés aux contrats déjà souscrits, et dont la survenance est incluse dans la période de garantie, doivent être provisionnés, ainsi que leur recours et frais de gestion. Cependant, le fait générateur est délicat à dater en construction. La notion de survenance peut donc être interprétée de manière différente selon les assureurs. Certains l'assimilent à la déclaration.

En pratique deux types de provisions pour sinistres doivent être constituées.

- La provision pour sinistres à payer (PSAP) : valeur estimative des dépenses en principal et en frais, tant internes qu'externes, nécessaires au règlement de tous les sinistres survenus et non payés, y compris les capitaux constitutifs des rentes non encore mises à la charge de l'entreprise (Art. R343-7 du Code des Assurances).
- La provision pour sinistres non encore manifestés (PSNEM) : elle couvre les sinistres liés à des DOC antérieures à la date d'inventaire, mais non survenus.

Bien que la distinction soit faite en pratique pour des raisons de simplification de la lecture, d'un point de vue juridique et comptable, la PSNEM fait également partie des PSAP.

Ces provisions et leurs modalités de calcul sont précisées par les articles 143-13, 143-14 et 143-15 du Règlement N°2015-11 de l'ANC.



Figure 1.2-1 Provisions pour sinistres en assurance construction

1.2.1 Provision pour sinistres non encore manifestés

Provision spécifique à l'assurance construction, la PSNEM a été introduite par le législateur pour pallier une succession de résultats techniques négatifs sur le marché français. Cette dernière couvre les sinistres liés à des dates de DOC antérieures à la date d'inventaire mais non survenus.

La PSNEM, calculée selon les méthodes fixées par arrêté du ministre de l'Économie, estime le coût des sinistres non encore manifestés et qui sont susceptibles de se manifester d'ici à l'expiration de la période de prescription décennale.

En effet, le caractère décennal des garanties de l'assurance construction nécessite le provisionnement des sinistres non encore manifestés qui seront ultérieurement rattachés à l'exercice de souscription (DOC). Cette vision permet de capter, sur une profondeur d'historique donnée, la cadence de survenance des sinistres par DOC.

Deux méthodes de calcul de la PSNEM sont possibles, une approche dite réglementaire et une approche dite économique. Ces deux méthodes sont détaillées dans le présent chapitre.

1.2.2 PSNEM réglementaire

Défini par l'Art. 143-14 du Règlement N2015-11 de l'ANC, le calcul de la PSNEM s'effectue par DOC, séparément pour la DO et la RCD selon deux approches, une approche basée sur les primes et une approche basée sur les sinistres. Dans les deux cas, le calcul est fait par vision DOC à l'aide de coefficients réglementaires prédéfinis. Pour chaque année de DOC, la provision retenue est égale à la valeur la plus élevée des montants calculés par les deux approches.

Approche par les primes : méthode rétrospective

Pour un écart de millésime n entre la date d'inventaire et la DOC, la méthode prospective consiste à provisionner une proportion, notée B_n , du montant des primes émises et des primes restant à émettre, nettes des primes à annuler et des frais d'acquisition. Cette proportion est déterminée par un coefficient, noté b_n , suivant le barème réglementaire ci-dessous :

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
b_n	1	1	0,95	0,85	0,75	0,65	0,55	0,45	0,35	0,25	0,2	0,15	0,1	0,05

Il est à noter qu'en cas de sous tarification, la provision retenue par l'approche rétrospective n'est pas suffisante. En d'autres termes, si la prime initiale, complétée par les éventuels produits financiers, ne couvre pas la charge totale des sinistres de la DOC, une autre approche doit être considérée.

Approche par les sinistres : méthode prospective

Pour des chantiers d'ancienneté n et qui se sont manifestés à la date de l'inventaire, la méthode des sinistres consiste à provisionner, par DOC, une proportion, notée A_n du coût total, estimé dossier par dossier, des sinistres afférents aux garanties décennales d'assurance construction délivrées diminué des recours encaissés ou à encaisser. Cette proportion est déterminée par un coefficient, noté a_n , suivant le barème réglementaire ci-dessous :

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
a_n	1	1	0,95	0,85	0,75	0,65	0,55	0,45	0,35	0,25	0,2	0,15	0,1	0,05

En conclusion, la provision pour sinistres non encore manifestés est définie comme suit :

$$PSNEM = \sum_{n=0}^{13} PSNEM_n = \sum_{n=0}^{13} \max(a_n \times A_n; b_n \times B_n).$$

1.2.3 PSNEM économique

Selon le modèle réglementaire de l'article 143-15 du Règlement N2015-11 de l'ANC, le calcul de la PSNEM est une contrainte minimale :

« Lorsque les éléments d'information disponibles conduisent à estimer un coût des sinistres non encore manifestés supérieur au montant résultant de l'application de la méthode prévue au 2 de l'article 143-3 du présent règlement, l'entreprise constitue des provisions à due concurrence du coût estimé. »

La méthode réglementaire ne permet pas de prendre en compte les caractéristiques propres au portefeuille de l'assureur concerné, une société peut ainsi construire son propre barème à partir des données de son portefeuille, en estimant ses propres cadences de manifestation.

Si ce barème est plus prudent que le barème réglementaire, elle doit retenir le montant économique.

Dans cette optique, il est possible d'estimer la charge ultime des sinistres non manifestés, sur la base des statistiques de l'assureur. Cette estimation est appelée PSNEM économique.

Méthode Nombre × Coût moyen

La méthode dite Nombre × Coût moyen, s'articule autour de trois étapes clés :

- Extrapolation des coûts moyens bruts de recours des survenances futures;
- Estimation des nombres de sinistres ultimes des DOC passés;
- Calcul de la charge ultime pour les DOC passés.

a) Extrapolation des coûts moyens bruts de recours des survenances futures :

Au cours de cette étape, une projection selon la méthode Chain Ladder est réalisée sur les triangles Survenance x Développement de charges et de nombres (toutes DOC confondues). A partir des triangles liquidés, les coûts moyens à l'ultime sont calculés par division des charges ultimes par les nombres de sinistres ultimes.

Pour tout année de survenance $i \in \{1, \dots, N\}$, on a

$$\text{coût moyen}_i = \frac{\hat{C}_{i,n}}{\hat{N}_{i,n}}$$

Où,

- $\hat{C}_{i,n}$ représente la charge ultime correspondant à l'année de survenance i ;
- $\hat{N}_{i,n}$ représente le nombre de sinistres ultimes correspondant l'année de survenance i .

En émettant l'hypothèse que le coût moyen d'une année de déclaration n est égal au coût moyen ultime de l'année de survenance n , une extrapolation des coûts moyens ultimes aux survenances futures est opérée après identification d'une tendance et l'application de l'inflation.

b) Estimation du nombre de sinistres ultimes des DOC passés :

L'estimation du nombre de sinistres ultimes pour les DOC passées se fait via une projection de type Chain Ladder du triangle DOC \times Développement du nombre de sinistres (toutes survenances confondues). Ce qui permet d'en déduire une estimation du nombre de sinistres par année de déclaration.

Pour tout année de DOC $k \in \{1, \dots, N\}$, on a

$$\widehat{N}_{k,n} = \widehat{\lambda}_{n-k+1} \times \dots \times \widehat{\lambda}_{n-1} \times \widehat{N}_{k,n-k+1}$$

Où,

- $\widehat{N}_{k,n-k+1}$ représente le nombre de sinistres à l'exercice comptable étudié;
- $\widehat{\lambda}_{n-k+1} \times \dots \times \widehat{\lambda}_{n-1}$ représentent les estimations des facteurs de développement de la méthode Chain Ladder pour le triangle des nombres DOC \times délai de survenance;
- $\widehat{N}_{k,n}$ représente le nombre de sinistres ultime correspondant à l'année de DOC k.

c) Calcul de la charge ultime pour les DOC passées :

Pour chaque année de déclaration, l'estimation du coût moyen de l'année est multipliée par l'estimation du nombre de sinistres pour cette même année. Cela donne le coût ultime des sinistres correspondant aux DOC passées. Ce montant inclut donc :

- La charge ultime des sinistres déclarés = Charge D/D + IBNER;
- La charge des sinistres non déclarés mais survenus = IBNYR;
- La charge des sinistres non déclarés, non survenus = PSNEM.

Le schéma ci-dessous illustre les étapes de la méthode Nombre \times Coût moyen .

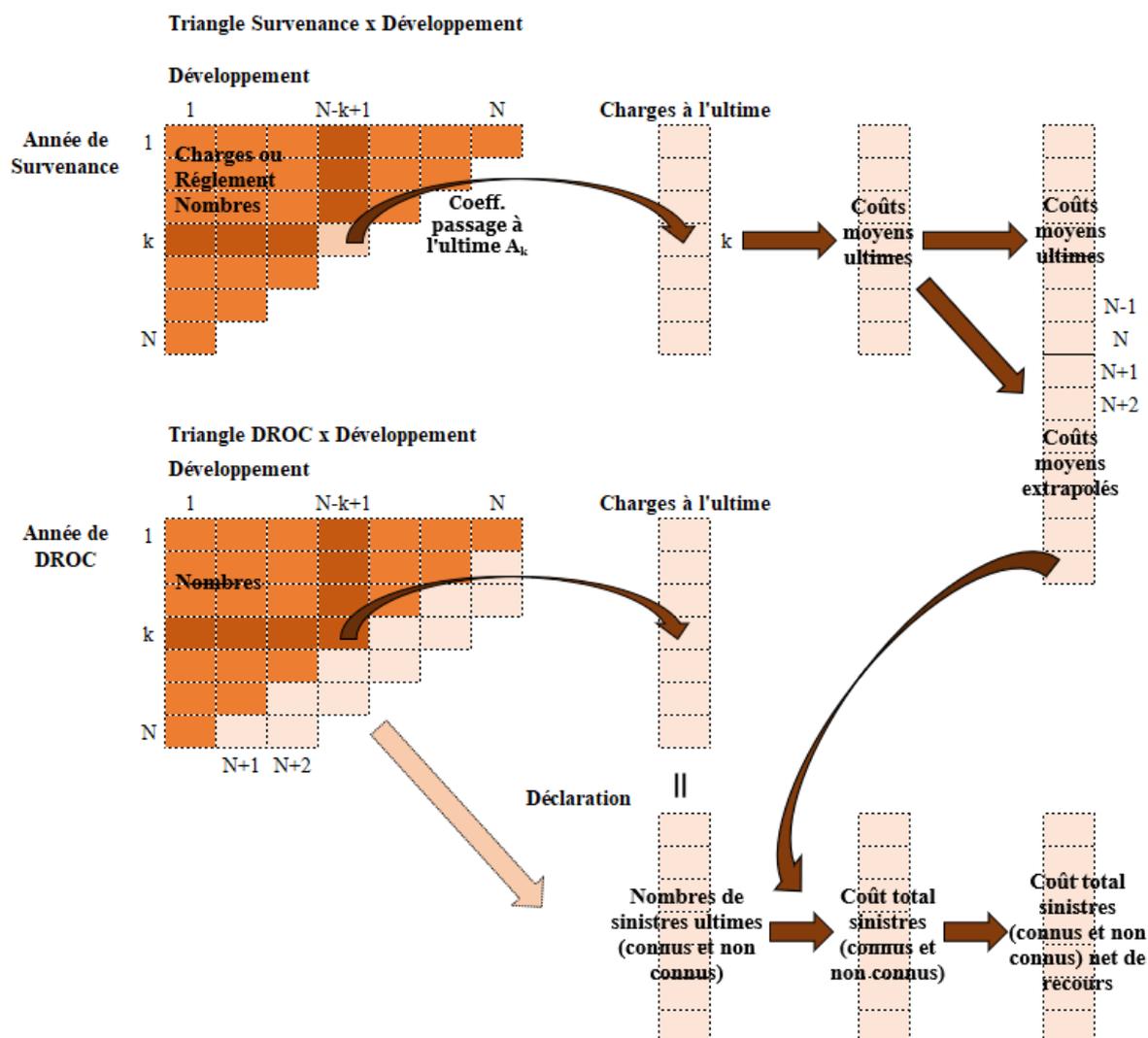


Figure 1.2-2 Synthèse de la méthode « Nombre \times Coût moyen »

d) Limites de la méthode :

La méthode Nombre \times Coût moyen permet d'obtenir une estimation de la charge à l'ultime. Toutefois elle est limitée par les points ci-dessous :

- Impossibilité de différenciation entre les IBNR et la PSNEM;
- Nécessite des données de bonne qualité, sur une longue période, avec un rattachement par DOC, survenance et déclaration corrects;
- Manipulation compliquée des 3 dimensions : DOC, survenance, développement;
- Emission d'hypothèses fortes sur l'évolution du coût moyen sur les survenances futures.

Méthode de projection actuarielle

La méthode de projection dite actuarielle de la PSNEM est une méthode de projection en trois temps. Articulée autour de 3 étapes clés, cette méthode utilise deux liquidations basées selon la méthode Chain Ladder. Une projection est appliquée au triangle Survenance × Développement et l'autre au triangle DOC × Délai de survenance transformé.

a) Première étape : Calcul des IBNR

Au cours de cette étape, une projection Chain Ladder du triangle Survenance × Développement toutes DOC confondues est réalisée. Cette dernière permet non seulement d'obtenir une estimation des IBNR mais aussi de calculer les coefficients de passage à l'ultime. A noter qu'une hypothèse est implicitement émise, à savoir que pour une profondeur d'historique n, les sinistres sont supposés être clos après n années de développement.

Mathématiquement, cela est traduit comme suit :

$$IBNR = \sum_{i=1}^n IBNR_i$$

Où, pour tout $i \in \{1, \dots, N\}$, on a

- $IBNR_i = \hat{C}_{n-i+1, n} \times (1 - \hat{f}_{n-i+1} \times \dots \times \hat{f}_{n-1})$
- $\hat{f}_{n-i+1} \times \dots \times \hat{f}_{n-1}$ représentent les estimations des facteurs de développement de la méthode Chain Ladder pour le triangle des nombres Survenance × Développement.

b) Deuxième étape : Ventilation des IBNR par DOC

En vue d'obtenir une charge totale des sinistres déjà manifestés à la date d'inventaire par DOC, il est nécessaire de pouvoir répartir la PSAP par date de souscription.

L'identification par DOC et par survenance de chaque dossier facilite l'extraction de la base sinistres. Les IBNR, quant à eux, sont rattachés aux années de survenance toutes DOC confondues.

A cet effet, il est nécessaire de disposer d'une méthode permettant le basculement d'une vision par survenance à une vision par DOC. Un moyen de ventiler les IBNR par DOC est d'utiliser une clé de répartition, par exemple les provisions D/D brutes de recours.

Une méthode alternative consiste à utiliser les coefficients de passage à l'ultime par exercice de survenance estimés comme suit :

$$A_k = \frac{\text{Charges de sinistres à l'ultime de l'année k}}{\text{Charge de sinistres à l'ultime de l'année k} - \text{IBNR de l'année k}}$$

La figure ci-dessous schématise les étapes de la ventilation des IBNR par la méthode utilisant les coefficients de passage à l'ultime.

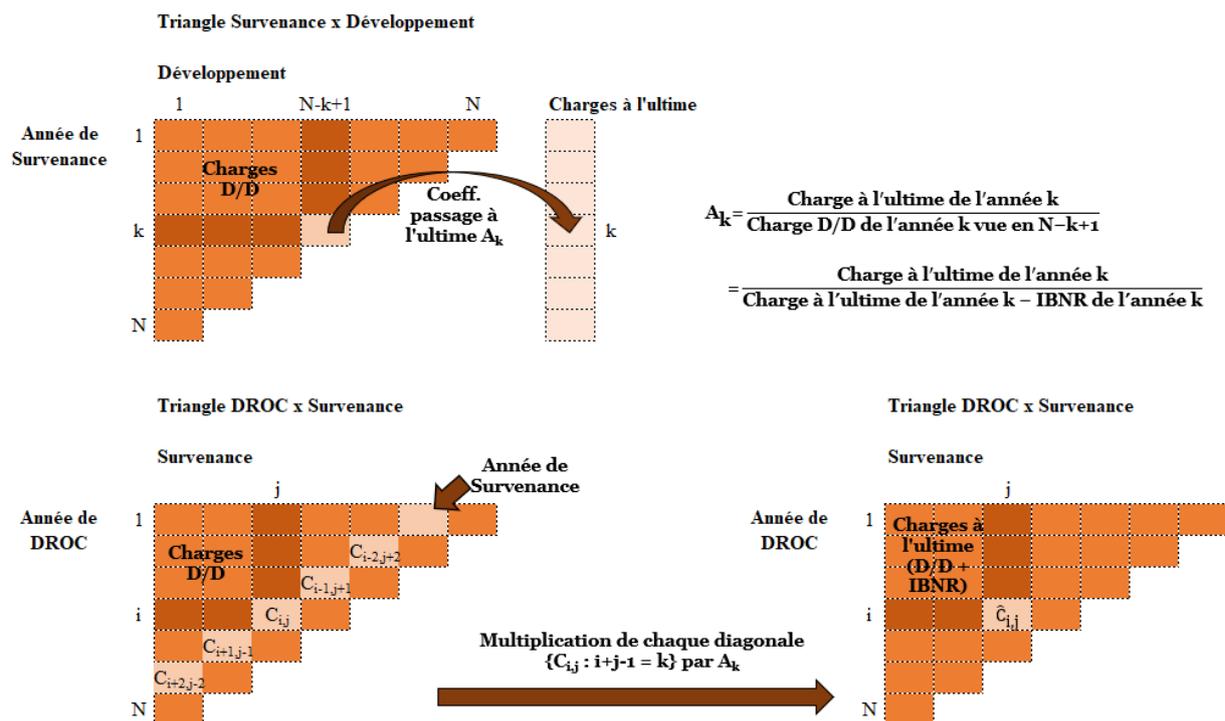


Figure 1.2-3 Mécanisme de ventilation des IBNR par DOC

c) Estimation des PSNEM

En dernière étape, le triangle DOC \times Survenance incluant les IBNR (charges ultimes) est liquidé par une projection type Chain Ladder. Cette liquidation correspond au développement dans le temps des survenances des charges, vu par année de DOC. Cela correspond aux montants de PSNEM par DOC.

$$PSNEM = \sum_{k=1}^n PSNEM_k$$

Où, pour tout $k \in \{1, \dots, N\}$, on a

$$- PSNEM_k = \hat{E}_{n-k+1, n} \times (1 - \hat{\lambda}_{n-k+1} \times \dots \times \hat{\lambda}_{n-1})$$

d) Limite de la méthode

La méthode actuarielle présentée permet une estimation de la PSNEM. Toutefois elle est limitée par les points suivants :

- Idéalement, les estimations doivent se faire d'abord en brut de recours ;
- Afin de passer en valeur nette de recours, un taux de recours doit être estimé (par exemple en extrapolant les taux de recours observés sur les DOC suffisamment matures). Ce taux est en moyenne de 50 % en RC décennale ;

- Les DOC récentes ont tendance à être sous-estimées, d'où la nécessité d'ajuster les résultats sur ces DOC, par exemple en retenant la PSNEM réglementaire sur ces DOC;

Cette méthode n'est applicable qu'en présence de triangles de type Survenance × Développement et DOC × Survenance.

1.3 Marché français de l'assurance construction

1.3.1 Historique du marché

Bien que les produits de l'assurance construction ne soient pas des produits nouveaux, la loi Spinetta en a redéfini les contours.

En effet, depuis l'apparition de la loi Spinetta, la garantie Dommages Ouvrage est gérée par capitalisation, ce qui n'a été le cas, pour la garantie RCD, qu'à partir de 1983.

Sur le plan technique et jusqu'à 1995, le secteur a connu un grand déséquilibre avec des résultats en dégradation continue. Cependant, une amélioration entre 1996 et 2006 a été constatée suivie de plusieurs années de fluctuations oscillant entre des résultats positifs et négatifs entre 2001 et 2004. Sur la période 2005-2016, le marché a connu une succession de résultats techniques bénéficiaires.

L'évolution de la tarification en assurance construction a connu deux phases majeures. Dans une première phase datant des années 80, plusieurs acteurs de l'assurance se sont lancés dans le secteur de la construction à la suite de l'instauration du régime par capitalisation et encouragés par un contexte économique favorable. L'introduction progressive de nouveaux acteurs et la concurrence engendrée ont causé une baisse considérable des tarifs. Il a fallu ainsi 10 ans aux assureurs pour acquérir une maturité suffisante sur le régime par capitalisation et constater une importante sous-tarification ayant engendré des déficits atteignant jusqu'à 160 % des primes. Ces déficits provenaient, notamment, d'une sous-estimation de la charge sinistre d'une part et d'un élargissement du champ des garanties décennales par la jurisprudence, d'autre part.

En 1995, au vu des circonstances, l'obligation de constitution d'une provision spécifique à l'assurance construction a vu le jour. Sous la dénomination PSNEM (Provision pour Sinistres Non Encore Manifestés), cette provision a contribué à l'augmentation des provisions techniques et à alourdir le passif des assureurs.

Par ailleurs, profitant du boom économique précédant le début de la récession des années 2000, le rendement des actifs a permis de compenser les pertes techniques constatées.

1.3.2 Poids économique

En 2020, le marché de l'assurance de biens et de responsabilité (ABR) a totalisé 60,1 Mds € de cotisations en affaires directes ce qui correspond à une hausse de 2,4 % par rapport à 2019.

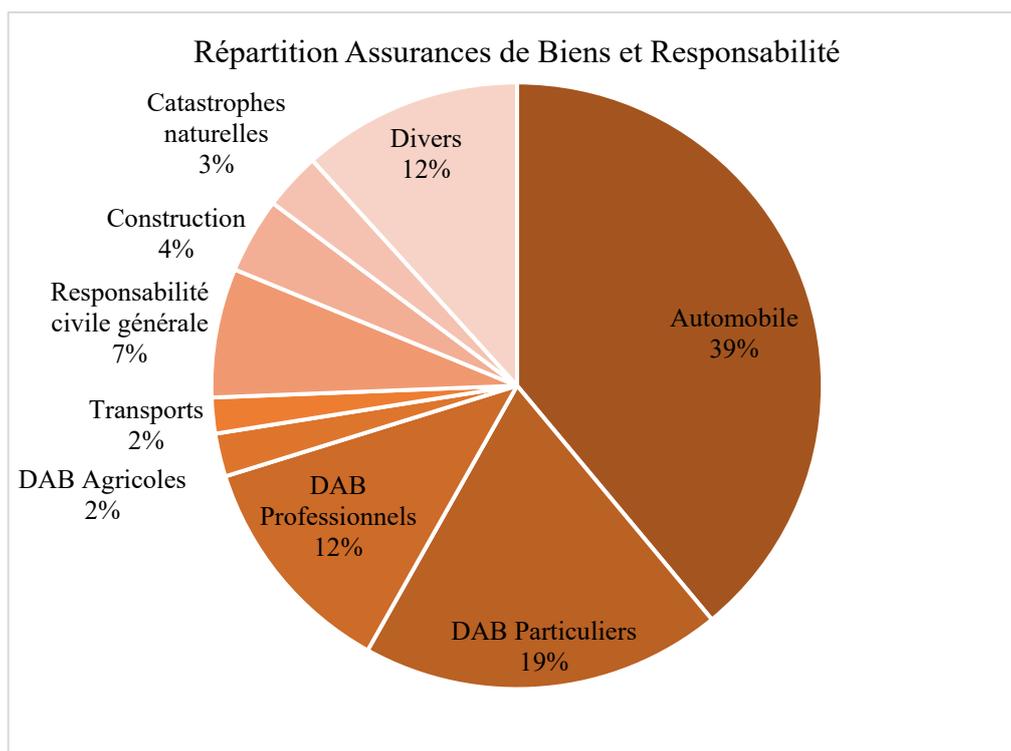


Figure 1.3-1 L'activité des assurances de biens et de responsabilité en 2020 (source : FFA)

Comme le montre la figure 1.2-1, la branche construction a comptabilisé un montant de cotisations s'élevant à 2,3 Mds € correspondant à 4 % des cotisations ABR. On constate néanmoins un léger recul par rapport à 2019.

Le marché de l'assurance construction enregistre en 2020 son 8^{ème} recul en 12 ans. L'ensemble des cotisations sur les 13 dernières années est présenté dans la figure ci-dessous.

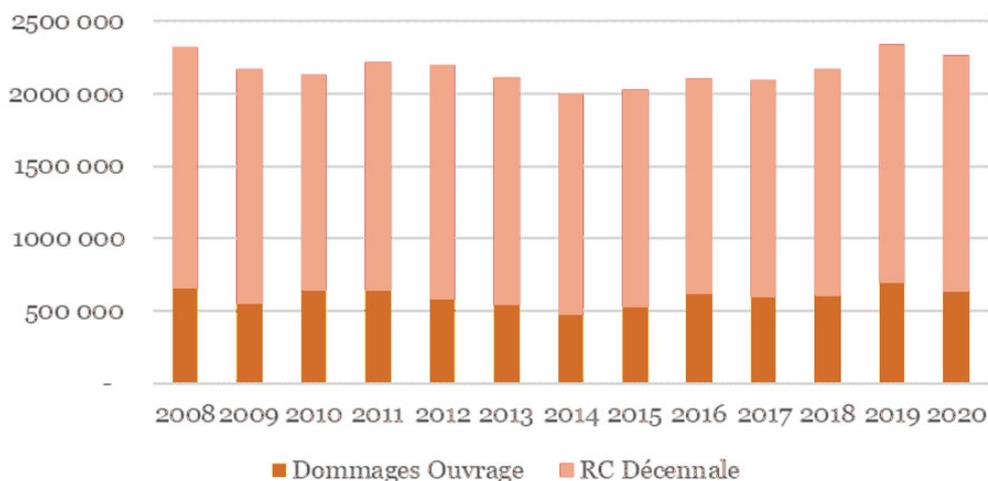


Figure 1.3-2 Cotisations selon le périmètre du marché (Sources : ACPR, FFA)

Le marché de l'assurance construction est marqué par un ralentissement de l'activité depuis 2008, accentué par l'effet de la crise économique. La reprise de l'activité constatée depuis 2017 a été freinée par la crise sanitaire.

En termes de chiffre d'affaires, la RC Décennale représente 70%-75% des primes émises.

1.3.3 Matière assurable

Mises en chantier et nombre de permis de construire

Après dix ans de crise, un rebond du nombre de mises en chantier ainsi que de celui des permis de construire présageait une reprise économique du secteur du BTP. Cette dernière a non seulement été stoppée en 2018 mais la baisse de l'activité s'est creusée en 2020 avec la crise sanitaire.

La figure ci-dessous, présentant l'évolution du nombre de logements mis en chantier sur les 22 dernières années.

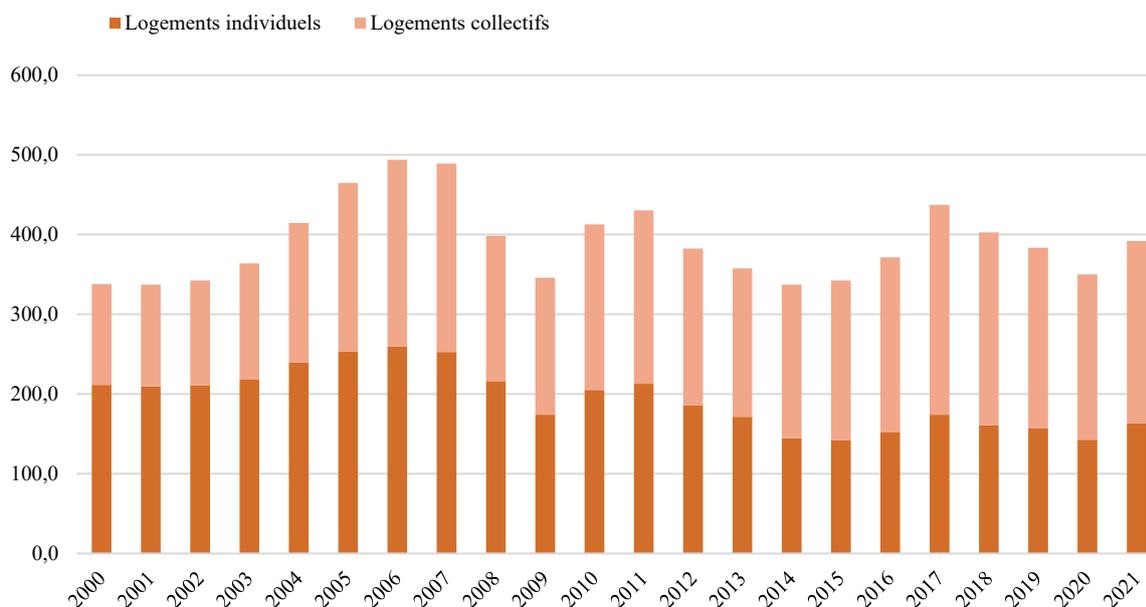


Figure 1.3-3 Evolution du nombre de logements mis en chantier

Depuis 2020, on constate une d'une tendance haussière pour les permis de construire ainsi que la mise en chantier. Plus de 500 000 permis de construire ont été délivrés entre mai 2021 et avril 2022.

Par ailleurs, afin de juger de la stabilité de cette reprise dans le temps, le graphique ci-dessous présente, sur une période de 10 ans, l'évolution du nombre de permis de construire délivrés en France.

Résultat et situation comptable

Malgré le rebond observé en 2018 et 2019, les encaissements de l'assurance construction reprennent leur décroissance en 2020 (-3,3% après +8,0% en 2019).

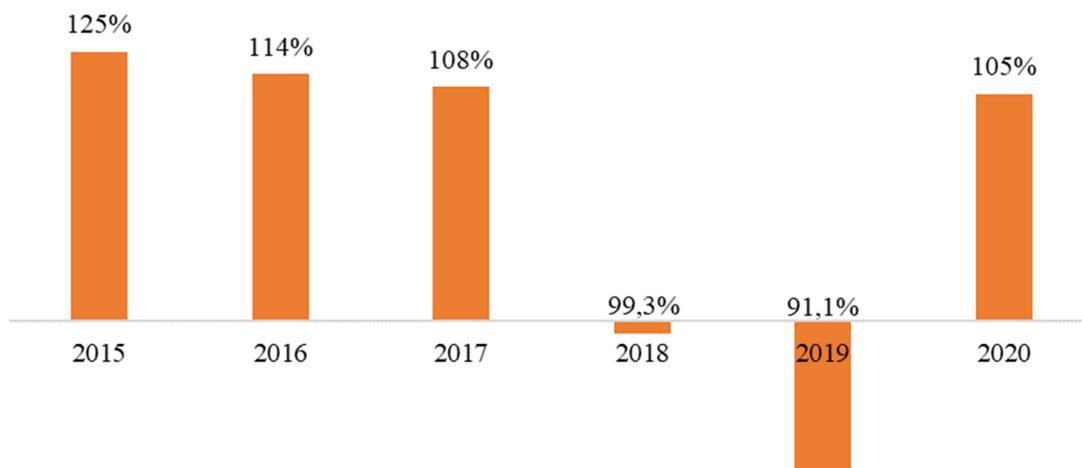


Figure 1.3-4 Evolution du ratio combiné construction (net de réassurance)

La RC Décennale recule de 0,8% et la Dommages Ouvrage de 9,3%. La branche a pâti du recul de l'activité avec une baisse sur l'année du nombre de logements autorisés de 13,6% et du nombre de logements commencés de 7,8%.

Les charges de prestations sont orientées à la hausse (+23,0%), tirées par la RC Décennale (+27,3%). Le ratio combiné augmente, passant à 105,2% en 2020 contre 91,1% en 2019.

1.3.4 Indices

A l'instar des branches longues, en assurance construction, l'impact de l'inflation n'est pas à négliger.

Cette dernière traduit l'évolution du coût des réparations des ouvrages ayant subi des préjudices et étant couverts par les garanties.

En pratique, l'impact de l'inflation est traité d'une manière prospective et rétrospective. Il est appliqué en un premier temps pour redresser les coûts des sinistres passés afin d'analyser des montants nets d'inflation dans une même devise euro (en euros constants) au lieu d'avoir une vision historique d'inflation des montants bruts (en euros courants). Dans un deuxième temps, en émettant une hypothèse sur un indice d'inflation future, une estimation des sinistres futurs tenant compte de l'impact de l'inflation est possible.

En assurance construction, l'évolution des coûts de réparation des ouvrages est liée à la conjoncture du secteur du bâtiment. Toutefois, il existe plusieurs indices spécifiques au secteur qui reflètent cette conjoncture et sur lesquels se basent les acteurs du marché.

Les index BT

Les index BT sont des indices composés à partir de sources diverses (salaires, matériaux, matériel, transport, énergie, ...), basés sur des données publiques de calculs d'évolutions.

Les index BT représentent 37 indices, mesurant mensuellement et hors TVA, l'évolution générale (index BT01) et par spécialité (index BT02 et suivants) le coût des facteurs de production incorporés dans le bâtiment.

L'indice FFB de la construction

L'indice FFB du coût de la construction est un indice trimestriel. Il est calculé par la Fédération Française du Bâtiment à partir du prix de revient d'un immeuble de la région Parisienne. Il enregistre les variations de coût des différents éléments qui entrent dans la composition de l'ouvrage sans prendre en compte la valeur du terrain.

L'indice Insee de la construction

L'indice INSEE du coût de la construction (ICC) permet de suivre l'évolution des prix des marchés de construction des logements neufs. Il est calculé trimestriellement par l'INSEE et concerne uniquement le territoire métropolitain. Les principales variables utilisées sont les montants des marchés de construction et les dépenses d'acquisition du terrain.

Indice INSEE de référence des loyers

L'indice de référence des loyer (IRL) est l'indice de base pour la revalorisation des loyers en cours de bail. En effet, il remplace depuis janvier 2006 l'indice ICC dans cette tâche. Le IRL est publié trimestriellement et est calculé comme la moyenne, sur les douze derniers mois, de l'indice des prix à la consommation hors tabac et hors loyer.

Indice CAC Bâtiment et matériaux

Référence sur la bourse européenne Euronext, l'indice CAC Bâtiment et matériaux est calculé à partir des valeurs boursières des acteurs du bâtiment (Vinci, Saint-Gobain, Lafarge, Bouygues, etc.) cotés au sein du CAC40.

1.4 Synthèse

Une opération de construction est une opération visant à édifier un ouvrage dans un site selon le choix du maître d'ouvrage et conformément à la vision du concepteur. Elle se décompose en quatre phases clés : avant travaux, déclaration de date d'ouverture du chantier (DOC), pendant les travaux et réception. Tout au long de l'opération interviennent plusieurs acteurs allant du maître d'ouvrage aux bureaux d'études passant par les promoteurs et les architectes.

Promulguée le 4 janvier 1978, la Loi Spinetta est la loi structurante du marché de l'assurance construction. En outre, elle a instauré d'une part la notion de présomption de responsabilité de plein droit des constructeurs et d'autre part le mécanisme « à double détention » composé de deux assurances obligatoires devant être souscrites avant l'ouverture du chantier :

- La Dommages Ouvrages (DO), souscrite par le maître d'ouvrage, Elle garantit en dehors de toute recherche de responsabilité (préfinancement rapide), le paiement des travaux de réparation nécessaires à la correction des malfaçons constatées une fois l'ouvrage construit. La garantie débute à la réception des travaux pour une période de 10 ans ;
- La Responsabilité Civile Décennale (RCD), souscrite par tous les constructeurs du fait de la présomption de responsabilité de ces derniers qui s'exerce pendant dix ans. Elle couvre tous les participants à l'acte de construire qui doivent assurer leur responsabilité propre.

Le marché de l'assurance construction en France représente 4% des cotisations de l'assurance de bien et de responsabilité totalisant 2,3 Mds € en 2020. Ce dernier est marqué par un ralentissement de l'activité depuis 2008, accentué par l'effet de la crise économique. Une reprise envisagée en 2017 a été freinée par la crise sanitaire. Le secteur est dit « condensé » avec 80% des cotisations détenues par les cinq premiers assureurs (Mutuelle spécialisée et Assureurs généralistes).

En pratique deux types de provisions pour sinistres doivent être constituées.

- La provision pour sinistres à payer (PSAP) : valeur estimative des dépenses en principal et en frais, tant internes qu'externes, nécessaires au règlement de tous les sinistres survenus et non payés, y compris les capitaux constitutifs des rentes non encore mises à la charge de l'entreprise (Art. R343-7 du Code des Assurances).
- La provision pour sinistres non encore manifestés (PSNEM) : elle couvre les sinistres liés à des DOC antérieures à la date d'inventaire, mais non survenus.
- Bien que la distinction soit faite en pratique pour des raisons de simplification de la lecture, d'un point de vue juridique et comptable, la PSNEM fait également partie des PSAP.

Les PSAP sont calculées par exercice de survenance et couvrent :

- Les sinistres connus évalués dossier par dossier (jugement d'expert, montant garanti, etc.), le coût d'un dossier comprenant toutes les charges externes individualisables (frais d'expertises notamment) ;

- Les sinistres survenus et non déclarés (IBNR) évalués au global selon des méthodes statistiques (pas de préconisation de méthodes).

La PSNEM peut être calculée par la méthode réglementaire à l'arrêté de décembre selon les modalités définies par l'article 143-14 du Règlement N°2015-11 de l'ANC. Le calcul s'effectue par DOC, séparément pour la DO et la RCD.

Le calcul de la PSNEM selon le modèle réglementaire de l'article 143-14 du Règlement N°2015-11 de l'ANC est une contrainte minimale. Une société peut par exemple construire son propre barème à partir des données de son portefeuille, en estimant ses propres cadences de manifestation. Si ce barème est plus prudent que le réglementaire, elle doit retenir le montant économique.

La méthode réglementaire ne permet pas de prendre en compte les caractéristiques propres au portefeuille de l'assureur concerné. Il est possible d'estimer la charge ultime des sinistres non manifestés, sur la base des statistiques de l'assureur : estimation d'une PSNEM économique.

La méthode « classique » de calcul de la PSNEM économique est une méthode qui s'articule en trois étapes :

- 1ère étape : Calcul des IBNR par projection du triangle Survenance \times Développement ;
- 2ème étape : Ventilation des IBNR par DOC ;

3ème étape : Estimation de la PSNEM par projection du triangle DOC \times Survenance ventilé (contenant les IBNR).

Chapitre II : Méthode d'estimation des provisions en assurance construction

L'estimation précise des provisions est essentielle dans le domaine de l'assurance construction. Ce chapitre se penche sur les méthodes d'estimation utilisées pour évaluer les provisions pour sinistres. Nous aborderons deux volets essentiels : les méthodes classiques et le développement de nouveaux modèles 3D pour l'assurance construction.

Avant d'entrer dans les détails des méthodes une première étape d'analyse des données nous permettra de mieux appréhender les spécificités de l'environnement 3D et de poser les bases pour la suite de notre analyse.

2.1 Analyse préliminaire des données

L'analyse préliminaire des données constitue une étape cruciale dans tout projet actuariel. Avant d'entrer dans les détails des méthodes de provisionnement, il est essentiel de comprendre la qualité et la structure des données que nous manipulons. Ce paragraphe se penche sur cette phase exploratoire, où nous examinerons, notamment, les caractéristiques des données en assurance construction.

2.1.1 Présentation et qualité des données

Avant d'entamer le calcul des provisions, une analyse des données est réalisée. La qualité des données étant un facteur non négligeable dans le processus de provisionnement, une étude de cette dernière sert, en outre, à expliquer les résultats trouvés.

Pour des considérations de confidentialité, la source des données demeure anonyme. Cela étant, les données considérées représentent le portefeuille d'une mutuelle française pour la garantie RC décennale brute de recours.

Nous avons récupéré des états statistiques établis et contrôlés sous un format informatique. Ils détaillent l'intégralité de la liquidation des exercices 1996 à 2021, par DOC et par survenance. A partir de ces états, présentant pour les affaires directes des données brutes de réassurance, nous avons constitué :

- Des triangles de liquidation des paiements bruts de recours encaissés au 31 décembre 2021,
- Des triangles de liquidation des charges de sinistres dossier/dossier bruts de recours à encaisser au 31 décembre 2021,

Les éléments utilisés ont fait l'objet de rapprochement avec les données comptables.

Compte tenu de la démarche de provisionnement, plusieurs variables doivent figurer dans la base de données, notamment une date de survenance, une date d'inventaire, une date de DOC et un montant de provision. Une séparation de la variable montant de provision en charge dossier/dossier et montant de paiement permet de pousser les analyses plus loin. Pour les branches présentant un taux de recours élevé, la séparation des provisions entre brut et net de recours permet de capter les différences de comportement entre ces deux derniers.

Cela étant dit, plusieurs contrôles doivent être réalisés :

- Vérification sur le format/les valeurs des données : élimination de toute valeur aberrante ;
- Vérification de l'exhaustivité dans la base : réconcilier la base avec la balance générale ;
- Contrôle de cohérence des dates : vérifier que les dates de survenance des sinistres sont ultérieures à la date DOC, etc. ;
- Réconciliation des charges de sinistres ainsi que les paiements entre la vision par année de survenance et la vision par année de DOC ;

Ces contrôles étant réalisés, quatre triangles sont construits, deux triangles Survenance \times Développement et deux triangles DOC \times Survenance.

2.1.2 Vision Survenance \times Développement (vision PSAP)

La vision Survenance \times Développement permet de visualiser le déroulement des sinistres au cours du temps en fonction de l'année de survenance. Pour chaque année de survenance le déroulement des sinistres est vu par année de développement, définie comme l'écart de millésime entre l'année de survenance et l'année d'inventaire.

Afin de juger la qualité des données triangulées, nous traçons la cohorte des charges et la cohorte des règlements. Ces cohortes représentent graphiquement les données de charges ou de règlements en montant cumulé en fonction des années de développement. Ces derniers permettent en outre de repérer, le cas échéant, une rupture de gestion et valider la cohérence de la gestion au fil des années.

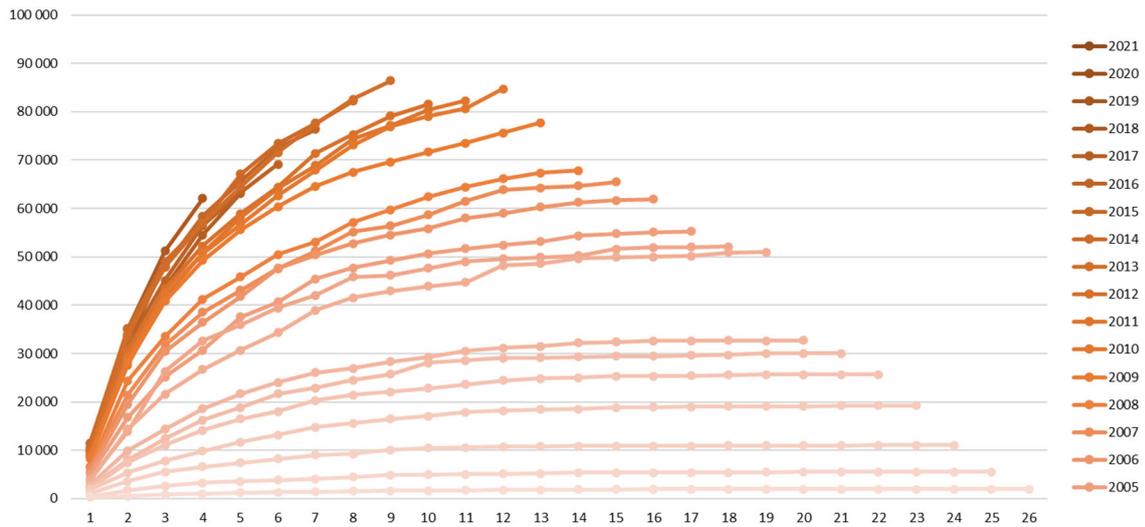


Figure 2.1-1 Cohorte des règlements (en k€) en vision Survenance x Développement

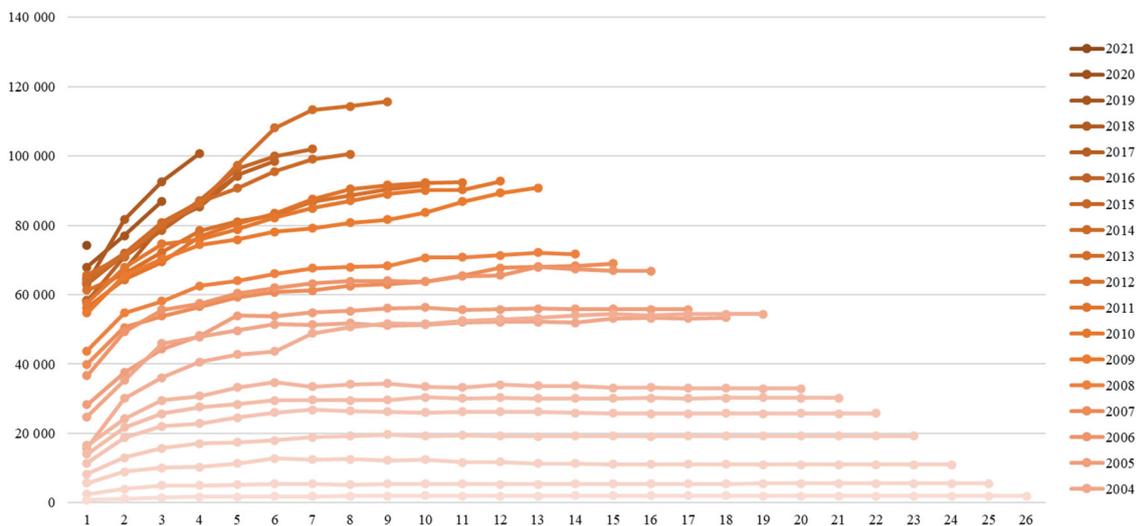


Figure 2.1-2 Cohorte des charges (en k€) en vision Survenance x Développement

On note que les années les plus anciennes semblent être à l'ultime. A partir de 15 ans de développement les paiements ainsi que les charges cumulés se stabilisent. Ceci permet de valider l'hypothèse implicite de Chain Ladder stipulant que la taille des triangles à disposition permet de liquider totalement une survenance donnée. C'est-à-dire, pour le cas étudié, toutes les survenances sont à l'ultimes après 26 ans de développement.

L'allures des différentes cohortes sont similaires. Ceci traduit une gestion homogène dans le temps. Indépendamment des volumes en jeu, la gestion est similaire sur toutes les années de survenance. Ceci permet de conforter l'idée d'estimer la gestion future à travers la gestion passée.

Afin d'approfondir l'analyse de la sinistralité, nous étudions la répartition des charges de sinistres par année de survenance.

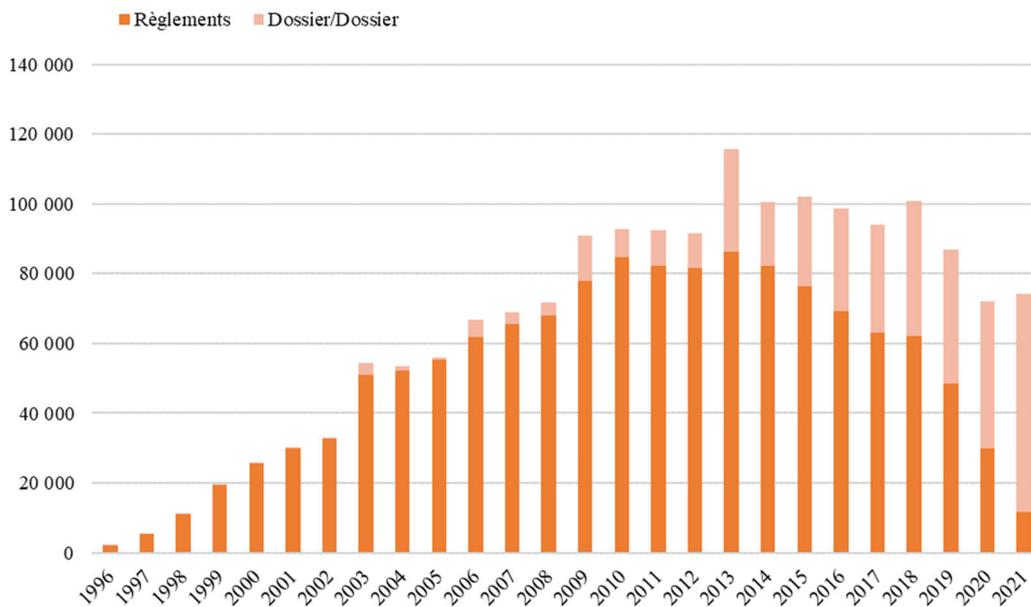


Figure 2.1-3 Répartition des charges par année de survenance

On note que le volume des charges a fortement augmenté au cours du temps. Ceci peut être expliqué par l'expansion du portefeuille d'une part et par l'inflation des coûts de construction d'autre part. On note également que les montants de dossier/dossier convergent vers zéro lorsqu'on tend vers les survenances les plus anciennes. Ceci traduit que les dossiers des survenances les plus anciennes sont clos et que la taille des triangles à disposition permet de liquider les survenances dans leurs totalités.

2.1.3 Vision DOC × Survenance (vision PSNEM)

La vision DOC × Survenance permet de visualiser le déroulement des sinistres au cours du temps en fonction de l'année de DOC. Pour chaque année de DOC le déroulement des sinistres est vu par délai de survenance, définie comme l'écart de millésime entre l'année de DOC et l'année de survenance.

A l'image de l'analyse de la vision PSAP et afin de juger la qualité des données triangulées, nous traçons la cohorte des charges et la cohorte des règlements. Ces cohortes représentent graphiquement les données de charges ou de règlements en montant cumulé en fonction des délais de survenance. Ces derniers permettent en outre de repérer, le cas échéant, un comportement atypique du portefeuille, notamment une dérive de la sinistralité.

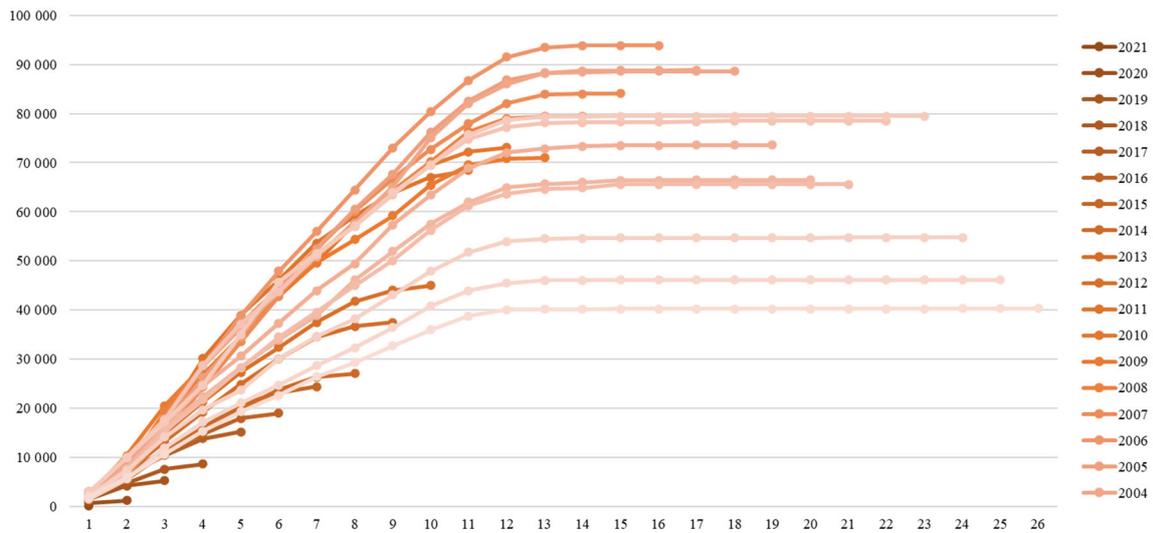


Figure 2.1-4 Cohorte des règlements (en k€) en vision DOC x Délai de Survenance

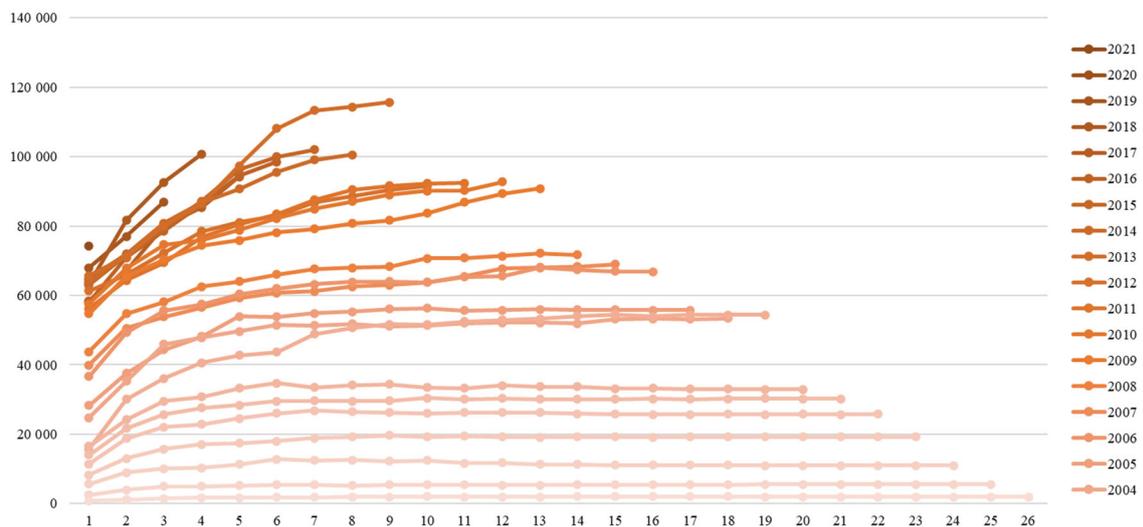


Figure 2.1-5 Cohorte des charges (en k€) en vision DOC x Délai de Survenance

On note que les DOC les plus anciennes semblent être à l'ultime. A partir de 10-12 ans de délai de survenance les paiements ainsi que les charges cumulés se stabilisent. Ceci permet de valider l'hypothèse implicite de Chain Ladder stipulant que la taille des triangles à disposition permet de liquider totalement une survenance donnée. C'est-à-dire, pour le cas étudié, toutes les DOC sont à l'ultimes après 26 ans de développement.

L'allures des différentes cohortes sont similaires. Ceci traduit une gestion homogène dans le temps. Indépendamment des volumes en jeu, la gestion est similaire sur toutes les années de survenance. Ceci permet de conforter l'idée d'estimer la gestion future à travers la gestion passée.

Afin d'approfondir l'analyse de la sinistralité, nous étudions le profil de survenance. Ce dernier permet de visualiser par année de DOC et en moyenne la survenance des sinistres. Ainsi, on peut identifier les périodes de survenance, à partir de l'entame du chantier, sur la durée de couverture décennale.

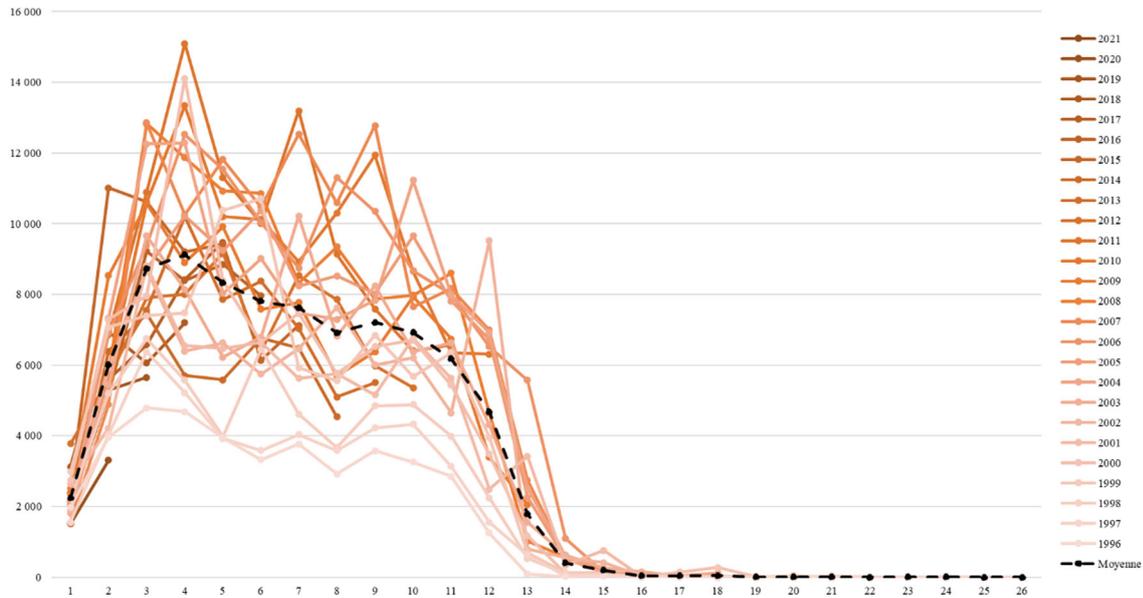


Figure 2.1-5 Profil de survenance

On note que le profil de survenance moyen présente deux « bosses ». La première bosse correspond à un délai de survenance entre 2 et 4 années. Ceci peut être interprété par le fait que les sinistres commencent à survenir, en moyenne, après la fin du chantier (durée moyenne des chantier 1-2 ans) et la levée des réserves après réception. La deuxième bosse est entre 8 et 12 années, ce qui correspond en réalité à la fin de la période de couverture décennale décalée de la durée du chantier.

Le profil de survenance étudié correspond au comportement des portefeuilles d'assurance construction observés sur le marché français.

2.2 Méthodes classiques de provisionnement Non-Vie

Ce paragraphe se concentre sur les méthodes classiques de provisionnement, des approches qui ont fait leurs preuves au fil des années. Nous explorerons ces méthodes en détail, en mettant l'accent sur leur utilisation pour évaluer les provisions pour sinistres. Ces approches traditionnelles constituent la base de notre réflexion avant d'aborder le développement d'un modèle stochastique spécifique à l'assurance construction.

2.2.1 Méthodes déterministes

Dans ce chapitre nous présenterons un panel de méthodes de provisionnement déterministe couramment utilisées en assurance non-vie. Nous nous focalisons cependant sur la méthode Chain Ladder étant l'approche la plus pertinente dans notre étude et la base du développement du modèle 3D.

Méthode Chain Ladder

a) Présentation

La méthode Chain Ladder représente la référence absolue des méthodes de provisionnement Non-Vie. Très simple à mettre en place, cette technique tient également sa popularité au nombre limité d'hypothèses qu'elle nécessite.

Basée sur un modèle multiplicatif, la méthode Chain Ladder suppose que, pour n'importe quelle année de survenance, l'évolution des coûts de sinistres entre deux années de développement successives est linéaire. A cet effet, le développement des sinistres de chaque année de survenance est supposé suivre une cadence de développement.

Le triangle Survenance \times Développement des montants de sinistres est défini par son abscisse i des années de survenance et par son ordonnée j des années de développement. Les diagonales $i + j$ représentent les exercices comptables.

La méthode Chain Ladder s'applique à des montants de sinistres agrégés représentés sous forme de triangles comme précédemment cité.

b) Notation

Pour une branche de durée maximale de n années (tous les sinistres sont supposés être clos après n années de développement), soit $C_{i,j}$ le montant des sinistres survenus l'année i et réglés après j années de développement.

A l'instar de la majorité des méthodes de provisionnement, Chain Ladder repose sur la visualisation en triangle de la vie d'un sinistre. Cette visualisation est illustrée dans le triangle de liquidation qui suit.

		Année de développement						
		1	...	j	...	n-i	...	n
Année de survenance	1	$C_{1,1}$...	$C_{1,j}$...	$C_{1,n-i}$...	$C_{1,n}$

	i	$C_{i,1}$...	$C_{i,j}$...	$C_{i,n-i}$		$C_{i,n}$

	n-j	$C_{n-j,1}$...	$C_{n-j,j}$				$C_{n-j,n}$

	n	$C_{n,1}$						$C_{n,n}$

Figure 2.2-1 Triangle de liquidation Survenance x Développement

Une ligne correspond à une année de survenance donnée et l'évolution, année de développement après année de développement, des règlements effectués est suivi en colonne.

c) Hypothèses

Les hypothèses de la méthode Chain Ladder sont définies comme suit :

(I) Il existe $(f_1, \dots, f_{n-1}) \in \mathbb{R}^{n-1}$ facteurs de développement tels que,

$$\mathbb{E}(C_{i,j+1} | C_{i,1}, \dots, C_{i,j}) = f_j C_{i,j} \quad (\forall i \in \{1, \dots, n\}, j \in \{1, \dots, n-1\}).$$

(II) L'indépendance des années de survenance, i.e.

$(C_{i,1}, \dots, C_{i,n})$ et $(C_{s,1}, \dots, C_{s,n})$ sont indépendants pour tout $(i, s) \in \{1, \dots, n\}^2$ tel que $i \neq s$.

Le calcul des provisions implique de compléter la colonne n correspondant aux charges/paiements ultimes par années de survenance. Ces derniers calculés, le montant de la provision R_i est défini pour chaque année de survenance i par :

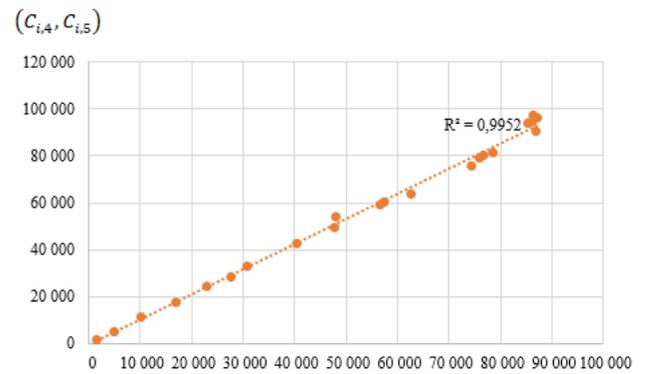
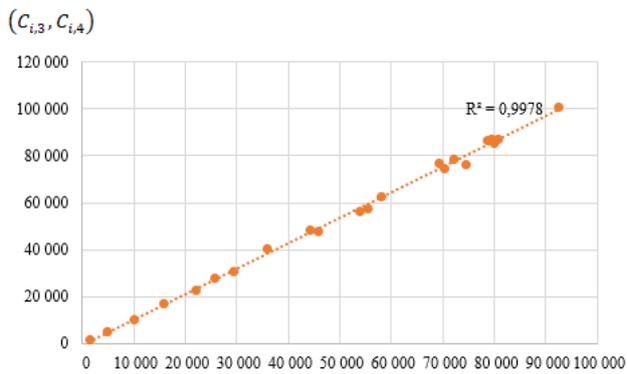
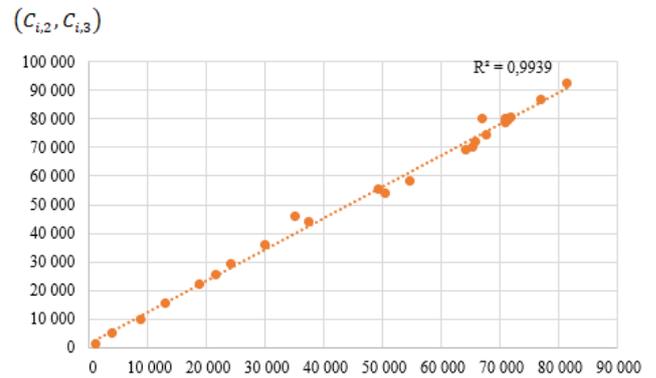
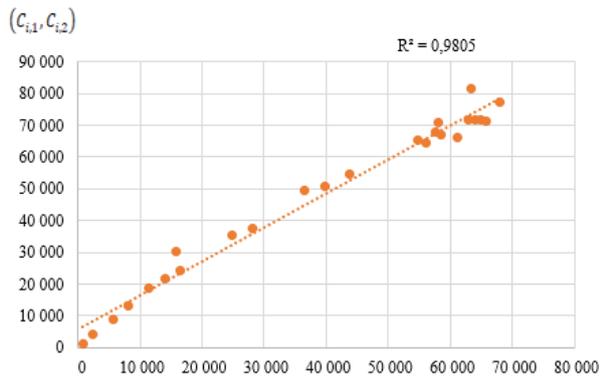
$$\text{pour tout } i \in \{1, \dots, n\}, \quad R_i = C_{i,n} - C_{i,n-i+1}$$

d) Validation des hypothèses

– Hypothèse de linéarité

Description : L'hypothèse est vérifiée à l'aide d'une étude graphique qui consiste à représenter pour chaque année de développement j l'ensemble des couples $(C_{i,j}, C_{i,j+1})$. Ces couples doivent être alignés sur une droite passant par l'origine et de pente f_j

Résultats : Nous observons graphiquement une tendance d'alignement de tous les couples de points. Ci-dessous un exemple pour les 4 premières années.



L'ensemble des graphiques du test de linéarité sont joints en annexe du mémoire (cf. [annexe 3](#)).

Par ailleurs, on peut voir dans le tableau suivant les résultats des régression linéaires effectuées.

Couples	Linéarité	R ²
(C _{i,1} , C _{i,2})	Acceptée	98,05%
(C _{i,2} , C _{i,3})	Acceptée	99,39%
(C _{i,3} , C _{i,4})	Acceptée	99,78%
(C _{i,4} , C _{i,5})	Acceptée	99,52%
(C _{i,5} , C _{i,6})	Acceptée	99,74%
(C _{i,6} , C _{i,7})	Acceptée	99,83%
(C _{i,7} , C _{i,8})	Acceptée	99,97%
(C _{i,8} , C _{i,9})	Acceptée	99,98%
(C _{i,9} , C _{i,10})	Acceptée	99,95%
(C _{i,10} , C _{i,11})	Acceptée	99,91%
(C _{i,11} , C _{i,12})	Acceptée	99,95%
(C _{i,12} , C _{i,13})	Acceptée	99,96%
(C _{i,13} , C _{i,14})	Acceptée	99,98%
(C _{i,14} , C _{i,15})	Acceptée	99,96%
(C _{i,15} , C _{i,16})	Acceptée	99,99%
(C _{i,16} , C _{i,17})	Acceptée	99,99%

Tableau 2.2-1 Test de linéarité - Triangle Survenance x Développement

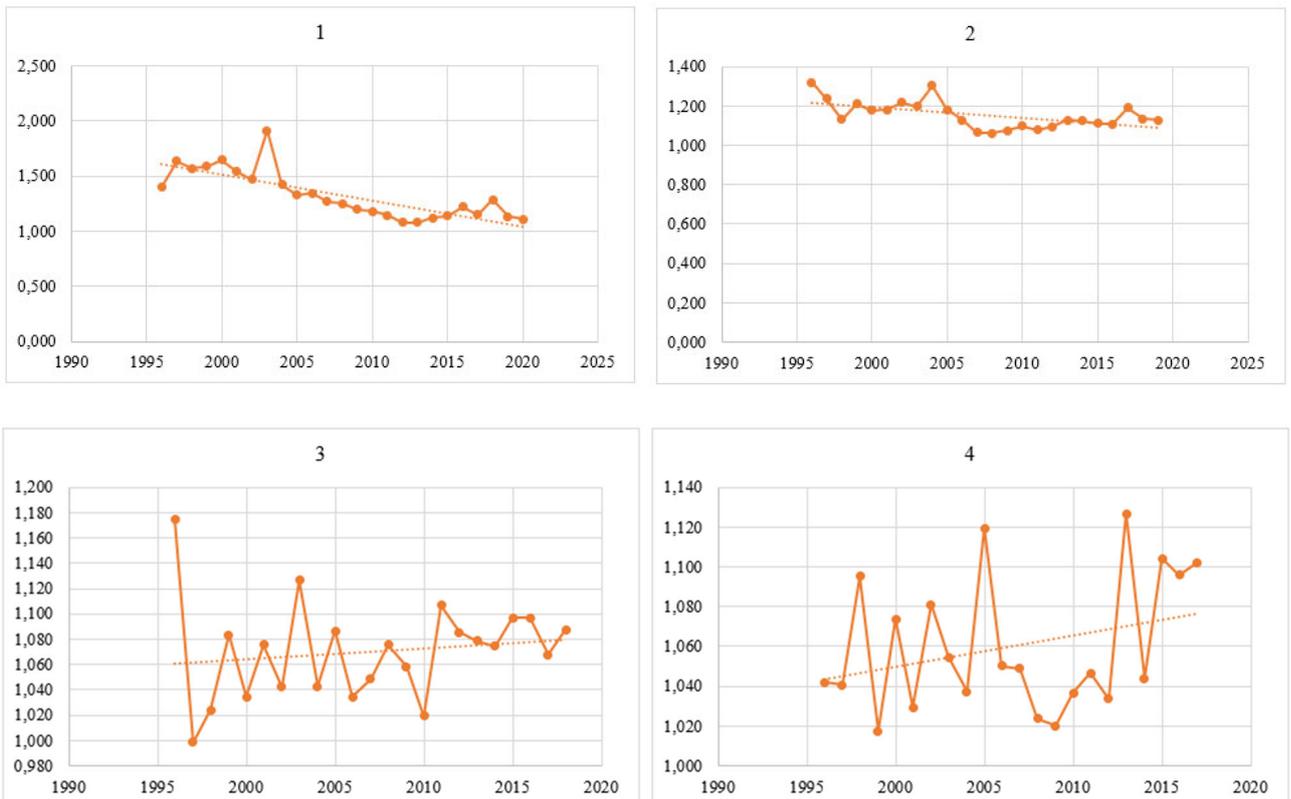
Conclusion : Le test est accepté sur toutes les années de développement et le coefficient R^2 est proche de 1, ce qui valide le test de linéarité.

– Hypothèse d'indépendance

L'hypothèse d'indépendance entre les années de survenance suppose l'indépendance entre les lignes du triangle de liquidation. Elle joue un rôle important car les triangles de provisionnement peuvent présenter des effets calendaires, dont l'apparition peut être liée à des changements dans la législation, la gestion du portefeuille de sinistres ou encore l'inflation, affectant plusieurs années de survenance de la même façon et provoquant ainsi une dépendance entre celles-ci.

Description : La validité de cette hypothèse est testée par un test graphique qui consiste à représenter la jème colonne du triangle (constituée des coefficients de développement) en fonction des années de survenances. L'indépendance des années de survenance signifierait que pour une colonne j du triangle, les coefficients oscillent sans structure aléatoire particulière autour du coefficient de passage moyen.

Résultats : À la lecture des graphiques (cf. [annexe 4](#)), les coefficients oscillent, sans structure aléatoire particulière, autour de la moyenne pour l'ensemble du triangle. Ci-dessous un exemple pour les 4 premières années.



Conclusion : Ce test visuel permet ainsi de conclure que l'hypothèse d'indépendance des années de survenance est globalement acceptable.

e) Estimation

$$\text{pour tout } i \in \{1, \dots, n\}^2, \quad f_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}.$$

Ces derniers sont utilisés en pratique dans le choix des coefficients estimés \hat{f}_j .

La sélection des coefficients $f_{i,j}$ pour le calcul des f_j découle de l'analyse du triangle des $f_{i,j}$ à la lumière de toutes les informations annexes concernant la sinistralité (sinistres exceptionnels, boni-mali, inflation, jurisprudence,...), nature du portefeuille ou encore gestion des sinistres.

La sélection des coefficients par jugement d'expert, illustrée dans Mack (1999), peut être introduite dans l'estimation des coefficients par l'ajout de poids $\omega_{i,j} \in [0,1]$ et d'un coefficient $C_{i,j}^\alpha$, $\alpha \in \{0,1,2\}$ associé aux facteurs de développement individuels.

$$\text{pour tout } j \in \{1, \dots, n-1\}, \quad \hat{f}_j = \frac{\sum_{i=1}^{n-j} f_{i,j} \omega_{i,j} C_{i,j}^\alpha}{\sum_{i=1}^{n-j} C_{i,j}}.$$

f) Résultat des estimations

Survenance	Diagonal	Facteur	IBNR	Ultime
1996	1 974 981	1,000	-	1 974 981
1997	5 509 785	1,000	-	5 509 785
1998	11 021 154	1,000	-	11 021 154
1999	19 247 346	1,000	- 1 676	19 245 670
2000	25 740 656	1,000	- 96	25 740 560
2001	30 158 294	1,001	17 473	30 175 767
2002	32 902 388	1,000	6 440	32 908 828
2003	54 442 418	1,000	3 093	54 445 511
2004	53 413 816	1,000	- 8 305	53 405 511
2005	55 787 979	1,001	78 448	55 866 427
2006	66 820 624	1,002	110 612	66 931 236
2007	68 944 914	1,000	7 894	68 952 807
2008	71 707 639	1,003	214 060	71 921 699
2009	90 853 396	1,001	134 650	90 988 047
2010	92 776 004	1,009	860 373	93 636 378
2011	92 369 271	1,023	2 168 184	94 537 455
2012	91 520 546	1,033	2 992 172	94 512 718
2013	115 687 009	1,042	4 901 476	120 588 485
2014	100 519 618	1,053	5 348 139	105 867 758
2015	101 976 091	1,070	7 111 685	109 087 776
2016	98 515 054	1,101	9 962 433	108 477 488
2017	94 139 585	1,147	13 874 490	108 014 076
2018	100 710 224	1,221	22 249 511	122 959 735
2019	86 910 967	1,309	26 860 660	113 771 627
2020	71 904 595	1,477	34 298 554	106 203 149
2021	74 311 279	1,797	59 223 416	133 534 695
Total	1 709 865 633		190 413 687	1 900 279 320

La méthode de Chain Ladder est basée sur l'hypothèse que les charges de sinistres suivent une tendance stable au fil du temps. Elle utilise les charges passés pour prédire les charges futures. Cependant, elle ne tient pas compte des variations imprévues et des changements de tendance.

Dans notre cas, nous pouvons observer que les provisions IBNR augmentent progressivement au fil des années, reflétant l'accumulation des sinistres non encore réglés. Cela souligne l'importance de modèles plus sophistiqués pour tenir compte des spécificités de l'assurance construction.

Méthode London Chain

Variante de la méthode Chain Ladder, la méthode London Chain suppose que le passage d'une année de développement à l'année qui la succède suit une tendance affine. Autrement dit, pour tout $i, j \in \{1, \dots, n\}$, il existe $(f_j, a_j) \in \mathbb{R}^2$ tel que

$$C_{i,j+1} = C_{i,j} \times f_j + a_j$$

Les coefficients f_j et a_j sont déterminés par la méthode des moindres carrés. Ils sont solution du problème d'optimisation suivant :

$$(\hat{f}, \hat{a}) = \underset{f, a \in \mathbb{R}^{n-1}}{\operatorname{Argmin}} \sum_{i=1}^n \sum_{j=1}^{n-1} (C_{i,j+1} - C_{i,j} \times f_j + a_j)^2$$

Méthode de Bornhuetter-Ferguson

a) Présentation

La méthode de Bornhuetter-Ferguson (B-F) peut être vue comme une variante de la méthode Chain Ladder qui intègre de l'information exogène au triangle. Cependant, elle ne se fonde non pas sur le comportement d'une année sur l'autre des sinistres comme dans Chain Ladder mais plutôt sur une estimation de la charge ultime et de l'évolution du montant des sinistres connus à ce jour vers cette charge ultime. Le plus fréquemment, l'information exogène correspond à un ratio S/P a priori.

L'idée étant d'ajuster la charge de sinistres à l'ultime par rapport à un S/P cible, la méthode applique un schéma de versement des indemnités tiré de Chain Ladder mais se réfère à une approche différente pour en déduire la charge totale définitive.

Les cadences de développement pour cette méthode sont définies par :

$$\text{pour tout } i \in \{1, \dots, n\}, j \in \{1, \dots, n-1\}, \quad \gamma_j = \frac{\mathbb{E}(C_{i,j})}{\mathbb{E}(C_{i,n})}$$

Elles correspondent à un pourcentage du montant des règlements effectués sur les j premières années. A l'image de Chain Ladder, les sinistres sont supposés être clos après n années de développement expliquant le fait que $\gamma_n = 1$.

En conservant les mêmes notations de la méthode Chain Ladder, notamment $f_j = \frac{\mathbb{E}(C_{i,j+1})}{\mathbb{E}(C_{i,j})}$, γ_j s'écrit :

$$\text{pour tout } j \in \{1, \dots, n\}, \quad \gamma_j = \frac{\mathbb{E}(C_{i,j})}{\mathbb{E}(C_{i,n})} = \frac{\mathbb{E}(C_{i,j})}{\mathbb{E}(C_{i,j+1})} \times \frac{\mathbb{E}(C_{i,j+1})}{\mathbb{E}(C_{i,j+2})} \times \dots \times \frac{\mathbb{E}(C_{i,n-1})}{\mathbb{E}(C_{i,n})} = \prod_{k=j}^{n-1} \frac{1}{f_k}$$

En pratique, ces cadences sont estimées par :

$$\text{pour tout } j \in \{1, \dots, n\}, \hat{\gamma}_j = \prod_{k=j}^{n-1} \frac{1}{\hat{f}_k}$$

b) Modèle

Le modèle de B-F est un modèle multiplicatif. Son hypothèse de base peut être présentée comme suit :

Il existe $\alpha = (\alpha_1, \dots, \alpha_n)$ et $\gamma = (\gamma_1, \dots, \gamma_n)$ avec $\gamma_n = 1$ tels que, pour tout $i, j \in \{1, \dots, n\}$

$$\mathbb{E}(C_{i,j}) = \gamma_j \alpha_i$$

$$\text{Avec } \alpha_i = \mathbb{E}(C_{i,n}) \text{ et } \gamma_j = \frac{\mathbb{E}(C_{i,j})}{\mathbb{E}(C_{i,n})}.$$

Avec ce modèle, la relation suivante est vérifiée :

$$\begin{aligned} \mathbb{E}(C_{i,j}) &= \mathbb{E}(C_{i,n-i+1}) + \left(\frac{\mathbb{E}(C_{i,j})}{\mathbb{E}(C_{i,n})} - \frac{\mathbb{E}(C_{i,n-i+1})}{\mathbb{E}(C_{i,n})} \right) \mathbb{E}(C_{i,n}) \\ &= \mathbb{E}(C_{i,n-i+1}) + (\gamma_j - \gamma_{n-i+1}) \mathbb{E}(C_{i,n}) \end{aligned}$$

Connaissant la prime P_i de l'année i et le loss ratio exogène $\frac{S}{P}$, la méthode d'estimation du vecteur retenue α est

$$\hat{\alpha}_i = \frac{S}{P} \times P_i$$

A noter que, l'estimation du loss ratio exogène peut être faite en utilisant d'autres méthodes (par exemple les résultats de la méthode Chain Ladder sur des exercices passés suffisamment matures), des hypothèses issues du Business Plan de l'assureur ou bien encore des données statistiques externes (S/P de marché relatif au segment analysé).

Cela étant défini, le développement du triangle des sinistres cumulés est donné par

$$\hat{C}_{i,j} = C_{i,n-i+1} + (\hat{\gamma}_j - \hat{\gamma}_{n-i+1}) \hat{\alpha}_i$$

Le montant des provisions peut alors être déduit pour chaque année de survenance i par

$$\hat{R}_i = \hat{C}_{i,n} - C_{i,n-i+1}.$$

On en déduit par sommation sur l'ensemble des exercices le montant de la provision totale :

$$\hat{R} = \sum_{i=1}^n \hat{R}_i = \sum_{i=1}^n (1 - \hat{\gamma}_{n-i+1}) \hat{\alpha}_i$$

Méthode de De Vylder

a) Présentation et modèle

A partir des triangles de montants de sinistres décumulés, le modèle de De Vylder est une méthode de modélisation factorielle des $C_{i,j}$ passant par l'estimation de facteurs propres à une année de développement d'une part et de facteurs propres à une année de survenance d'autre part.

Pour tout $i, j \in \{1, \dots, n\}$, cette modélisation peut être définie comme suit :

$$C_{i,j} = \alpha_i \times \beta_j \times \gamma_{i+j}$$

Où

- α_i est un facteur de développement fonction de l'année de survenance i ;
- β_j est un facteur de développement fonction de l'année de développement j ;
- γ_{i+j} est un facteur de développement correspondant à l'impact de l'année comptable $i+j$ traduisant notamment l'inflation.

Sans contrainte supplémentaire, le modèle admet une infinité de solutions de type $(x\alpha_i, \frac{\beta_j}{x}, \gamma_{i+j})$ pour tout $x \neq 0$.

En supposant une inflation constante, une contrainte supplémentaire de type $\sum_{j=1}^n \beta_j = 1$ peut être ajoutée.

Travaillant avec des montants correspondant à des valeurs corrigées de l'inflation γ est supposée égale à 1 par la suite.

Le modèle devient identifiable et les deux autres paramètres peuvent être facilement interprétés :

$$\text{pour tout } i \in \{1, \dots, n\}, \quad \sum_{j=1}^n C_{i,j} = \alpha_i \sum_{j=1}^n \beta_j = \alpha_i$$

Le coefficient α_i représente la charge de sinistres à l'ultime de l'année de survenance i . $\beta_j = \frac{C_{i,j}}{\alpha_i}$ correspond à la proportion réglée au titre de l'année de développement j de la charge totale.

Pour des soucis de simplification, l'inflation sera supposée constante dans ce qui suit.

b) Estimation

Le modèle peut être écrit sous forme de problème d'optimisation des moindres carrés :

$$(\hat{\alpha}, \hat{\beta}) = \underset{\alpha, \beta \in \mathbb{R}^n}{\text{Argmin}} \sum_{i=1}^n \sum_{j=1}^{n-i} (C_{i,j} - \alpha_i \beta_j)^2$$

Ce problème admet comme solution le système suivant :

$$\begin{cases} \alpha_i = \frac{\sum_{j=1}^{n-j} C_{i,j} \beta_j}{\sum_{j=1}^{n-j} \beta_j^2} \text{ pour tout } i \in \{1, \dots, n\} \\ \beta_j = \frac{\sum_{i=1}^{n-j} C_{i,j} \alpha_i}{\sum_{i=1}^{n-j} \alpha_i^2} \text{ pour tout } j \in \{1, \dots, n\} \end{cases}$$

Bien que ce système n'ait pas de solution explicite, il peut être résolu de manière récursive comme suit :

Pour un choix donné de solution initiales $(\beta_j^{(1)})_{j \in \{1, \dots, n\}}$, l'algorithme s'écrit pour $k \geq 1$:

$$\begin{cases} \alpha_j^{(k-1)} = \frac{\sum_{i=1}^{n-j} C_{i,j} \beta_j^{(k-1)}}{\sum_{i=1}^{n-j} (\beta_j^{(k-1)})^2} \text{ pour tout } i \in \{1, \dots, n\} \\ \beta_j^{(k-1)} = \frac{\sum_{i=1}^{n-j} C_{i,j} \alpha_j^{(k-1)}}{\sum_{i=1}^{n-j} (\alpha_j^{(k-1)})^2} \text{ pour tout } j \in \{1, \dots, n\} \end{cases}$$

À noter que le choix des solutions initiales influe directement sur la vitesse de convergence de l'algorithme mais n'influe en aucun cas sur sa convergence. Une solution initiale usuelle est :

$$\begin{cases} \beta_1^{(1)} = 1 \\ \beta_j^{(1)} = \frac{C_{1,j}}{C_{1,1}} \text{ pour tout } j \in \{2, \dots, n\} \end{cases}$$

Cela étant défini, le triangle estimé s'écrit pour tout $i, j \in \{1, \dots, n\}$

$$\hat{C}_{i,j} = \hat{\alpha}_i \times \hat{\beta}_j$$

Le montant des provisions peut alors être déduit pour chaque année de survenance i par

$$\hat{R}_i = \sum_{j=n-i+1}^n \hat{C}_{i,j}$$

On en déduit le montant de la provision totale R_b par sommation sur l'ensemble des exercices :

$$\hat{R} = \sum_{i=1}^n \hat{R}_i$$

Méthode de Verbeek-Taylor

a) Présentation et modèle :

La méthode de Verbeek-Taylor (ou modèle de séparation) est un modèle factoriel. Ce dernier s'applique au cas où l'inflation est inconnue et non constante.

Le modèle peut être défini par

$$C_{i,j} = \lambda_{i+j} \times y_j$$

Le modèle n'étant pas identifiable, une contrainte de type $\sum_{j=1}^n y_j = 1$ est ajoutée.

S'appliquant sur des triangles décumulés, $C_{i,j}$ représente le paiement incrémental pour l'exercice de survenance i et l'exercice de développement j . Le coefficient y_j correspond à la cadence de paiement liée à l'exercice de développement j . λ_{i+j} illustre un effet diagonal qui traduit les modifications de la gestion des sinistres ou la fluctuation de l'inflation. En supposant que la gestion des sinistres ne connaît pas de modification significative, la séquence des coefficients diagonaux permet de capturer l'inflation endogène au triangle (indices d'inflation).

b) Estimation :

L'estimation des paramètres se fait de proche en proche avec les formules suivantes :

$$\hat{\lambda}_{n+1} = \sum_{j=1}^n C_{i,n-i+1}$$

$$y_j = \frac{C_{1,n}}{\hat{\lambda}_{n+1}}$$

$$\hat{y}_j = \frac{\sum_{i=1}^{n+1-j} C_{i,j}}{\sum_{k=j}^n \hat{\lambda}_k} \text{ pour tout } j \in \{1, \dots, n\}$$

$$\hat{\lambda}_j = \frac{\sum_{i=1}^k C_{i,k-i}}{1 - \sum_{j=k+1}^n \hat{y}_j} \text{ pour tout } j \in \{1, \dots, n\}$$

Une extrapolation d'une courbe paramétrique peut par exemple être ajustée sur les valeurs passées ($i + j \leq n + 1$) ou l'usage d'hypothèse externe peut être utilisé pour estimer les valeurs futures.

Les montants à estimer $\hat{C}_{i,j}$ avec $i + j$ compris entre $n + 2$ et $2n$ peuvent ainsi être reconstitués en utilisant la formule suivante :

$$\hat{C}_{i,j} = \hat{\lambda}_k \times \hat{y}_j$$

Le montant de provisions peut alors être déduit pour chaque année de survenance i par

$$\hat{R}_i = \sum_{j=n-i+1}^n \hat{C}_{i,j}$$

On en déduit le montant de la provision totale R_b par sommation sur l'ensemble des exercices :

$$\hat{R} = \sum_{i=1}^n \hat{R}_i$$

2.2.2 Méthodes stochastiques

En complément des modèles déterministes, les modèles stochastiques offrent une estimation des provisions et une quantification de l'erreur de prédiction. Cette mesure permet notamment un pilotage du niveau de prudence des calculs des provisions techniques.

Ces modèles reposent sur une approche probabiliste permettant de mesurer la volatilité autour de la valeur moyenne (espérance + écart-type).

En pratique, un modèle stochastique est généralement défini par une loi décrivant l'évolution de la variable aléatoire étudiée.

A partir d'un Best Estimate (espérance), ces méthodes permettent l'évaluation de la volatilité autour de ce dernier. De surcroît, en émettant une hypothèse paramétrique de loi de probabilité, une distribution des provisions peut être obtenue ainsi que ses différents quantiles.

Un corollaire de ces méthodes peut être l'estimation des quantiles des distributions des provisions et la détermination d'une marge pour risque explicite.

Dans ce chapitre nous présenterons deux méthodes de provisionnement stochastique couramment utilisées en assurance non-vie. Nous nous focalisons cependant sur le modèle de Mack étant l'approche la plus pertinente dans notre étude et la base du développement du modèle 3D.

Quantification de l'erreur de prédiction

Afin d'évaluer l'incertitude du montant estimé, une approche stochastique nécessite la définition d'une mesure permettant de quantifier cette incertitude. Une des mesures utilisées en statistique est la moyenne quadratique de l'erreur prédite MSEP (Mean Squared Error of Prédiction).

Cette mesure capte l'écart moyen entre la valeur prédite et la valeur réelle. A travers le carré défini dans la mesure, plus l'estimation s'éloigne de la réalité plus elle est prise en compte dans le calcul de l'erreur.

En termes mathématiques, en désignant R la variable à prédire et R_b son estimateur, cette mesure est définie par :

$$MSEP(\hat{R}) = \mathbb{E} \left[(\hat{R} - R)^2 \right]$$

Cette quantité peut être réécrite comme suit :

$$MSEP(\hat{R}) = \mathbb{E} \left[(\hat{R} - \mathbb{E}(R))^2 \right] + \mathbb{E} \left[(R - \mathbb{E}(R))^2 \right]$$

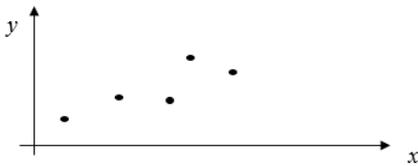
Ceci permet de distinguer deux quantités : une erreur d'estimation et une erreur de processus.

L'erreur de processus, $\mathbb{E} \left[(R - \mathbb{E}(R))^2 \right]$ correspond à un terme de variance lié à l'évaluation des paiements futurs (aléatoires) et mesurant leurs oscillations par rapport à la moyenne.

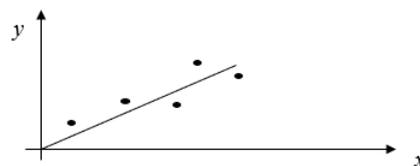
Une erreur d'estimation, $\mathbb{E} \left[(\hat{R} - \mathbb{E}(R))^2 \right]$ liée au modèle qui traduit l'incertitude provenant de l'estimation des paramètres et de leur espérance conditionnelle (les \hat{f}_j).

Le calcul de ces erreurs est illustré dans les graphiques qui suivent.

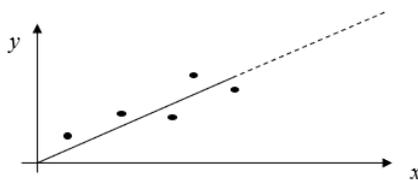
1 – Nous avons une série de données



2 – Nous en cherchons la tendance



3 – Nous prolongeons cette tendance



4 – Les valeurs futures sont « semées » autour de la droite de régression

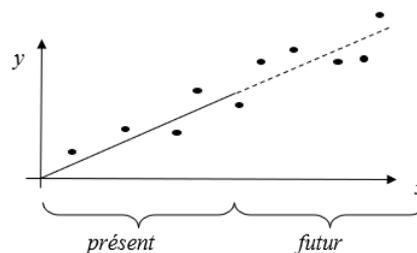


Figure 2.2-3 Illustration de l'erreur de prédiction

De manière schématique, les méthodes de prédiction opèrent en trois parties. Dans un premier temps, les données à disposition sont étudiées pour en tirer une tendance. Cette dernière est prolongée pour finalement faire office de référentiel pour les prédictions futures.

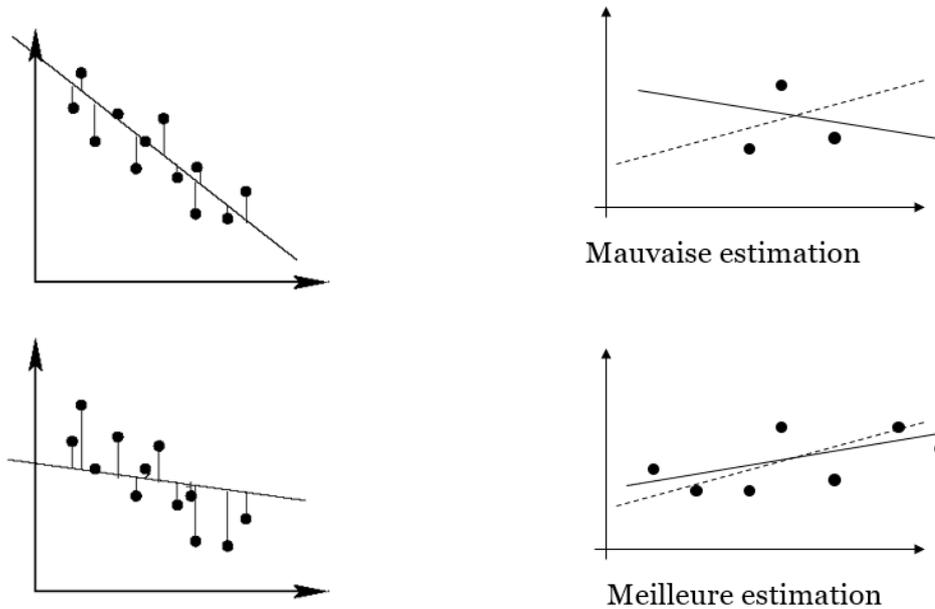


Figure 2.2-4 Illustration de l'erreur d'estimation

Illustré dans la figure 2.2-4, la quantification de l'erreur d'estimation mesure l'écart entre les valeurs prédites et la moyenne historique. En d'autres termes, la moyenne des distances entre la courbe de tendance et les points prédits sur la tranche future de la courbe.

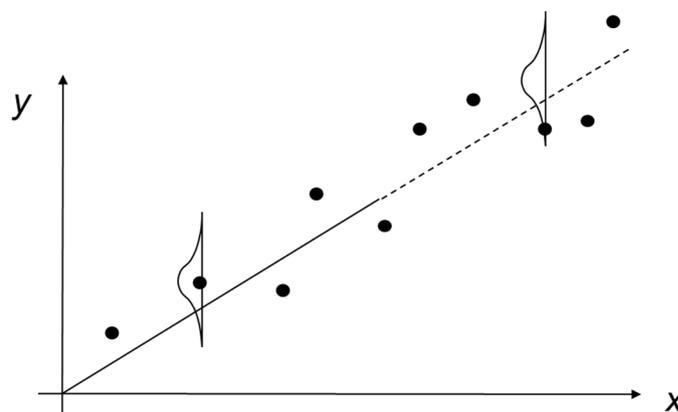


Figure 2.2-5 Quantification de l'erreur de processus

La quantification de l'erreur de processus, présentée dans la figure 2.2-5, mesure l'erreur directement liée au choix du modèle. Cette erreur représente l'écart entre la tendance et les points représentant les valeurs historiques.

Modèle de Mack

a) Présentation

Prenant comme point de départ la méthode Chain Ladder, la méthode de Mack est un modèle stochastique à formule fermée permettant de mesurer l'erreur autour de l'estimation de la provision. Il s'agit de construire un estimateur sans biais pour les facteurs de développement, la charge ultime, les provisions et pour la variance du montant des sinistres.

A cet effet, le triangle des données est assimilé à une variable aléatoire $(C_{i,j})_{i,j \in \{1, \dots, n\}}$ dont les valeurs telles que $i + j \leq n + 1$ sont observables. Par ailleurs, les facteurs de développement individuels $f_{i,j}$ sont des réalisations d'une variable aléatoire d'espérance f_j dont l'estimateur est noté \hat{f}_j .

Le modèle de Mack reprend les hypothèses de la méthode Chain Ladder définie dans le paragraphe [2.2.1](#).

Soit $D = \{C_{i,j} ; i + j \leq n + 1\}$ l'ensemble de l'information connue à la date d'inventaire. Sous les hypothèses **(I)** et **(II)**, le résultat suivant a été établi dans Mack (1993) :

$$\text{pour tout } i \in \{1, \dots, n\}, \quad \mathbb{E}(C_{i,n} | D) = C_{i,n-i+1} \times f_{n-i+1} \times \dots \times f_{n-1}$$

D'autre part, pour tout $j \in \{1, \dots, n\}$, \hat{f}_j est défini comme estimateur de f_j et s'écrit :

$$\text{pour tout } j \in \{1, \dots, n-1\}, \quad \hat{f}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}$$

De plus, sous les hypothèses **(I)** et **(II)**, Mack (1993) a démontré que les estimateurs \hat{f}_j pour tout $j \in \{1, \dots, n-1\}$, étaient sans biais et non corrélés.

Pour obtenir des estimateurs à variance minimale, une troisième hypothèse concernant la variance est ajoutée. Cette hypothèse qui décrit le caractère linéaire de l'évolution de la variance d'une année de développement à l'année de développement qui lui succède est formulée ainsi :

(III) Il existe $(\sigma_1, \dots, \sigma_n) \in \mathbb{R}^{n-1}$ tel que

$$\text{pour tout } i \in \{1, \dots, n\} \text{ et } j \in \{1, \dots, n-1\}, \quad \text{Var}(C_{i,j+1} | C_{i,1}, \dots, C_{i,j}) = \sigma_j^2 \times C_{i,j}$$

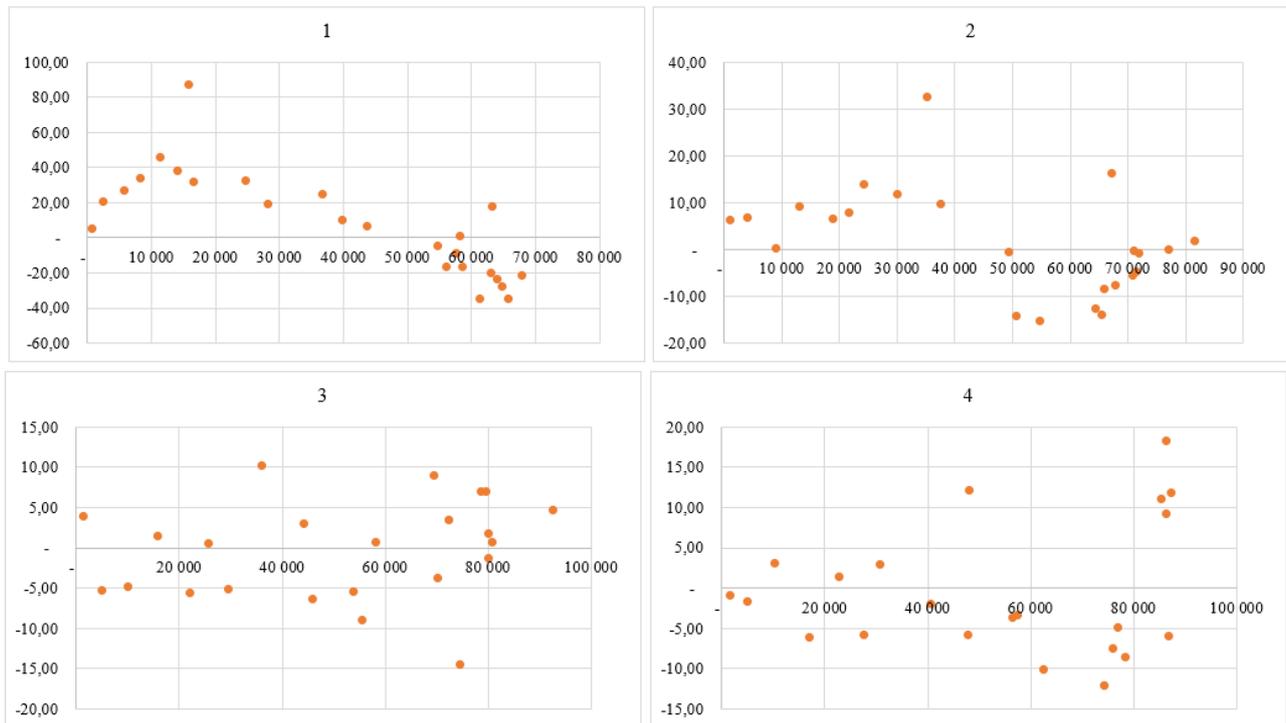
A noter que l'estimateur de Chain Ladder est de variance minimale si l'hypothèse **(III)** est vérifiée.

b) Validation de l'hypothèse

Description : Un test graphique permet de vérifier l'hypothèse de variance à l'aide d'un graphique représentant les résidus normalisés en fonction des montants cumulés $C_{i,j}$. L'hypothèse est rejetée si les résidus présentent une structure non-aléatoire.

$$\text{Les résidus sont définis de la façon suivante : } res_{i,j} = \frac{C_{i,j+1} - f_j C_{i,j}}{\sqrt{C_{i,j}}}$$

Résultats : L'interprétation des graphique (cf. [annexe 5](#)) des résidus est généralement délicate et subjective. Cela nécessite en règle générale une bonne compréhension préalable des données et des phénomènes de gestion susceptibles d'expliquer des résultats observés. Ci-dessous un exemple pour les 4 premières années.



Conclusion : Bien que les résidus ne soient pas uniformément répartis de part et d'autre de l'axe des abscisses sur toutes les années de développement, les graphiques ne dégagent pas une structure non-aléatoire. Ainsi, on peut conclure que l'hypothèse de variance n'est pas remise en cause.

c) Estimation

De plus, l'article de Mack (1993) introduit un estimateur sans biais de σ_j^2 :

$$\left\{ \begin{array}{l} \hat{\sigma}_j^2 = \frac{1}{n-j-1} \sum_{i=1}^{n-j} C_{i,j} \times \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2 \text{ pour tout } j \in \{1, \dots, n-2\} \\ \hat{\sigma}_{n-1}^2 = \min \left(\frac{\hat{\sigma}_{n-2}^4}{\hat{\sigma}_{n-3}^2}, \hat{\sigma}_{n-2}^2, \hat{\sigma}_{n-3}^2 \right) \end{array} \right.$$

Cette dernière expression provient du fait que le rapport de variance entre les années de développement $n-1$ et $n-2$ est supposé être égal au rapport de variance entre les années de développement $n-2$ et $n-3$:

$$\frac{\hat{\sigma}_{n-1}^2}{\hat{\sigma}_{n-2}^2} = \frac{\hat{\sigma}_{n-2}^2}{\hat{\sigma}_{n-3}^2}$$

De surcroît, Mack a supposé que les variances étaient décroissantes sur les développements $n-3$ à $n-1$.

Cela étant défini, l'estimation de la charge à l'ultime peut être calculée par la formule suivante :

$$\hat{C}_{i,n} = C_{i,n-i+1} \times \hat{f}_{n-i+1} \times \dots \times \hat{f}_{n-1}$$

En reprenant les notations de la méthode Chain Ladder, la *MSEP* de l'estimation de la provision \hat{R}_i pour l'année de survenance i est définie par :

$$\begin{aligned} MSEP(\hat{R}_i) &= \mathbb{E} \left[(\hat{R}_i - R_i)^2 \mid D \right] \\ &= \mathbb{E} \left[\left((\hat{C}_{i,n} - C_{i,n-i+1}) - (C_{i,n} - C_{i,n-i+1}) \right)^2 \mid D \right] \\ &= \mathbb{E} \left[(\hat{C}_{i,n} - C_{i,n})^2 \mid D \right] \\ &= MSEP(\hat{C}_{i,n}) \end{aligned}$$

Cette quantité est égale à la somme d'une erreur de processus et d'une erreur d'estimation :

$$MSEP(\hat{C}_{i,n}) = \text{Var}(C_{i,n} \mid D) + (\mathbb{E}(C_{i,n} \mid D) - \hat{C}_{i,n})^2$$

Ainsi l'erreur de prédiction est liée d'une part à une variabilité interne du modèle dite erreur de processus et d'autre part à une erreur d'estimation liée à l'incertitude sur l'estimation.

L'erreur quadratique de prédiction est estimée, pour chaque année de survenance i , sous forme de formule fermée par :

pour tout $i \in \{1, \dots, n\}$,

$$\boxed{\widehat{MSEP}(\hat{R}_i) = E \left[(\hat{R}_i - R_i)^2 \right] = \hat{C}_{i,n}^2 \sum_{j=n-i+1}^{n-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \left(\frac{1}{\hat{C}_{i,j}} + \frac{1}{\sum_{k=1}^{n-j} C_{k,j}} \right)}$$

L'erreur quadratique de prédiction pour le montant de provision totale $\hat{R} = \sum_{i=1}^n \hat{R}_i$ est quant à elle donnée par l'expression suivante :

$$\boxed{\widehat{MSEP}(\hat{R}) = \sum_{i=2}^n \left\{ \widehat{MSEP}(\hat{R}_i) + \hat{C}_{i,n} \left(\sum_{k=i+1}^n \hat{C}_{k,n} \right) \sum_{j=n-i+1}^{n-1} \frac{2\hat{\sigma}_j^2 / \hat{f}_j^2}{\sum_{l=1}^{n-j} C_{l,j}} \right\}}$$

d) Résultats de l'estimation

Les résultats de l'estimation du modèle de Mack sont présentés ci-dessous.

Survenance	IBNR	Erreur de processus	%	Erreur d'estimation	%	Total	%
2010	860 373	1 308 578	152%	614 218	71%	1 445 558	168%
2011	2 168 184	1 606 324	74%	708 842	33%	1 755 772	81%
2012	2 992 172	2 025 141	68%	828 984	28%	2 188 244	73%
2013	4 901 476	2 555 068	52%	1 139 276	23%	2 797 557	57%
2014	5 348 139	2 527 065	47%	1 034 523	19%	2 730 622	51%
2015	7 111 685	2 717 228	38%	1 102 768	16%	2 932 478	41%
2016	9 962 433	3 445 448	35%	1 268 648	13%	3 671 591	37%
2017	13 874 490	4 024 767	29%	1 398 272	10%	4 260 741	31%
2018	22 249 511	5 238 775	24%	1 819 697	8%	5 545 814	25%
2019	26 860 660	5 503 608	20%	1 787 462	7%	5 786 599	22%
2020	34 298 554	6 664 074	19%	1 953 465	6%	6 944 488	20%
2021	59 223 416	14 390 181	24%	4 178 976	7%	14 984 697	25%
Total	90 413 687	19 181 601	10%	14 894 117	8%	24 285 150	13%

Le modèle de Mack semble fournir des estimations raisonnables pour la plupart des années, mais il y a des variations significatives d'une année à l'autre. Les erreurs d'estimation et d'estimation des paramètres varient également, ce qui souligne l'importance de considérer ces deux aspects lors de l'évaluation des provisions.

Il est à noter qu'au vu du volume des IBNR antérieurs à 2010, nous affichons uniquement les résultats postérieurs. Afin de ne pas ajouter un biais à l'estimation, celles-ci sont inclus dans le calcul global de l'erreur de prédiction.

Il est crucial de comprendre les raisons derrière ces erreurs, notamment relativement à de variations saisonnières, de changements de tendance ou d'autres facteurs. Discuter ces résultats avec les équipes métiers, pourrait aider à identifier les forces et les faiblesses du modèle.

e) Limite du modèle

Une des limites du modèle de Mack 2D en assurance construction réside dans son traitement des IBNR. Dans le modèle de Mack 2D, le triangle de base est considéré comme étant déterministe tant dis que ce dernier contient les IBNR, résultat d'une estimation.

Cette hypothèse déterministe peut donc conduire à une sous-estimation de la variance des provisions, ce qui peut à son tour affecter la solvabilité et la stabilité financière de l'assureur.

Modèle linéaire Généralisé

a) Présentation Générale :

Un modèle linéaire est un modèle statistique dans lequel une variable réponse s'écrit comme la somme d'une combinaison linéaire de variables explicatives et d'un terme d'erreur d'espérance nulle et de variance constante. Une méthode des moindres carrés est utilisée pour l'estimation des paramètres du modèle. Par extension de ce modèle, le modèle linéaire généralisé (GLM) permet le traitement de données sous des hypothèses moins contraignantes (relation linéaire entre variable réponse et variables explicatives, distribution des erreurs ...).

Les modèles linéaires généralisés permettent d'étudier la liaison entre une variable dépendante ou réponse Y et un ensemble de variables explicatives notées X_1, \dots, X_K .

Les modèles linéaires généralisés comportent trois composantes :

- des observations indépendantes de la variable de réponse Y dont la loi de probabilité est issue de la famille exponentielle (composante aléatoire);
- un prédicteur linéaire $\eta_i = x_i' \times \beta$, combinaison linéaire des variables explicatives X_1, \dots, X_K (composante déterministe);
- une fonction de lien qui décrit la relation fonctionnelle entre le prédicteur linéaire et l'espérance mathématique de la variable de réponse.

Le modèle peut s'écrire comme suit.

$$\mathbb{E}(Y_i | X) = x_i' \times \beta .$$

où X est la matrice des effets, c'est-à-dire la matrice dont la $i^{\text{ème}}$ ligne est la transposée du vecteur x_i .

Comme indiqué précédemment, la loi de probabilité de la réponse Y_i doit appartenir à la famille exponentielle.

$$f(y_i | (\theta, \phi)) = \frac{\theta y_i - b(\theta)}{a(\phi)} - c(y_i, \phi) .$$

Le paramètre θ est appelé paramètre canonique et caractérise l'espérance de la loi.

Les fonctions a , b et c sont spécifiées en fonction du type de loi exponentielle et le paramètre ϕ (appelé paramètre de dispersion) est supposé connu.

$$\begin{aligned} \mu &= \mathbb{E}(Y_i) = b'(\theta) , \\ \text{Var}(Y_i) &= a(\phi)b''(\theta) . \end{aligned}$$

Les lois suivantes sont des lois appartenant à la famille exponentielle décrite précédemment :

- Loi de Poisson;
- Loi Normale;
- Loi Binomiale Négative, etc...

L'hypothèse de loi étant émise, un choix des variables explicative s'impose. Pour ce faire, étant le paramètre naturel de la loi, θ est estimé comme une fonction linéaire des variables explicatives :

$$\theta = x_i' \times \beta .$$

Cela étant, la fonction de lien s'écrit également $g(\mu) = b^{-1}(\mu)$, ce qui donne :

$$g(\mu) = x_i \times \beta .$$

b) GLM Poisson surdispersé :

Démontré dans Renshaw Verrall (2002), le modèle GLM Poisson surdispersé fournit une estimation des provisions égale à celles résultant de la méthode Chain Ladder. Pour cette raison, cette méthode a été privilégiée par la suite.

Utilisant un triangle de montants de sinistres incrémentaux noté $(X_{i,j})_{i,j}$, l'approche par GLM ODP (Over Dispersed Poisson) repose sur un modèle linéaire généralisé de type Poisson surdispersé avec une fonction de lien *Log*. Il est défini comme suit :

$$\begin{aligned} X_{i,j} &\hookrightarrow ODP(\mu_{i,j}, \phi), \\ \mathbb{E}(X_{i,j}) &= \mu_{i,j}, \\ \text{Var}(X_{i,j}) &= \phi\mu_{i,j}, \\ \text{Log}(\mu_{i,j}) &= c + \alpha_i + \beta_j. \end{aligned}$$

Lorsque le paramètre de dispersion ϕ est égal à 1, le modèle correspond au cas de la loi de Poisson. Le modèle ODP permet de considérer des valeurs supérieures à 1 et s'exonérer de l'hypothèse d'égalité de la variance et de l'espérance et peu réaliste dans le contexte du provisionnement Non-Vie.

A l'image du modèle linéaire, la qualité d'ajustement du modèle aux données se base notamment sur l'analyse des résidus. Les résidus généralement utilisés sont les résidus de Pearson :

$$\text{pour tout } i, j \in \{1, \dots, n\}, \quad r_{i,j}^{(P)} = \frac{X_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\text{Var}(\hat{\mu}_{i,j})}}.$$

La statistique du χ^2 généralisée de Pearson est

$$\chi^2 = \sum_{i+j \leq n} (r_{i,j}^{(P)})^2.$$

Les paramètres c , α_i et β_j sont estimés par la méthode du maximum de vraisemblance, tandis que le paramètre de dispersion ϕ est estimé séparément par

$$\frac{\chi^2}{m - p},$$

où, m est le nombre de données du triangle, soit $m = \frac{n(n+1)}{2}$ et p est le nombre de paramètres à estimer, soit $p = 2 \times n - 1$.

Dans la pratique, la résolution du système passe par une méthode numérique comme la méthode de Newton-Raphson.

c) Approche par rééchantillonnage – Bootstrap

Approche bien connue en actuariat, la méthode Bootstrap est une méthode de simulation fréquemment utilisée sur des jeux de données peu volumineux permettant, en provisionnement, d'obtenir une distribution des réserves. Contrairement aux autres approches, l'approche Bootstrap cherche non pas à calculer un indicateur de risque, la variance, mais à estimer la distribution complète.

Grâce à des simulations de Monte Carlo, le procédé du Bootstrap permet de réaliser des estimations sur la distribution des provisions, de construire des intervalles de confiance et tester les hypothèses.

Dans le cadre du provisionnement, le Bootstrap est habituellement utilisé pour évaluer les erreurs de prédiction pour différentes méthodes de provisionnement, notamment la méthode Chain Ladder et les GLM.

En partant de données réelles, le procédé de Bootstrap génère une multitude de nouveaux échantillons par tirage avec remise dans l'échantillon initial. Ces derniers ont les mêmes caractéristiques que l'échantillon de départ, ce qui permet d'estimer les paramètres du modèle.

Il est à noter que deux approches de Bootstrap peuvent être utilisées. La première approche consiste à directement rééchantillonner les données tandis que la seconde approche est basée sur le rééchantillonnage des résidus. Du fait que le Bootstrap requiert des données indépendantes et identiquement distribuées, la seconde approche est l'unique approche possible dans le cadre du provisionnement.

Les résidus considérés sont les résidus de Pearson définis précédemment.

d) Estimation

Dans le contexte du provisionnement, la méthode Bootstrap est ajustée sur un modèle linéaire généralisé comme défini précédemment. Pour être appliqué à la méthode Chain Ladder, on considère à nouveau un modèle multiplicatif Poisson ODP.

La procédure Bootstrap appliquée au cas du provisionnement peut être décrite comme suit :

- Estimation des paramètres de régression puis des projections $(\hat{x}_{i,j})_{i+j>n}$ afin d'évaluer les réserves \hat{R}
- Calcul des incréments espérés $(x_{i,j}^*)_{i+j\leq n}$
- Obtention des résidus de Pearson $(r_{i,j}^{(P)})_{i+j\leq n}$:
 - Calcul et ajustement des résidus de Pearson avec $m = \frac{n(n+1)}{2}$ le nombre de résidus et p le nombre de paramètres de la régression : $r_{i,j}^{(P)} = \sqrt{\frac{m}{m-p}} \times \frac{x_{i,j} - x_{i,j}^*}{\sqrt{\text{Var}(x_{i,j}^*)}}$.
 - Estimation du paramètre de dispersion ϕ selon l'expression suivante :

$$\hat{\phi} = \frac{1}{m-p} \sum_{i+j \leq n} (r_{i,j}^{(P)})^2.$$

- Rééchantillonnage répété N fois, on se place à la $k^{\text{ème}}$ itération
 - Obtention d'un triangle supérieur de résidus $(x_{i,j}^*)_{i+j \leq n}$ par rééchantillonnage des résidus initiaux
 - Calcul des incréments $x_{i,j}^*$ associés aux précédents résidus en utilisant la relation suivante. $x_{i,j}^{*k} = x_{i,j}^* + r_{i,j}^{*k} \times \sqrt{\text{Var}(x_{i,j}^*)}$.
 - Liquidation du triangle des incréments bootstrappés issus de l'étape précédente $\hat{x}_{i,j}^k$ et estimation des réserves \hat{R}^k
 - On conserve la valeur de \hat{R}^k et on passe à l'itération suivante $k + 1$
- Obtention de la distribution des réserves à l'aide du N-échantillon bootstrappé
 - Calcul de quantiles et construction d'intervalles de confiance
 - Estimation Bootstrap de l'erreur d'estimation : $MSE_{boot}(\hat{R})$
 - Estimation de l'erreur de prédiction : $MSEP_{boot}(\hat{R})$

$$\begin{aligned}
 MSE_{boot}(\hat{R}) &= \sqrt{\text{Var}(\hat{R}) + \frac{m}{m-p} \times MSE_{boot}^2(\hat{R})} \\
 &= \sqrt{\phi \hat{R} + \frac{m}{m-p} \times MSE_{boot}^2(\hat{R})}.
 \end{aligned}$$

2.3 Vers un Modèle de Mack 3D pour l'Assurance Construction

Alors que le modèle de Mack traditionnel se concentre sur les provisions de sinistres en deux dimensions, notre adaptation en trois dimensions tient compte des particularités inhérentes au secteur.

Habituellement, la variance est calculée à l'aide de méthodes stochastiques classiques, qui peuvent ne pas être adaptées à la nature du risque et qui ne tiennent pas compte du biais introduit par la ventilation des IBNR déjà estimés. Le calcul de la variance doit ainsi être conditionné à ces estimations.

Dans ce chapitre, nous introduirons l'environnement 3D de l'assurance construction et nous explorerons les implications de cette généralisation et les étapes nécessaires pour sa mise en œuvre.

2.3.1 Environnement 3D de l'assurance construction

En assurance construction la sinistralité évolue dans un repère muni de trois axes de temps : l'exercice de survenance, l'exercice de développement et l'année d'ouverture du chantier (DOC). La prise en compte de la DOC découle de l'obligation réglementaire de la comptabilisation de la PSNEM rattachée à chaque exercice de souscription.

Tandis que sur les branches traditionnelles d'assurance Non-Vie (Auto, Habitation ...), la sinistralité est représentée sous forme de triangle, les données de sinistralité en assurance construction sont représentées sous forme de pyramide. Cette représentation est illustrée dans le graphe ci-dessous.

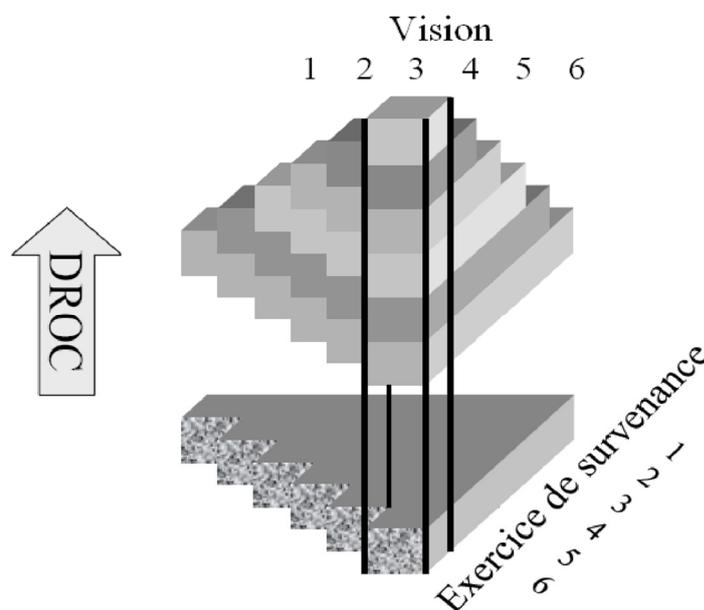


Figure 2.3-1 Visualisation des données en pyramide

Il est à noter que la vision par exercice de développement sous-entend l'écart de millésime entre la date d'inventaire et la date de survenance tandis que la vision par exercice de survenance sous-entend l'écart de millésime entre la date de DOC et l'année de survenance.

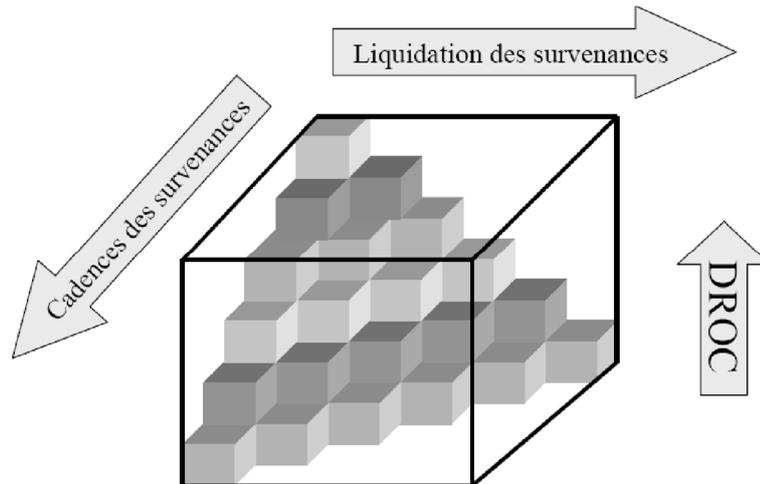


Figure 2.3-2 Visualisation des données en pyramide

Chaque strate de la pyramide correspond à un triangle de dimension $(n - k)^2$ rattaché à une date de

DOC k . Chaque brique de la pyramide correspond à un montant de règlement d'un sinistre souscrit à une date DOC donnée, survenue une année de survenance donnée et rattachée à une date de développement donnée. Il est à remarquer que l'entassement de toutes les couches de la pyramide produit un triangle Survenance \times Développement toutes DOC confondues. La liquidation de ce dernier permet de calculer les PSAP de la même manière qu'en deux dimensions.

2.3.2 Modèle GLM « 3D » : modèle multiplicatif Poisson ODP

Dans le cadre de notre étude, nous introduisons également le "Modèle GLM : modèle multiplicatif Poisson ODP". Dans les sections suivantes, nous allons détailler ce modèle et l'utiliser comme point de comparaison pour évaluer les performances du modèle de Mack 3D. Cette comparaison nous permettra de mieux comprendre les avantages et les limites de chaque approche.

a) Présentation :

On désigne par $C_{i,j,k}$ le montant incrémental de sinistres (règlements, charges ...) pour l'année de DOC i , le délai de survenance j et le délai de développement k . On dispose d'un échantillon constitué de réalisations des montants observés de la « pyramide » $(C_{i,j,k})_{i+(j-1)+(k-1) \leq n}$.

Notre modèle de provisionnement GLM « 3D » est une extension au cas tridimensionnel du modèle Chain Ladder ODP décrit dans England Verrall (2002)¹ et peut être défini comme suit :

¹ England, P.D., Verrall, R.J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, **8:3**, 443-518.

Les $C_{i,j,k}$ sont supposés mutuellement indépendants et sont distribués selon une loi de Poisson surdispersée (ODP) telle que :

pour tout $i, j, k \in \{1, \dots, n\}$,

$$\mathbb{E}(C_{i,j,k}) = C_0 \times x_i \times y \times z_i = e^{\delta + \gamma_i^{DROC} + \gamma_j^{SURV} + \gamma_k^{DVT}} \equiv: e^{\eta_{ijk}}$$

$$\text{Var}(C_{ijk}) = \phi \cdot m_{ijk}$$

Les paramètres à estimer² sont $\delta, (\gamma_i^{DROC})_{2 \leq i \leq N}, (\gamma_j^{SURV})_{2 \leq j \leq N}, (\gamma_k^{DVT})_{2 \leq k \leq N}$. Sous forme vectorielle :

$$\beta: \equiv \begin{bmatrix} \delta \\ \gamma_2^{DROC} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_N^{DROC} \\ \gamma_2^{SURV} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_N^{SURV} \\ \gamma_2^{DVT} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_N^{DVT} \end{bmatrix}$$

ϕ désigne le paramètre de dispersion et la fonction de lien g est le logarithme népérien : $m_{ijk} = g^{-1}(\eta_{ijk}) = e^{\eta_{ijk}}$. Dans la suite, on emploiera la notation $g^{-1}(x) \equiv [g^{-1}(x_1) \ \dots \ g^{-1}(x_n)]'$ pour tout $x = [x_1 \ \dots \ x_n]'$.

b) Estimation des paramètres :

Le vecteur des paramètres β est estimé avec l'estimateur du maximum de vraisemblance (EMV), noté $\hat{\beta}$. Celui-ci est calculé au moyen de la méthode de Newton-Raphson ou la méthode du score de Fisher.

Le paramètre de dispersion ϕ doit également être estimé dans le cas d'un modèle ODP. Il est estimé séparément par un estimateur $\hat{\phi}$: estimateur basé sur le Chi-deux de Pearson ou estimateur basé sur la déviance. Ces deux estimateurs conduisent généralement à des résultats très proches.

² Pour des raisons d'identifiabilité du modèle, le nombre de modalités pour lequel on estime un paramètre est égal à N-1 pour chacune des 3 variables temporelles.

c) Prédiction des C_{ijk} et estimation des provisions :

$$\hat{C}_{ijk} = \hat{m}_{ijk} = e^{\hat{\eta}_{ijk}}$$

$$\text{avec } \hat{\eta}_{ijk} = \hat{\delta} + \hat{\gamma}_i^{DROC} + \hat{\gamma}_j^{SURV} + \hat{\gamma}_k^{DVT}.$$

Lorsque les C_{ijk} sont les montants de charges dossier à dossier³, la prédiction des montants non observés correspond à la somme des IBNR et de la PSNEM :

$$\hat{R} = \sum_{i+(j-1)+(k-1) > N} \hat{C}_{ijk}$$

$$\hat{R} = [1 \quad \dots \quad 1] \cdot g^{-1}(M \cdot \hat{\beta}) \equiv: f(M \cdot \hat{\beta})$$

$$\text{avec } M: \equiv \begin{array}{c} \begin{array}{c} \mathbf{1} \\ \vdots \\ \mathbf{1} \end{array} \quad \begin{array}{c} \text{DOC : } 2 \leq i \leq N \\ \mathbb{I}_{\{i+(j-1)+(k-1) > N\}} \end{array} \quad \begin{array}{c} \text{SURV : } 2 \leq j \leq N \\ \mathbb{I}_{\{i+(j-1)+(k-1) > N\}} \end{array} \quad \begin{array}{c} \text{DVT : } 2 \leq k \leq N \\ \mathbb{I}_{\{i+(j-1)+(k-1) > N\}} \end{array} \end{array}$$

Calcul de l'erreur de prédiction :

$$MSEP(\hat{R}) = \mathbb{E} \left[(R - \hat{R})^2 \right] \approx \text{Var}(R) + \text{Var}(\hat{R})$$

avec

$$R: \equiv \sum_{i+(j-1)+(k-1) > N} C_{ijk}$$

d) Erreur de processus :

Les montants incrémentaux étant supposés indépendants, on peut écrire :

$$\text{Var}(R) = \sum_{i+(j-1)+(k-1) > N} \text{Var}(C_{ijk})$$

On peut alors estimer cette quantité par :

³ Lorsque les C_{ijk} sont les montants de règlements, la prédiction correspond à la somme de la PSAP et de la PSNEM.

$$\widehat{Var}(R) = \hat{\phi} \sum_{i+(j-1)+(k-1) > N} \hat{m}_{ijk} = \hat{\phi} \sum_{i+(j-1)+(k-1) > N} \hat{c}_{ijk}$$

e) Erreur d'estimation :

$$Var(\hat{R}) = Var(f(M \cdot \hat{\beta}))$$

On rappelle que l'EMV d'un modèle linéaire généralisé est asymptotiquement normal et que :

$$\hat{\beta} \xrightarrow[n \rightarrow \infty]{} \mathcal{N}(\beta, I^{-1})$$

où I^{-1} désigne l'inverse de la matrice d'information de Fisher :

$$I: \equiv \left[\left(-\mathbb{E} \left(\frac{\partial^2 \ell}{\partial \beta_k \partial \beta_l} \right) \right)_{k,l} \right]$$

avec ℓ la fonction de log-vraisemblance de l'échantillon.

En utilisant la méthode delta et la normalité asymptotique de l'EMV $\hat{\beta}$, on peut estimer cette $Var(\hat{R})$ par :

$$\widehat{Var}(\hat{R}) = \widehat{Var}(f(M \cdot \hat{\beta})) = \left(\nabla f(M \cdot \hat{\beta}) \right)' \cdot Cov(M \cdot \hat{\beta}) \cdot \nabla f(M \cdot \hat{\beta})$$

où $Cov(M \cdot \hat{\beta})$ désigne la matrice de variance-covariance du vecteur $M \cdot \hat{\beta}$.

On a $f(x) = \sum_i e^{x_i}$. Par conséquent, le vecteur des dérivées partielles de f est égal à :

$$\nabla f = g^{-1}$$

Par ailleurs :

$$Cov(M \cdot \hat{\beta}) = M \cdot Cov(M \cdot \hat{\beta}) \cdot M' = M \cdot I^{-1} \cdot M'$$

On en déduit que :

$$\widehat{Var}(\hat{R}) = g^{-1}(M \cdot \hat{\beta}) \cdot M \cdot I^{-1} \cdot M' \cdot \left(g^{-1}(M \cdot \hat{\beta}) \right)'$$

Dans l'expression ci-dessus, $g^{-1}(M \cdot \hat{\beta})$ est en fait un vecteur dans lequel sont rangés les montants incrémentaux futurs estimés (dont le total est égal à la somme \hat{R} des IBNR et de la PSNEM).

f) MSEP

Notre estimateur de la MSEP s'écrit par conséquent :

$$\widehat{MSEP}(\hat{R}) = \hat{\phi} \sum_{i+(j-1)+(k-1) > N} \hat{C}_{ijk} + g^{-1}(M \cdot \hat{\beta}) \cdot M \cdot I^{-1} \cdot M' \cdot (g^{-1}(M \cdot \hat{\beta}))'$$

2.3.3 Modèle de Mack 3D

a) Présentation

Mentionnée dans la sous-section [2.1.3](#), la méthode de projection actuarielle de la PSNEM se base sur deux liquidations indépendantes des triangles Survenance \times Développement et DOC \times Survenance. Sachant que la ventilation de ce dernier se fait au moyen des coefficients de passage à l'ultime estimés préalablement, appliquer le modèle de Mack séparément pour les deux triangles omet l'aléa introduit par cette ventilation. A cet effet, le triangle DOC \times Survenance est modélisé comme étant conditionné par rapport à l'information détenue.

En gardant les notations du paragraphe [2.2.2](#), le modèle de Mack 3D est modélisé comme suit :

Soit $(CD_{k,l})_{k,l}$ le triangle décumulé DOC \times Délai de Survenance des paiements non ventilés pour une date de DOC k et un délai de survenance l .

Pour une profondeur d'historique n , les coefficients de passage à l'ultime sont définis par :

$$\hat{A}_k = \prod_{i=n-k+1}^{n-1} \hat{f}_i = \hat{f}_{n-k+1} \times \dots \times \hat{f}_{n-1}.$$

Ce qui permet de définir le triangle ventilé par,

$$E_{k,l} = \sum_{s=1}^l CD_{k,s} \times \hat{A}_{k+s-1}$$

La PSNEM est donc déduite pour tout année de DOC k par,

$$PSNEM_k = E_{k,n} - E_{k,n-k+1}$$

b) Hypothèses

A l'image du Mack 2D, les hypothèses suivantes doivent être vérifiées :

(I bis) Il existe $(\lambda_1, \dots, \lambda_{n-1}) \in \mathbb{R}^{n-1}$ facteurs de développement tels que,

$$\mathbb{E}(E_{k,l+1} \mid CD_{k,1}, \dots, CD_{k,l}) = \lambda_l E_{k,l} \quad (\forall k \in \{1, \dots, n\}, l \in \{1, \dots, n-1\}).$$

(II bis) L'indépendance des années de DOC, i.e.

$(CD_{k,1}, \dots, CD_{k,n})$ et $(C_{s,1}, \dots, C_{s,n})$ sont indépendants pour tout $(k, s) \in \{1, \dots, n\}^2$ tel que $k \neq s$.

c) Validation des hypothèses du modèle

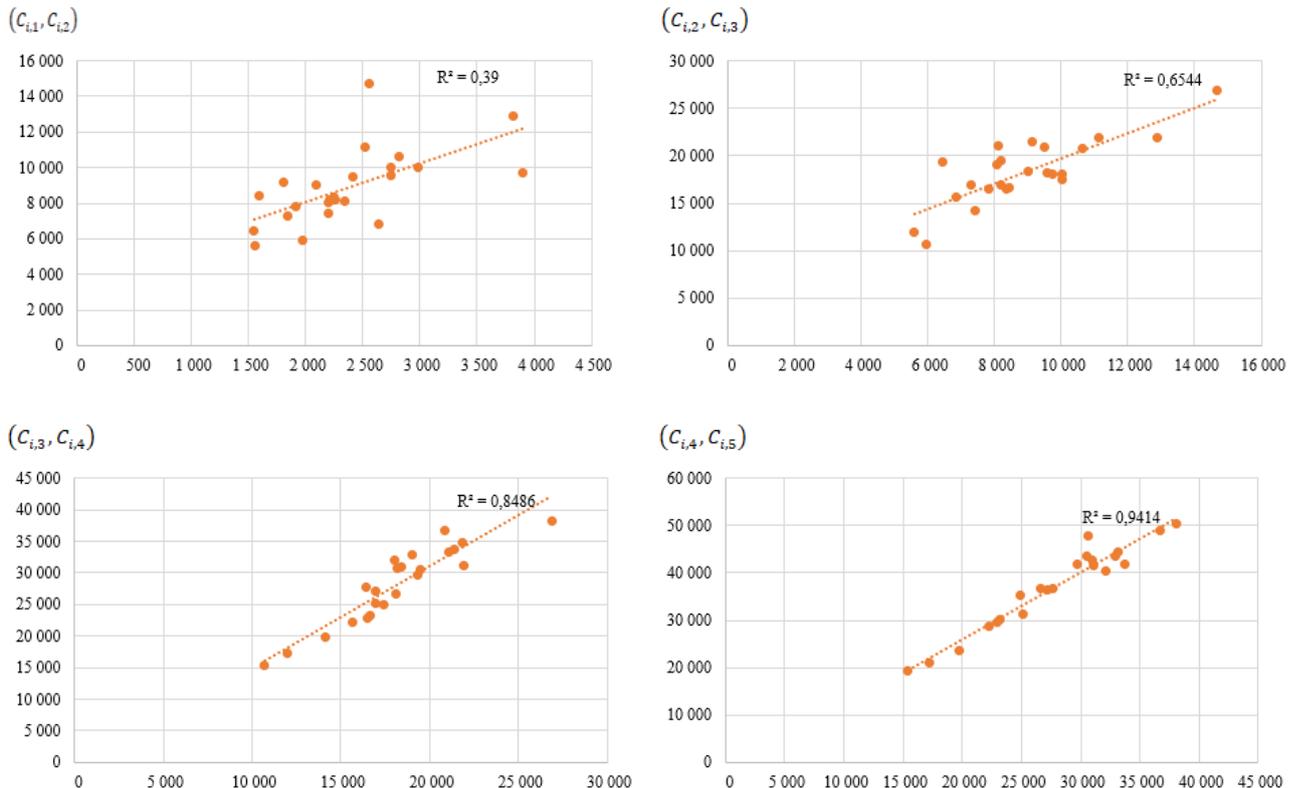
L'objectif de cette section est de valider les hypothèses sous-jacentes au modèle, à savoir les 3 hypothèses sous-jacentes au modèle de Mack 3D.

En effet, le modèle de Mack repose les trois hypothèses suivantes :

- Hypothèse de linéarité ;
- Hypothèse de variance ;
- Hypothèse d'indépendance des années de survenance.
- Hypothèse de linéarité

Description : L'hypothèse est vérifiée à l'aide d'une étude graphique qui consiste à représenter pour chaque délai de survenance j l'ensemble des couples $(C_{i,j}, C_{i,j+1})$. Ces couples doivent être alignés sur une droite passant par l'origine et de pente f_j

Résultats : Nous observons graphiquement une tendance d'alignement de tous les couples de points. Ci-dessous les résultats des 4 premières années.



Cet alignement est moins prononcé pour les couples $(C_{i,1}, C_{i,2})$ et les couples $(C_{i,2}, C_{i,3})$ mais garde cependant une tendance linéaire. On peut voir dans le tableau suivant les résultats des régression linéaires effectuées.

Couples	Linéarité	R ²
(C _{i,1} , C _{i,2})	Acceptée	39,00%
(C _{i,2} , C _{i,3})	Acceptée	65,44%
(C _{i,3} , C _{i,4})	Acceptée	84,86%
(C _{i,4} , C _{i,5})	Acceptée	94,14%
(C _{i,5} , C _{i,6})	Acceptée	97,43%
(C _{i,6} , C _{i,7})	Acceptée	98,05%
(C _{i,7} , C _{i,8})	Acceptée	99,06%
(C _{i,8} , C _{i,9})	Acceptée	99,25%
(C _{i,9} , C _{i,10})	Acceptée	99,57%
(C _{i,10} , C _{i,11})	Acceptée	99,90%
(C _{i,11} , C _{i,12})	Acceptée	99,41%
(C _{i,12} , C _{i,13})	Acceptée	99,71%
(C _{i,13} , C _{i,14})	Acceptée	99,99%
(C _{i,14} , C _{i,15})	Acceptée	99,99%
(C _{i,15} , C _{i,16})	Acceptée	100,00%
(C _{i,16} , C _{i,17})	Acceptée	100,00%

Tableau 2.3-1 Test de linéarité - Triangle DOC x Survenance

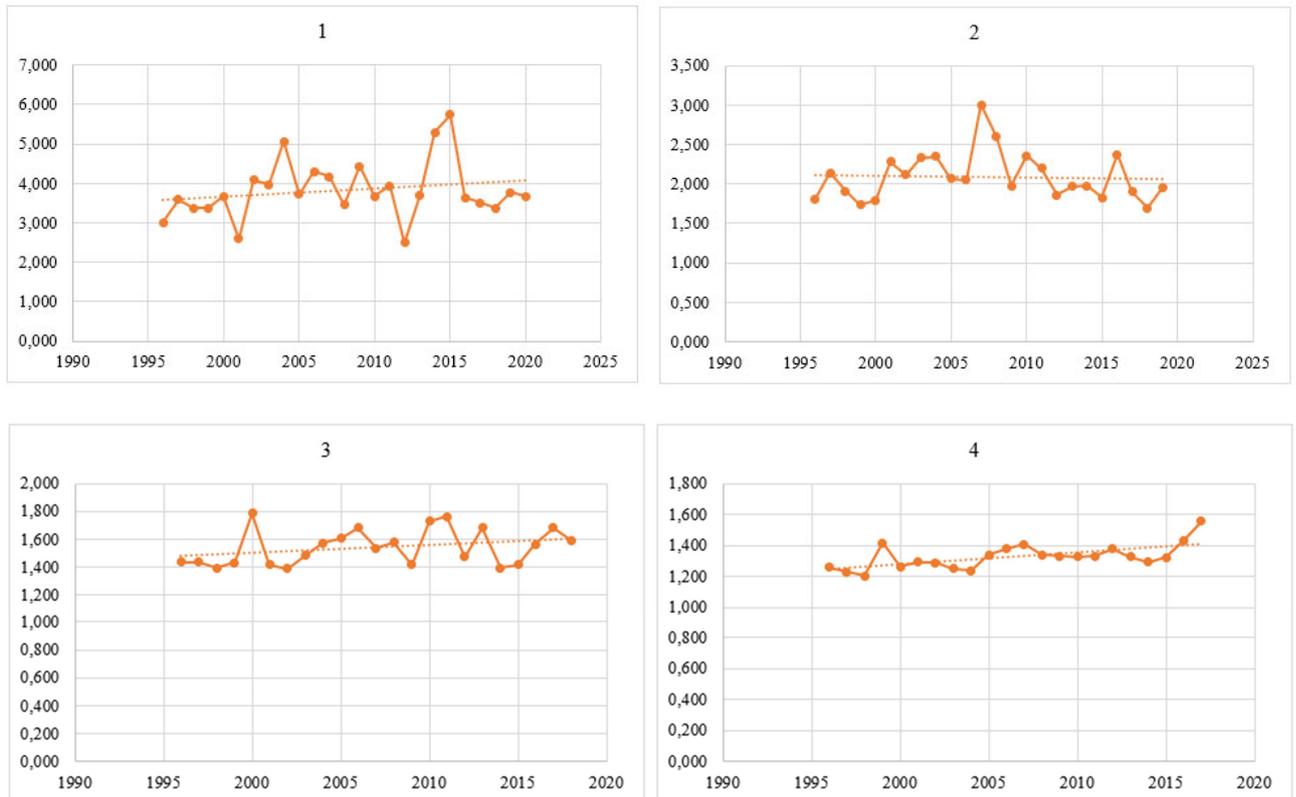
Conclusion : Le test est globalement accepté sur toutes les délais de survénance et le coefficient R^2 est proche de 1, ce qui valide le test de linéarité. Les graphiques du test de linéarité sont joints en annexe du mémoire (cf. [annexe 6](#)).

– Hypothèse d'indépendance :

L'hypothèse d'indépendance entre les années de DOC suppose l'indépendance entre les lignes du triangle de liquidation. Elle joue un rôle important car les triangles de provisionnement peuvent présenter des effets diagonaux (survenance), dont l'apparition peut être liée à une dérive de sinistralité affectant plusieurs années de DOC de la même façon et provoquant ainsi une dépendance entre celles-ci.

Description : La validité de cette hypothèse est testée par un test graphique qui consiste à représenter la jème colonne du triangle (constituée des coefficients de passage) en fonction des années de DOC. L'indépendance des années de DOC signifierait que pour une colonne j du triangle, les coefficients oscillent sans structure aléatoire particulière autour du coefficient de passage moyen.

Résultats : À la lecture des graphiques (cf. [annexe 7](#)), les coefficients oscillent, sans structure aléatoire particulière, autour de la moyenne pour l'ensemble du triangle. Ci-dessous les résultats des 4 premières années.



Conclusion : Ce test visuel permet ainsi de conclure que l'hypothèse d'indépendance des années de DOC est globalement acceptable.

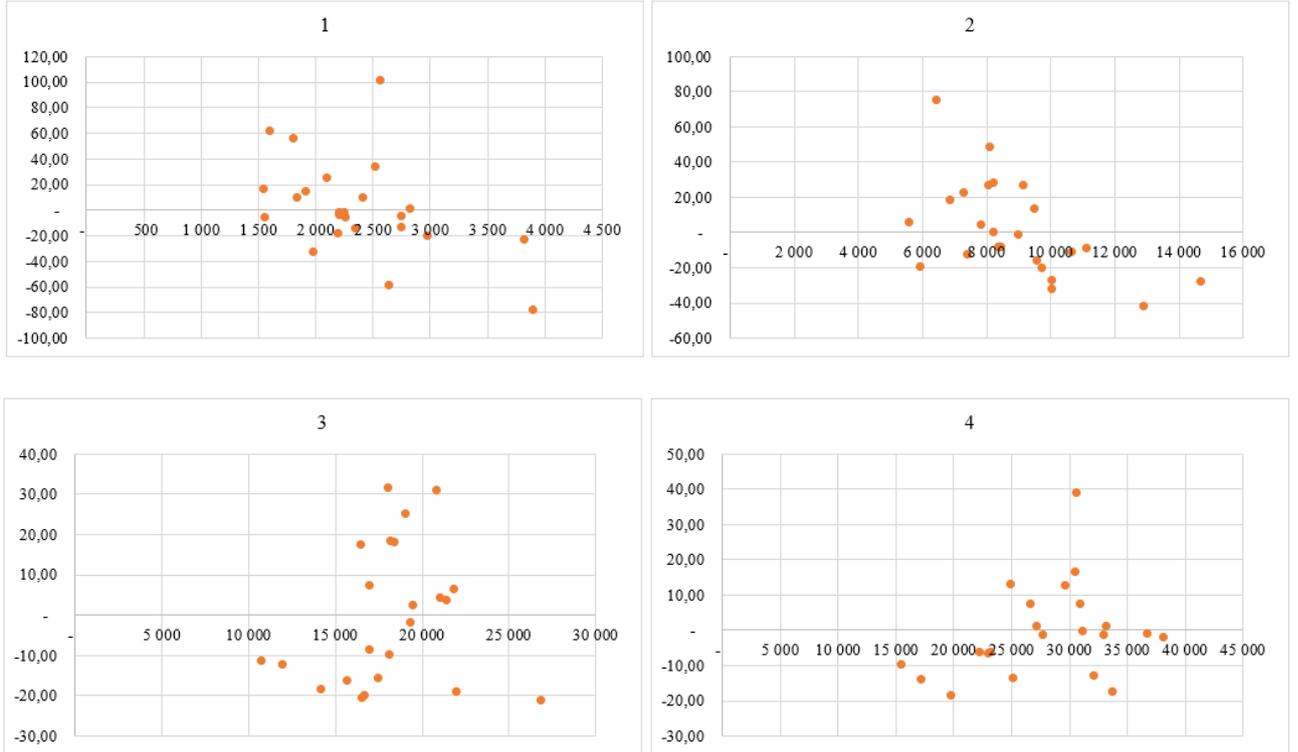
– Hypothèse de variance :

L'hypothèse de variance suppose que $\text{Var}(C_{i,j+1} | C_{i,1}, \dots, C_{i,j})$ est proportionnel à $C_{i,j}$. C'est-à-dire qu'il existe σ_j^2 tel que $\text{Var}(C_{i,j+1} | C_{i,1}, \dots, C_{i,j}) = \sigma_j^2 \times C_{i,j}$. En d'autres termes, on suppose que la volatilité, pour un délai de survenance donné, est identique quelle que soit l'année de DOC.

Description : Un test graphique permet de vérifier l'hypothèse de variance à l'aide d'un graphique représentant les résidus normalisés en fonction des montants cumulés $C_{i,j}$. L'hypothèse est rejetée si les résidus présentent une structure non-aléatoire.

Les résidus sont définis de la façon suivante : $res_{i,j} = \frac{C_{i,j+1} - f_j C_{i,j}}{\sqrt{C_{i,j}}}$

Résultats : L'interprétation des graphique (cf. [annexe 8](#)) des résidus est généralement délicate et subjective. Cela nécessite en règle générale une bonne compréhension préalable des données et des phénomènes de gestion susceptibles d'expliquer des résultats observés. Ci-dessous les résultats des 4 premières années.



Conclusion : Bien que les résidus ne soient pas uniformément répartis de part et d'autre de l'axe des abscisses sur toutes les délais de survénance, les graphiques ne dégagent pas une structure non-aléatoire. Ainsi, on peut conclure que l'hypothèse de variance n'est pas remise en cause.

d) Estimation

Soit $F = \{(CD_{k,l}) ; k + l \leq n + 1\}$ l'ensemble de l'information connue du triangle DOC \times Survénance à la date d'inventaire.

Sous les hypothèses **(I bis)**, **(II bis)**, le résultat suivant est vérifié :

pour tout $k \in \{1, \dots, n\}$,

$$\mathbb{E}(E_{k,n} | F) = \left(\sum_{s=1}^{n-k+1} CD_{k,s} \times A_{k+s-1} \right) \times \lambda_{n-i+1} \times \dots \times \lambda_{n-1}$$

En pratique $\hat{\lambda}_1, \dots, \hat{\lambda}_{n-1}$, estimateurs de $\lambda_1, \dots, \lambda_{n-1}$, sont définis par :

pour tout $l \in \{1, \dots, n - 1\}$,

$$\hat{\lambda}_l = \frac{\sum_{k=1}^{n-l} E_{k,l+1}}{\sum_{k=1}^{n-l} E_{k,l}} = \frac{\sum_{k=1}^{n-l} \sum_{s=1}^{l+1} CD_{k,s} \times \hat{A}_{k+s-1}}{\sum_{k=1}^{n-l} \sum_{s=1}^l CD_{k,s} \times \hat{A}_{k+s-1}}$$

Et $(\sum_{s=1}^{l+1} CD_{k,s} \times A_{k+s-1})$ est estimé par $(\sum_{s=1}^{l+1} CD_{k,s} \times \hat{A}_{k+s-1})$.

Ce qui permet de d'estimer le montant de la charge à l'ultime :

$$\begin{aligned}\hat{E}_{k,n} &= \left(\sum_{s=1}^{n-k+1} CD_{k,s} \times \hat{A}_{k+s-1} \right) \times \hat{\lambda}_{n-k+1} \times \dots \times \hat{\lambda}_{n-1} \\ &= E_{k,n-k+1} \times \hat{\lambda}_{n-k+1} \times \dots \times \hat{\lambda}_{n-1}\end{aligned}$$

Pour obtenir des estimateurs à variance minimale, une troisième hypothèse concernant la variance est ajoutée. Cette hypothèse qui décrit le caractère linéaire de l'évolution de la variance d'une année de survénance à l'année de survénance qui lui succède est formulée ainsi :

(I bis) Il existe $(\varphi_1, \dots, \varphi_{n-1}) \in \mathbb{R}^{n-1}$ facteurs de développement tels que,

$$\text{Var}(E_{k,l+1} \mid CD_{k,1}, \dots, CD_{k,l}) = \varphi_l^2 E_{k,l} \quad (\forall k \in \{1, \dots, n\}, l \in \{1, \dots, n\}).$$

La variance introduite par cette dernière hypothèse est minimale pour l'estimateur suivant :

$$\left\{ \begin{array}{l} \hat{\varphi}_l^2 = \frac{1}{n-l-1} \sum_{k=1}^{n-l} E_{k,l} \times \left(\frac{E_{k,l+1}}{E_{k,l}} - \hat{\lambda}_l \right)^2 \text{ pour tout } l \in \{1, \dots, n-2\} \\ \hat{\varphi}_{n-1}^2 = \min \left(\frac{\hat{\varphi}_{n-2}^4}{\hat{\varphi}_{n-3}^2}, \hat{\varphi}_{n-2}^2, \hat{\varphi}_{n-3}^2 \right) \end{array} \right.$$

e) Calcul de l'erreur de prédiction

Afin de calculer l'erreur de prédiction sur l'estimation de la PSNEM, la MSEP de cette dernière est introduite comme suit.

pour tout $k \in \{1, \dots, n\}$

$$\begin{aligned}MSEP(\widehat{PSNEM}_k) &= \mathbb{E} \left[(\widehat{PSNEM}_k - PSNEM_k)^2 \mid F \right] \\ &= \mathbb{E} \left(\left[(\hat{E}_{k,n} - E_{k,n-k+1}) - (E_{k,n} - E_{k,n-k+1}) \right]^2 \mid F \right) \\ &= \mathbb{E} \left([\hat{E}_{k,n} - E_{k,n}]^2 \mid F \right) \\ &= MSEP(\hat{E}_{k,n}).\end{aligned}$$

Elle peut être réécrite comme suit :

$$MSEP(\hat{E}_{k,n}) = \text{Var}(E_{k,n} \mid F) + (\mathbb{E}(E_{k,n} \mid D) - \hat{E}_{k,n})^2$$

Après développement (cf. [annexe](#)), la formule suivante est vérifiée :

$$\boxed{MSEP(\widehat{PSNEM}_k) = \hat{E}_{k,n}^2 \sum_{l=n-k+1}^{n-1} \frac{\hat{\varphi}_l^2}{\hat{\lambda}_l^2} \left(\frac{1}{\hat{E}_{k,l}} + \frac{1}{\sum_{s=1}^{n-l} C_{s,l}} \right) + \text{Var}_k(E_{k,n-k+1}) \times \hat{B}_k^2.}$$

Avec,

- $\hat{B}_k^2 = \hat{\lambda}_{n-k+1} \times \dots \times \hat{\lambda}_{n-1} = \prod_{i=n-k+1}^{n-1} \hat{\lambda}_i$;
- $\text{Var}_k(E_{k,n-k+1}) = \text{Var}(E_{k,n-k+1} \mid CD_{k,1}, \dots, CD_{k,n-k+1})$

Il est à noter que cette formule tient compte d'un nouveau terme de variance, en sus de la MSEP issue du modèle classique de Mack.

Dans le cas d'une projection Chain Ladder classique, les données situées sur la dernière diagonale étant déterministes, ce terme est nul. Dans le cas spécifique de l'assurance construction, le triangle étudié est le triangle DOC \times Survenance contenant une ventilation des IBNR estimées dans le triangle Survenance \times Développement. Concrètement, on rajoute à l'erreur de la projection du triangle DOC \times Survenance une erreur des IBNR retrouvée sur la diagonale, qu'on appellera erreur « 3D ».

Le \hat{B}_k^2 étant connu, on cherche à estimer l'entité $\text{Var}_k(E_{k,n-k+1})$. Après développement (cf. annexe), on obtient :

$$\widehat{\text{Var}}_k(E_{k,n-k+1}) = \sum_{s=1}^{n-k+1} CD_{k,s}^2 \times \frac{\widehat{\text{Var}}(\hat{C}_{k+s-1,n})}{C_{k+s-1,n-k-s+2}^2}$$

Avec,

$$\widehat{\text{Var}}(\hat{C}_{k,n}) = \hat{C}_{k,n}^2 \sum_{j=n-k+1}^{n-1} \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{\hat{C}_{i,j}}$$

Ainsi,

$$\text{erreur « 3D »}^2 = \hat{B}_k^2 \times \sum_{s=1}^{n-k+1} CD_{k,s}^2 \times \frac{\hat{C}_{k+s-1,n} \sum_{j=n-k+s}^{n-1} \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{\hat{C}_{i,j}}}{C_{k+s-1,n-k-s+2}^2}$$

Et,

$$\widehat{MSEP}(PSNEM_k) = \hat{E}_{k,n}^2 \sum_{l=n-k+1}^{n-1} \frac{\hat{\varphi}_l^2}{\hat{\lambda}_l^2} \left(\frac{1}{\hat{E}_{k,l}} + \frac{1}{\sum_{s=1}^{n-l} C_{s,l}} \right) + \text{erreur « 3D »}^2$$

L'application du modèle de Mack 3D permet ainsi d'obtenir une estimation de l'erreur de prédiction pour l'estimation combinée des IBNR et de la PSNEM à partir d'une formule fermée.

2.4 Synthèse

L'analyse initiale des données est une étape clé dans tout projet actuariel, permettant de comprendre la qualité et la structure des données. Les données utilisées pour cette étude correspondent à des données réelles d'un portefeuille construction d'une mutuelle sur le marché français. Pour des considérations de confidentialité, les données ont été anonymisées.

Plusieurs contrôles ont été réalisés afin de fiabiliser les données. Des statistiques descriptives ont permis de mettre en évidence la cohérence de ces dernières et confirment que le portefeuille étudié présente des caractéristiques observées sur le marché de l'assurance construction en France.

La gestion en capitalisation des garanties DO et RCD et le caractère décennal de la couverture obligent un suivi spécifique des sinistres. Les sinistres étant rattachés à la DOC (exercice de souscription), il faut considérer une nouvelle dimension et donc une représentation en 3D de type « pyramide » : Survenance \times Développement \times DOC. Pour une DOC donnée, on dispose comme pour les branches « classiques » des données de survenance par développement.

Afin d'appliquer les modèles développés dans ce mémoire, les hypothèses nécessaires ont été testées, notamment les hypothèses de linéarité, d'indépendance et de variance. Travaillant sur un environnement 3D, ces dernières ont été testées pour une vision Survenance \times Développement (vision PSAP) et pour une vision DOC \times Survenance (vision PSNEM). Les différents tests d'hypothèses sont acceptés.

Les méthodes traditionnelles de provisionnement, qui ont été utilisées pendant de nombreuses années pour évaluer les provisions pour sinistres, ne sont pas assez flexibles pour s'adapter aux complexités de l'assurance construction.

Le modèle de Mack classique, qui se concentre sur les provisions pour sinistres en deux dimensions, est adapté dans ce travail pour tenir compte des spécificités de l'assurance construction. Cette adaptation à trois dimensions permet de mieux modéliser les risques associés.

Le calcul de la variance, généralement effectué à l'aide de méthodes stochastiques classiques, doit être conditionné aux estimations des sinistres survenus mais non déclarés (IBNR) pour tenir compte du biais introduit par leur ventilation. C'est pourquoi le modèle de Mack 3D proposé dans ce mémoire est une généralisation d'un modèle de Mack conditionnel à l'estimation des IBNR. Le résultat de cette généralisation est présenté ci-après.

$$\widehat{MSEP}(PSNEM_k) = \hat{E}_{k,n}^2 \sum_{l=n-k+1}^{n-1} \frac{\hat{\phi}_l^2}{\hat{\lambda}_l^2} \left(\frac{1}{\hat{E}_{k,l}} + \frac{1}{\sum_{s=1}^{n-l} C_{s,l}} \right) + \text{erreur « 3D »}^2.$$

Dans le cadre de notre étude, nous introduisons également le “Modèle GLM : modèle multiplicatif Poisson ODP” utilisé comme point de comparaison pour évaluer les performances du modèle de Mack 3D. Cette comparaison nous permettra de mieux comprendre les avantages et les limites de chaque approche. Le résultat du modèle est présenté ci-après.

$$\widehat{MSEP}(\hat{R}) = \hat{\phi} \sum_{i+(j-1)+(k-1) > N} \hat{C}_{ijk} + g^{-1}(M \cdot \hat{\beta}) \cdot M \cdot I^{-1} \cdot M' \cdot \left(g^{-1}(M \cdot \hat{\beta}) \right)'$$

Chapitre III : Mise en pratique du modèle de Mack 3D

Ce chapitre est consacré à la mise en œuvre du modèle de Mack 3D dans le contexte réel.

Nous présenterons en premier lieu les résultats de l'estimation de la PSNEM économique via le modèle de Mack. Ensuite, nous détaillerons les résultats du calcul des erreurs de prédiction à la fois en 2D et en 3D. Ceci nous permettra d'illustrer l'impact de la généralisation du modèle.

Enfin, nous comparerons les résultats obtenus avec les résultats d'un GLM 3D via différentes hypothèses de distribution.

3.1 Calcul de la PSNEM

Les hypothèses étant validées, nous appliquons le modèle Mack 3D au portefeuille étudié. La projection des triangles de charges, nous mène à une estimation des IBNR et de la PSNEM comme suit.

Provisions			
DOC	IBNR	PSNEM	IBNR+PSNEM
1996	20 028	-	20 028
1997	23 811	-	23 811
1998	52 035	-	52 035
1999	122 375	14 883	137 257
2000	309 081	21 854	330 935
2001	563 889	20 614	584 503
2002	955 634	39 256	994 890
2003	1 303 118	55 723	1 358 841
2004	2 444 098	79 240	2 523 338
2005	3 282 726	174 969	3 457 695
2006	5 758 597	277 628	6 036 225
2007	7 541 853	365 175	7 907 027
2008	8 578 590	658 223	9 236 813
2009	10 598 293	1 184 011	11 782 304
2010	18 106 068	4 712 439	22 818 506
2011	17 943 832	12 930 290	30 874 122
2012	14 645 716	17 993 483	32 639 199
2013	13 963 263	26 269 831	40 233 094
2014	12 153 293	30 605 536	42 758 829
2015	16 395 554	55 561 053	71 956 607
2016	16 170 682	63 986 256	80 156 938
2017	15 192 132	76 434 391	91 626 523
2018	11 459 488	85 945 381	97 404 869
2019	7 695 800	90 910 122	98 605 922
2020	3 358 302	82 253 310	85 611 612
2021	1 775 426	161 649 794	163 425 220
Total	190 413 682	712 143 461	902 557 144

Tableau 3.1-1 Résultats des provisions pas année de DOC

Les IBNR ont été estimés par année de survenance et ventilés ultérieurement par année de DOC. Afin d'appliquer la méthode de Mack 3D développée dans ce mémoire, aucun ajustement n'a été réalisé pour l'estimation des IBNR et de la PSNEM.

3.2 Résultat 2D

Afin d'estimer l'impact de la généralisation de Mack en 3D, nous estimons l'erreur de prédiction via la méthode traditionnelle. Les résultats sont présentés ci-dessous.

DOC	Provisions	Erreur de processus	Erreur d'estimation	RMSE (2D)	CV (2D)
1996	20 028	-	-	-	0%
1997	23 811	-	-	-	0%
1998	52 035	-	-	-	0%
1999	137 257	22 005	16 659	27 600	20%
2000	330 935	23 790	17 991	29 827	9%
2001	584 503	22 196	15 373	27 000	5%
2002	994 890	34 543	20 436	40 136	4%
2003	1 358 841	41 603	23 819	47 939	4%
2004	2 523 338	48 237	30 191	56 906	2%
2005	3 457 695	127 071	55 191	138 539	4%
2006	6 036 225	145 649	67 155	160 385	3%
2007	7 907 027	161 962	73 255	177 758	2%
2008	9 236 813	316 627	114 648	336 745	4%
2009	11 782 304	446 993	142 806	469 251	4%
2010	22 818 506	1 840 253	603 945	1 936 823	8%
2011	30 874 122	3 370 787	1 074 300	3 537 842	11%
2012	32 639 199	3 154 324	879 912	3 274 753	10%
2013	40 233 094	3 549 104	964 765	3 677 895	9%
2014	42 758 829	3 863 959	957 218	3 980 759	9%
2015	71 956 607	5 453 784	1 594 909	5 682 208	8%
2016	80 156 938	6 477 860	1 786 885	6 719 794	8%
2017	91 626 523	7 740 344	2 092 404	8 018 171	9%
2018	97 404 869	9 963 420	2 544 215	10 283 131	11%
2019	98 605 922	12 944 776	3 051 295	13 299 535	13%
2020	85 611 612	17 656 158	3 586 021	18 016 644	21%
2021	163 425 220	34 967 548	9 341 890	36 193 927	22%
Total	902 557 144	44 559 184	20 887 966	49 212 072	5%

Tableau 3.2-1 Résultats du modèle Mack 2D

En appliquant un modèle de Mack classique sur les données, nous obtenons une estimation de l'erreur quadratique (RMSE) de 50 m€, soit 5% des provisions. Cette estimation est le résultat du modèle de Mack sur le triangle de PSNEM sans tenir compte de la volatilité induite par l'estimation des IBNR ventilé dans le triangle.

Comme expliqué précédemment, cette approche est critiquable du fait que l'application directe du modèle de Mack suppose que le triangle étudié est déterministe, ce qui n'est pas le cas. Ainsi, le modèle sous-estime fortement la volatilité des provisions en ne tenant pas compte de la volatilité induit par l'estimation des IBNR et leur ventilation.

3.3 Résultat 3D

Nous présentons ci-après les résultats obtenus grâce à l'application du modèle de Mack 3D, qui a été développé tout au long de ce mémoire. Ce modèle, adapté aux spécificités de l'assurance construction, a permis d'obtenir des estimations plus précises de la volatilité de l'estimation des provisions des garanties construction. Cette présentation des résultats du modèle de Mack 3D vise à démontrer l'impact de l'intégration de l'erreur de prédiction lié à la ventilation des IBNR dans l'estimation global de l'erreur.

DOC	Provisions	RMSE (2D)	Erreur 3D	RMSE (3D)	CV (3D)
1996	20 028	-	9 845 874	9 845 874	49161%
1997	23 811	-	11 249 721	11 249 721	47245%
1998	52 035	-	13 394 562	13 394 562	25742%
1999	137 257	27 600	19 739 512	19 739 519	14381%
2000	330 935	29 827	20 460 783	20 460 791	6183%
2001	584 503	27 000	17 276 178	17 276 184	2956%
2002	994 890	40 136	18 761 712	18 761 723	1886%
2003	1 358 841	47 939	20 127 812	20 127 826	1481%
2004	2 523 338	56 906	24 746 178	24 746 197	981%
2005	3 457 695	138 539	23 695 876	23 695 941	685%
2006	6 036 225	160 385	26 926 146	26 926 230	446%
2007	7 907 027	177 758	27 054 212	27 054 311	342%
2008	9 236 813	336 745	24 326 151	24 326 422	263%
2009	11 782 304	469 251	20 613 452	20 613 947	175%
2010	22 818 506	1 936 823	24 729 583	24 736 957	108%
2011	30 874 122	3 537 842	25 218 670	25 241 542	82%
2012	32 639 199	3 274 753	18 931 950	18 952 387	58%
2013	40 233 094	3 677 895	18 336 315	18 361 678	46%
2014	42 758 829	3 980 759	15 529 426	15 558 899	36%
2015	71 956 607	5 682 208	22 280 647	22 337 658	31%
2016	80 156 938	6 719 794	19 719 374	19 800 168	25%
2017	91 626 523	8 018 171	19 122 558	19 236 693	21%
2018	97 404 869	10 283 131	19 417 561	19 583 531	20%
2019	98 605 922	13 299 535	19 172 360	19 413 650	20%
2020	85 611 612	18 016 644	20 510 159	20 821 292	24%
2021	163 425 220	36 193 927	39 260 674	40 356 801	25%
Total	902 557 144	49 212 072	109 796 492	111 765 723	12%

Tableau 3.3-1 Résultats du modèle Mack 3D

L'application du modèle de Mack introduit un terme supplémentaire dans le calcul de la RMSE. Ce terme, appelé erreur 3D, permet de capter la volatilité de l'estimation des IBNR, leur ventilation et leur projection. Ainsi, nous obtenons, une RMSE de 112 m€, soit 12% des provisions.

Bien que les années anciennes présentent des montants d'IBNR non nuls, l'estimation de l'erreur de Mack 2D sur celles-ci est nulle. Ce phénomène est corrigé par le modèle Mack 3D.

On note également que l'erreur 3D est plus importante sur les années récentes. Ceci est expliqué du fait que la projection est réalisée sur une assiette de charge y compris IBNR. Ainsi, une erreur d'estimation de cette dernière est répercutée à horizon de projection (caractère multiplicatif de la méthode Chain Ladder).

3.4 Résultat du modèle GLM

Afin de comparer les résultats obtenus par le modèle Mack 3D développé dans ce mémoire, nous mettons en place un modèle GLM 3D décrit au paragraphe [2.3.2](#). Usuellement en provisionnement Non-Vie, le choix de la loi a priori se porte sur une loi de Poisson surdispersée (ODP). Nous avons cependant sélectionné trois lois candidates pour comparaison : Poisson ODP, Tweedie et Binomiale Négative.

Les résultats des différents modèles sont donnés dans le tableau ci-dessous.

Loi	Provisions	RMSE	CV
Poisson ODP	1 042 002 712	94 005 919	9%
Binomiale Négative	745 462 783	361 381 624	48%
Tweedie	899 170 234	147 806 550	16%

Tableau 3.4-1 Résultats GLM 3D

On note que le cas Poisson ODP du modèle GLM 3D estime un montant de provision supérieur aux provisions estimées par le modèles Mack (1 042 m€ vs 903 m€) tandis que le cas Tweedie sous-estime légèrement cette dernière (899 m€ vs 903 m€). Le cas de la loi Binomiale Négative semble nettement sous-estimer le montant des provisions.

L'erreur quadratique estimée par le cas Poisson ODP est légèrement inférieur à celle estimée par le modèle Mack 3D tandis que le cas Tweedie estime une RMSE légèrement supérieure à cette dernière. A l'image de l'estimation des provisions, on note que le cas de la loi Binomiale Négative semble éloigné des autres résultats.

Dans le but de juger de la qualité des estimations des différents modèles GLM 3D, on définit les indicateurs statistiques suivants.

- RMSE du modèle : Racine carrée de la moyenne des résidus au carré. Cet indicateur correspond à l'erreur moyenne de la prédiction, ligne par ligne. Attention à ne pas confondre la RMSE de la provision calculée précédemment avec la RMSE du modèle GLM ;
- R^2 : Coefficient de détermination défini comme étant la proportion de variance expliquée par le modèle. Quand $R^2 = 1$ le modèle prédit parfaitement les données ;
- Top-Décile-Lift : métrique qui exprime l'incidence dans les 10% des données avec les prédictions les plus élevées du modèle se comparant à l'incidence globale de l'échantillon. Un Top-Décile-Lift égal à 1 indique un modèle aléatoire. Un Top-Décile-Lift égale à 3 indique que dans les 10% de prédictions les plus élevées, 3 fois plus de cas positifs sont identifiés par le modèle que ce qui serait attendu pour une sélection aléatoire d'instances ;
- Log Loss : Fonction de perte logistique ;
- Quotient : Le quotient est défini comme le rapport de la provision prédite sur la provision déterministe de la base. Idéalement le quotient doit être supérieur 1.

Le résultat du calcul des différents indicateurs est présenté ci-dessous.

Loi	RMSE du modèle (€)	R ²	TopDecile Lift	Log Loss	Quotient
Poisson ODP	415 856	89%	6,293	4,026	1,000
Binomiale Négative	812 692	57%	6,318	4,026	1,245
Tweedie	482 993	85%	6,338	4,026	1,011

Tableau 3.4-2 Indicateurs des modèles GLM 3D

Les modèles avec les lois Poisson ODP et Tweedie présente à la fois une RMSE acceptable et un R² proche de 1. Leurs TopDecileLift, Log Loss et Quotient indiquent à la fois une bonne performance des modèles mais également une comparabilité des deux résultats. Se basant principalement sur la plus petite RMSE et le R² le plus proche de 1, le cas du Poisson ODP semble être le modèle le plus performant.

On note par ailleurs, que la performance du modèle Binomiale Négative est faible. L'erreur quadratique de ce dernier est nettement supérieure aux erreurs des autres modèles et son R² est très faible et inférieur à 60%. Ces constatation nous pousse à ne pas retenir ce dernier et l'exclure des comparaisons qui suivent.

3.5 Comparaison des résultats

Ce paragraphe est dédié à la comparaison des différents résultats obtenus avec les différents modèles.

Loi/Modèle	Provisions	RMSE	CV
Mack 3D	902 557 144	111 765 723	12%
Poisson ODP	1 042 002 712	94 005 919	9%
Tweedie	899 170 234	147 806 550	16%

Tableau 3.5-1 synthèses des résultats

On note que l'estimation des provisions est comparable entre les trois modèles retenus avec une légère surestimation dans le cas du modèle GLM 3D Poisson ODP.

On constate que les erreurs quadratiques calculées à la fois par le modèle Mack 3D et les modèles GLM 3D sélectionnés sont comparables et nettement supérieurs à l'erreur quadratique calculée par le modèle Mack 2D (50 m€). Ce constat permet de conforter le fait que l'application d'un modèle de Mack classique en assurance construction est erronée et de confirmer que le modèle de Mack 3D corrige cette erreur.

On note également qu'en pourcentage des provisions, l'erreur quadrique calculée par le modèle de Mack 3D est encadrée entre les deux erreurs calculées par les modèles GLM 3D sélectionnés. Ceci permet de situer un ordre de grandeur de la volatilité qu'on souhaite capter lors de l'estimation des provisions d'une part et d'autre part de confirmer la bonne estimation de l'erreur quadratique par le modèle de Mack 3D.

3.6 Synthèse

Les résultats du modèle de Mack 3D, permet de confirmer une sous-estimation de la volatilité calculée par un modèle Mack classique, qui en négligeant la volatilité induite par l'estimation des IBNR, introduit une inexactitude dans l'estimation de l'erreur quadratique des provisions.

L'estimation de l'erreur quadratique des provisions par des modèles GLM 3D permet d'encadrer l'estimation de l'erreur quadratique calculée par le modèle de Mack 3D. Ces résultats permettent, d'une part de situer un ordre de grandeur de la volatilité qu'on souhaite capter lors de l'estimation des provisions et d'autre part de confirmer la bonne estimation de l'erreur quadratique par le modèle de Mack 3D.

Conclusion

Dans un contexte économique de relance du secteur du BTP, une évaluation précise de la rentabilité de l'assurance construction est plus que jamais nécessaire. Du fait de la représentation tridimensionnelle des données de sinistralité de cette branche, cette exercice semble particulièrement complexe. L'avènement de la nouvelle norme IFRS 17, laquelle introduit la détermination et la comptabilisation d'un ajustement pour risque non financier (ou Risk Adjustment) motive d'autant plus les développements des modèles de provisionnement stochastiques adaptés aux spécificités de la branche.

Après un recensement des méthodes de provisionnement en assurance Non-Vie, une modélisation d'une nouvelle méthode de provisionnement stochastique a été réalisée. Prenant en compte les spécificités de l'assurance construction et ses trois axes de temps, cette dernière permet de calculer l'incertitude de prédiction des provisions par formule fermée, permettant ainsi une extension du modèle classique de Mack à des données « 3D ».

Cette extension de la formule de Mack a été comparée avec d'autres approches telles que la famille des GLM 3D dans le cas des données de RC décennale d'un assureur français. La qualité d'ajustement de ces modèles ainsi que leurs performance sont jugées satisfaisants et permettent un comparaison des résultats avec le modèle Mack 3D.

Au vu des résultats obtenus lors de l'application numérique, l'application du modèle de Mack 3D introduit un terme supplémentaire dans le calcul de l'erreur de prédiction. Ce terme qu'on appelle « erreur 3D », permet de capter la volatilité de l'estimation des IBNR, leur ventilation et leur projection. Dans le cadre d'un modèle de Mack classique cet effet est négligé en supposant que le triangle étudié pour l'estimation des provisions est déterministe, ce qui n'est le cas. Les résultats montrent que le modèle de Mack 3D permet de compenser une nette sous-estimation de l'erreur de prédication en comparaison à un modèle de Mack classique non adapté aux spécificités de l'assurance construction.

Les résultats de l'application numérique des modèles de GLM de 3D permettent d'encadrer l'estimation de l'erreur de prédiction calculée par le modèle de Mack 3D. Ces estimations permettent, d'une part de situer un ordre de grandeur de la volatilité qu'on souhaite capter lors de l'estimation des provisions et d'autre part de confirmer la bonne estimation de l'erreur quadratique par le modèle de Mack 3D.

Cette comparaison des approches stochastiques dans le cas très spécifique et complexe de l'assurance construction demanderait cependant à être poursuivie via l'analyse d'autres jeux de données, afin de confirmer ou d'infirmer les premiers résultats obtenus dans ce mémoire.

Par ailleurs, une piste possible d'amélioration consisterait à pouvoir inclure un jugement d'expert dans la méthode de Mack 3D développé dans ce mémoire via la prise en compte de pondérations appliquées aux coefficients de développement individuels à l'instar de la généralisation proposée dans Mack (1999). En ce qui concerne les modèles GLM 3D, d'autres lois pourraient être également considérées, notamment des lois dites "zero-inflated".

Il est à noter que l'incertitude comprise dans les triangles couvre également divers risques notamment juridiques, politiques, d'inflation, etc. Dans le cadre de ce mémoire, ces risques, parfois exogènes aux triangles, ne sont pas pris en compte, notamment dû à la date des triangles considérés (arrêtés au 31.12.2021). L'intégration de ces risques pourrait faire l'objet d'un nouveau sujet de mémoire sur lequel, le modèle proposé pourrait être étendu afin d'intégrer ces derniers.

Bibliographie

1. Ouvrage

[1] *McCullagh, P., Nelder, J. A. (1989). Generalized linear models (Vol. 37). CRC press.*

2. Articles

[2] *Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. ASTIN Bulletin : The Journal of the IAA, 23(2), 213-225.*

[3] *Mack, T. (1994). Measuring the variability of chain ladder reserve estimates. In Casualty Actuarial Society Forum (Vol. 1, No. 1, pp. 101-83). Edward W. Frees, Richard A. Derrig and Glenn Meyers Excerpt More information.*

[4] *Mack, T. (1999). The standard error of chain ladder reserve estimates : Recursive calculation and inclusion of a tail factor. ASTIN Bulletin : The Journal of the IAA, 29(2), 361-366.*

[5] *Murphy, D. M. (2007). Chain ladder reserve risk estimators. In CAS E-Forum Summer.*

[6] *Tee, L. (2013). Stochastic reserving methods in non-life insurance (Doctoral dissertation, Tartu Ulikool)."*

[7] *England, Peter D., and Richard J. Verrall (2002). "Stochastic claims reserving in general insurance." British Actuarial Journal 8.3*

3. Mémoires et Rapports

[8] *Ridha Boulahia, Analyse des provisions techniques dans le cadre des travaux de la Fonction Actuarielle d'un assureur non-vie.*

[9] *DEPLANTE Brice, Provisionnement et Tarification en Dommages Ouvrage.*

[10] *Clément BOURRY, Evaluation des provisions techniques et du capital économique associé au risque de réserve en assurance construction.*

[11] *Ruben Nataf, les méthodes de provisionnement en assurance construction.*

[12] *Thomas CHAPUIS, Provisionnement en assurance construction MLG et mesure de l'aléa.*

[13] *Thomas HAEGEL, Solvabilité II - Impact de la nouvelle directive sur une société d'assurance non-vie spécialisée dans l'assurance construction : Application et calibrage de la formule standard.*

[14] *ACPR, Quelques statistiques concernant le marché français de l'assurance construction (garanties dommages-ouvrage et RC décennale). n86 – Février 2018.*

- [15] *ACPR, Les chiffres du marché français de la banque et de l'assurance 2016.*
- [16] *FFA, LES ASSURANCES DE BIENS ET DE RESPONSABILITE. Données clés 2020.*
- [17] *FFA, RAPPORT ANNUEL 2020.*

Annexes

Annexe 1 : Développement du calcul de la MSEP

On reprend ici les notations du paragraphe [2.3.3](#).

▪ Résultat 1 :

Soit $F = \{(CD_{k,l}) ; k + l \leq n + 1\}$ l'ensemble de l'information connue du triangle DOC \times Survenance à la date d'inventaire.

Sous les hypothèses **(I bis)**, **(II bis)**, le résultat suivant est vérifié :

$$\text{pour tout } k \in \{1, \dots, n\}, \mathbb{E}(E_{k,n} | F) = \left(\sum_{s=1}^{n-k+1} CD_{k,s} \times A_{k+s-1} \right) \times \lambda_{n-i+1} \times \dots \times \lambda_{n-1}$$

Preuve :

Définition 1 :

Soit $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ et T une variable aléatoire définie sur (Ω, \mathcal{A}) à valeur dans (F, \mathcal{F}) . L'espérance conditionnelle de X sachant T , noté $\mathbb{E}(X | T)$ est définie par :

$$\mathbb{E}(X | T) = \mathbb{E}(X | \sigma(T)),$$

où $\sigma(T)$ est la sou-tribu de \mathcal{A} engendrée par T .

Définition 2 :

On se donne une famille $(Y_\gamma)_{\gamma \in \mathcal{C}}$ de variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ indexée par un ensemble \mathcal{C} quelconque. La tribu engendrée par la famille $(Y_\gamma)_{\gamma \in \mathcal{C}}$, notée $\Gamma = \sigma(Y_\gamma ; \gamma \in \mathcal{C})$ est définie comme la plus petite tribu sur Ω rendant Y_γ mesurable pour tout $\gamma \in \mathcal{C}$. Autrement dit,

$$\Gamma = \sigma\left(Y_\gamma^{-1}(B) ; \gamma \in \mathcal{C} ; B \in \mathcal{B}(\mathbb{R}^d)\right).$$

Propriété (Espérances emboîtées) :

Soit $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ et \mathcal{B} une sous-tribu de \mathcal{A} .

Si \mathcal{B}' est une sous-tribu de \mathcal{B} , on a

$$\mathbb{E}(\mathbb{E}(X | \mathcal{B}) | \mathcal{B}') = \mathbb{E}(X | \mathcal{B}') = \mathbb{E}(\mathbb{E}(X | \mathcal{B}') | \mathcal{B}).$$

Notation :

Pour k entier, on note,

$$\mathbb{E}_k(X) = \mathbb{E}(X | E_{k,1}, \dots, E_{k,n-k+1}).$$

D'après **(II bis)**, on a

$$\mathbb{E}(E_{k,n} | F) = \mathbb{E}(E_{k,n} | E_{k,1}, \dots, E_{k,n-k+1}) = \mathbb{E}_k(E_{k,n}).$$

Ainsi, en appliquant **(I bis)**, on a

$$\begin{aligned} \mathbb{E}(E_{k,n} | F) &= \mathbb{E}_k(E_{k,n}) \\ &= \mathbb{E}_k\left(\mathbb{E}(E_{k,n} | E_{k,1}, \dots, E_{k,n-1})\right) \\ &= \mathbb{E}_k(E_{k,n-1} \times \lambda_{n-1}) \\ &= \mathbb{E}_k(E_{k,n-1}) \times \lambda_{n-1} \\ &= \dots \\ &= \mathbb{E}_k(E_{k,n-k+1}) \times \lambda_{n-k+1} \times \dots \times \lambda_{n-1} \\ &= E_{k,n-k+1} \times \lambda_{n-k+1} \times \dots \times \lambda_{n-1} \end{aligned}$$

□

▪ Résultat 2 :

En reprenant les notation ci-dessus, on a

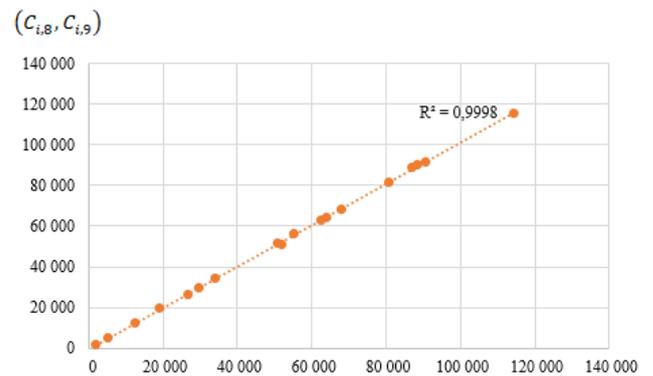
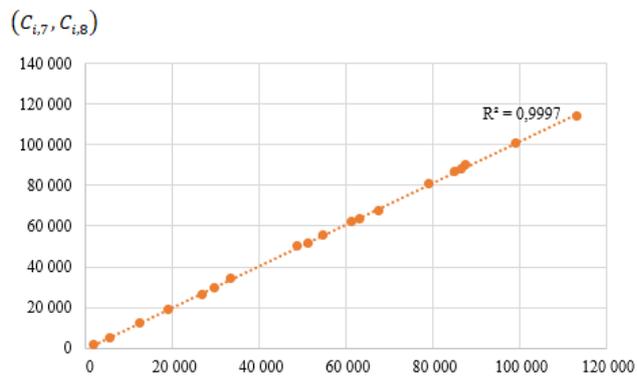
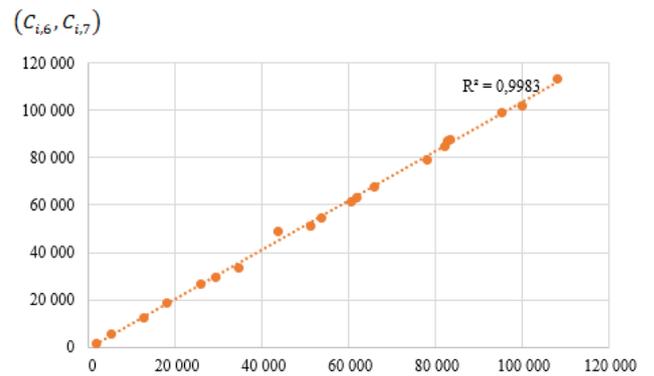
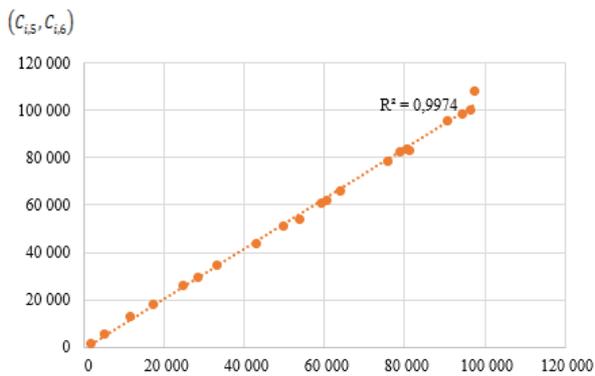
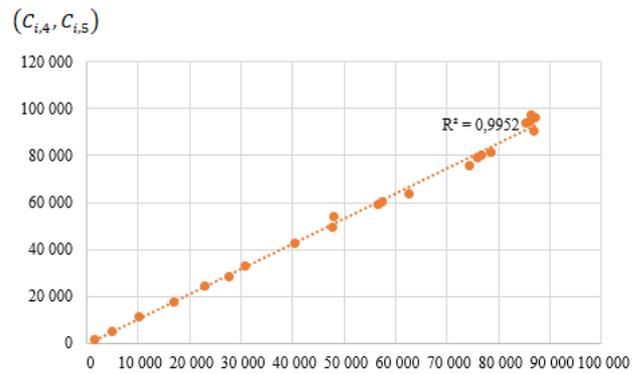
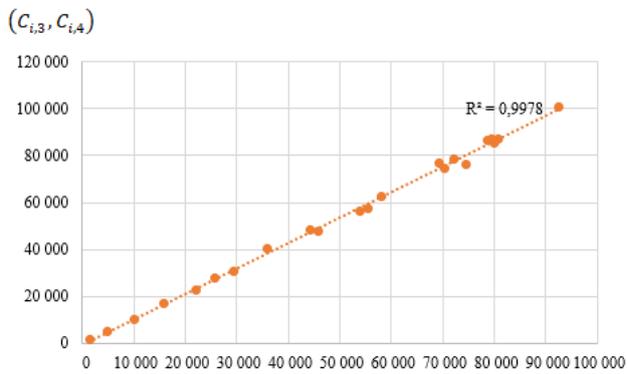
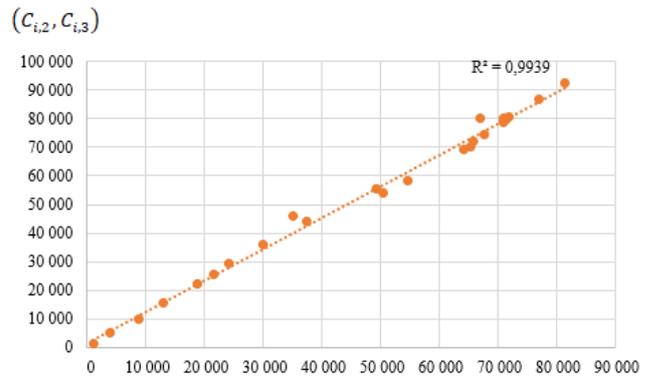
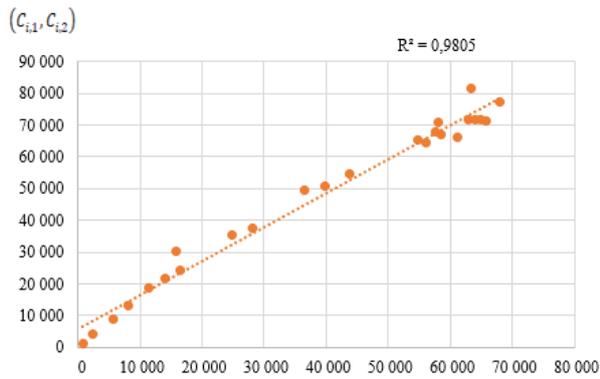
$$\text{Var}_k(E_{k,n-k+1}) = \sum_{s=1}^{n-k+1} CD_{k,s}^2 \times \frac{\text{Var}(\hat{C}_{k+s-1,n})}{C_{k+s-1,n-k-s+2}^2}$$

Preuve :

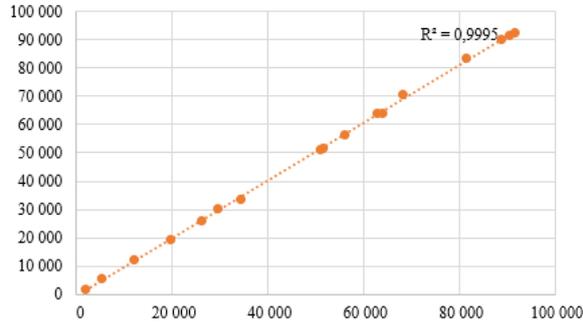
$$\begin{aligned} \text{Var}_k(E_{k,n-k+1}) &= \text{Var}_k\left(\sum_{s=1}^{n-k+1} CD_{k,s} \times \hat{A}_{k+s-1}\right) \\ &= \sum_{s=1}^{n-k+1} \text{Var}_k(CD_{k,s} \times \hat{A}_{k+s-1}) + 2 \sum_{1 \leq s < p \leq n-k+1} \text{Cov}(CD_{k,s} \times \hat{A}_{k+s-1}, CD_{k,p} \times \hat{A}_{k+p-1}) \\ &= \sum_{s=1}^{n-k+1} CD_{k,s}^2 \times \text{Var}_k(\hat{A}_{k+s-1}) + 2 \sum_{1 \leq s < p \leq n-k+1} CD_{k,s}^2 \times CD_{k,p}^2 \times \text{Cov}(\hat{A}_{k+s-1}, \hat{A}_{k+p-1}) \\ &= \sum_{s=1}^{n-k+1} CD_{k,s}^2 \times \frac{\text{Var}_k(\hat{C}_{k+s-1,n})}{C_{k+s-1,n-k-s+2}^2} \\ &\quad + 2 \sum_{1 \leq s < p \leq n-k+1} CD_{k,s}^2 \times CD_{k,p}^2 \times \frac{\text{Cov}(\hat{C}_{k+s-1,n}, \hat{C}_{k+p-1,n})}{C_{k+s-1,n-k-s+2} \times C_{k+p-1,n-k-p+2}} \\ &= \sum_{s=1}^{n-k+1} CD_{k,s}^2 \times \frac{\text{Var}(\hat{C}_{k+s-1,n})}{C_{k+s-1,n-k-s+2}^2} \end{aligned}$$

□

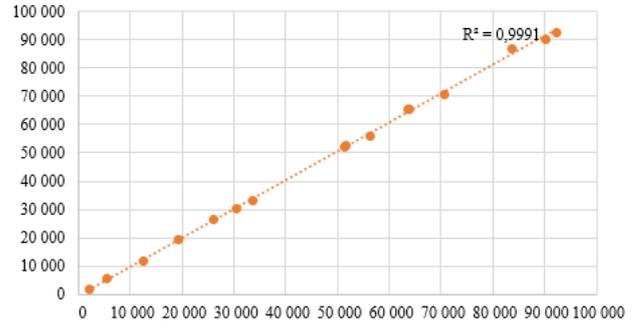
Annexe 3 : Graphiques test de linéarité vision Survenance × Développement



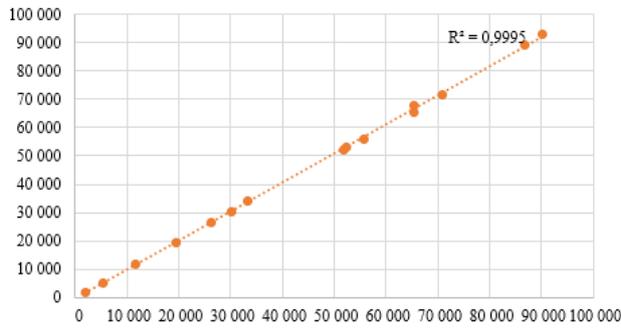
$(C_{i,9}, C_{i,10})$



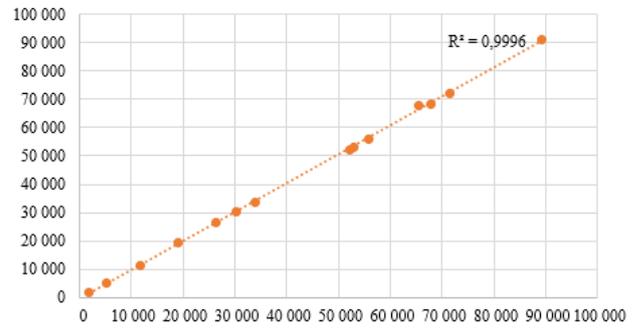
$(C_{i,10}, C_{i,11})$



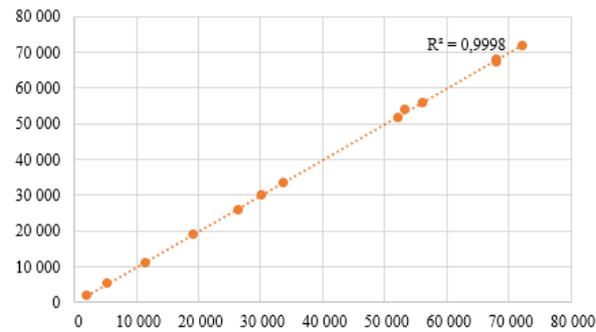
$(C_{i,11}, C_{i,12})$



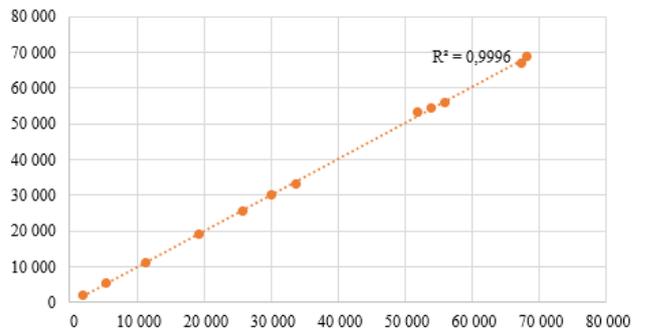
$(C_{i,12}, C_{i,13})$



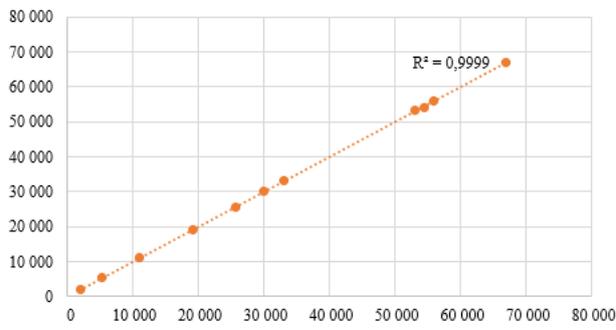
$(C_{i,13}, C_{i,14})$



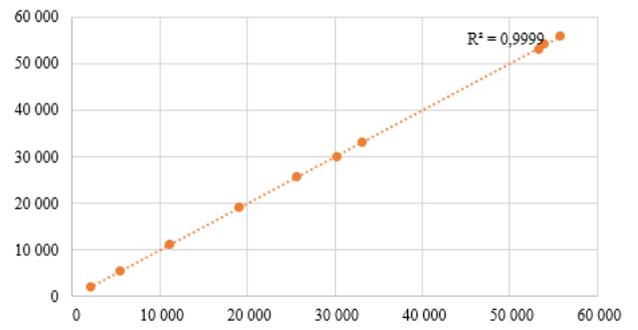
$(C_{i,14}, C_{i,15})$



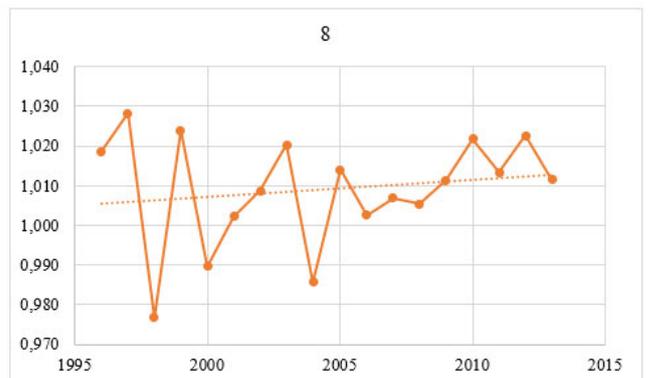
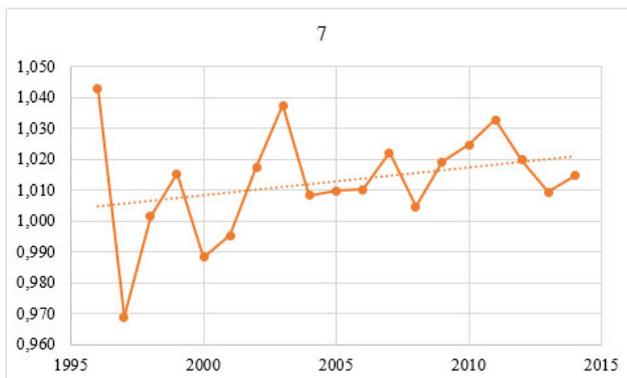
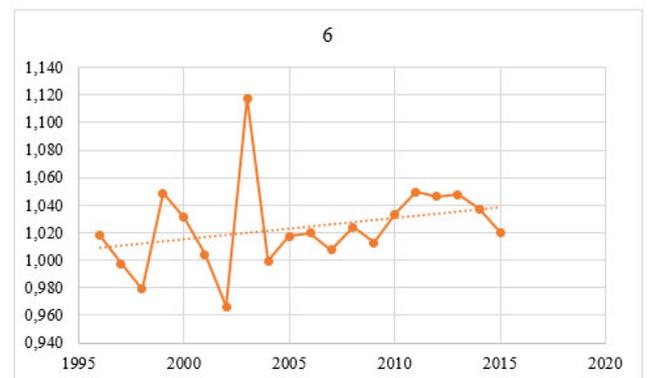
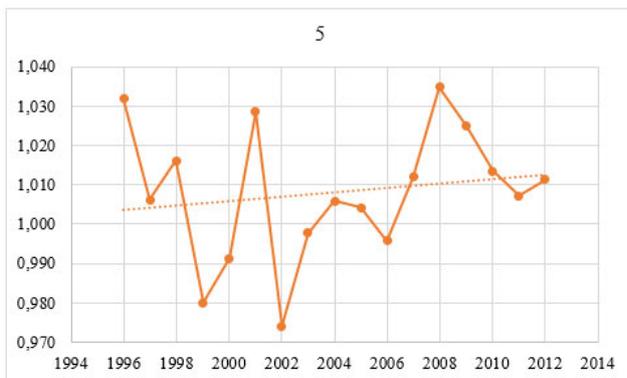
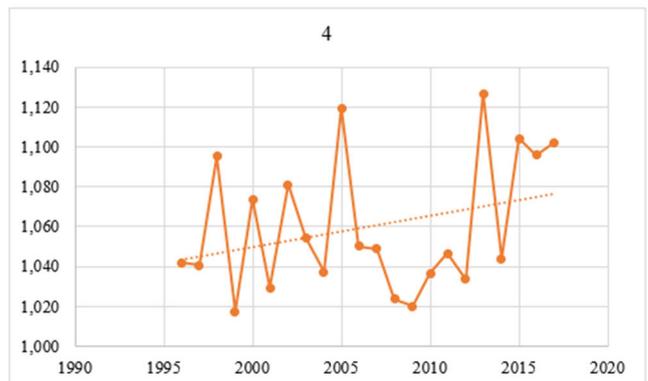
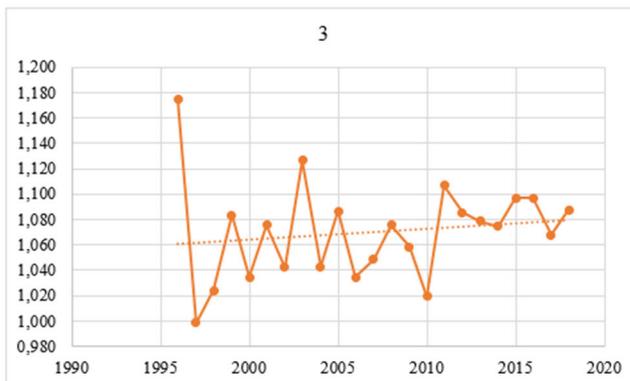
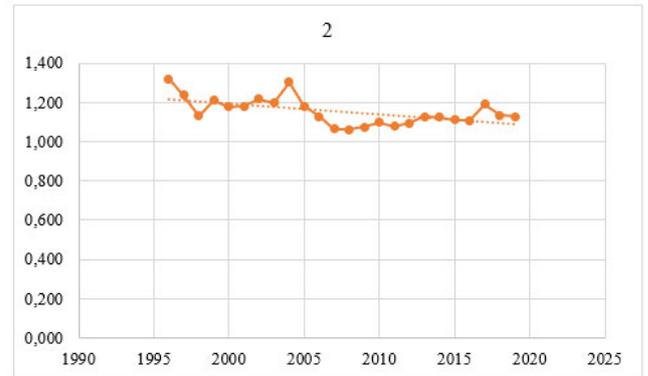
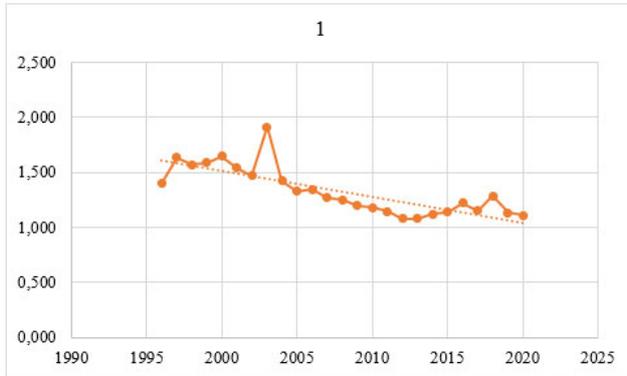
$(C_{i,15}, C_{i,16})$

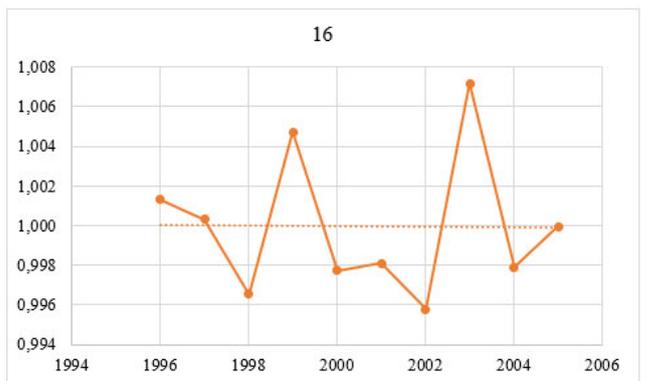
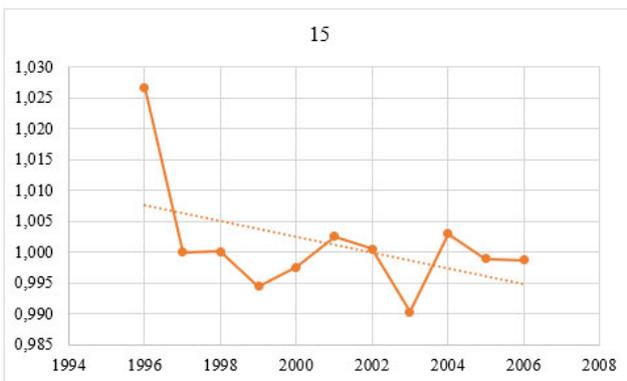
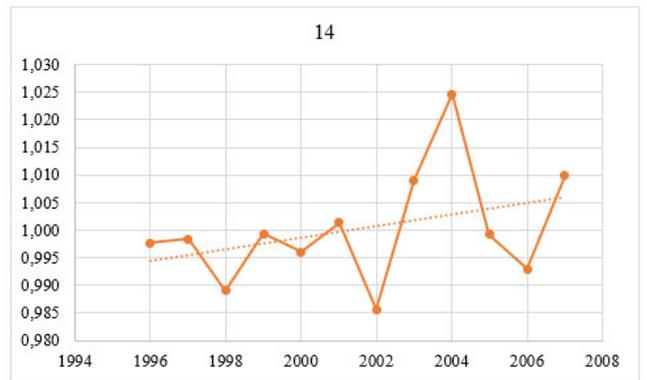
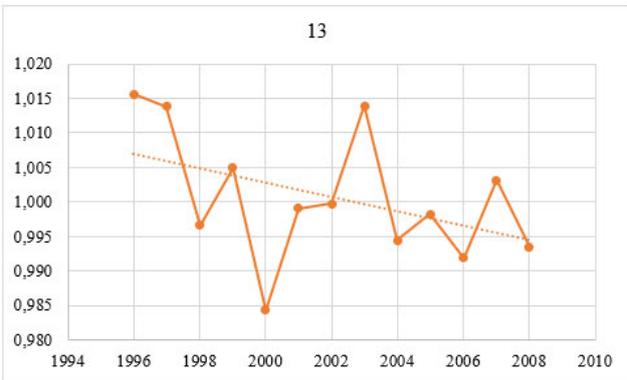
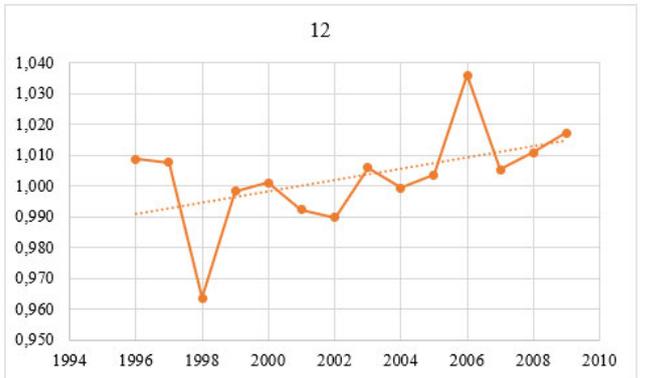
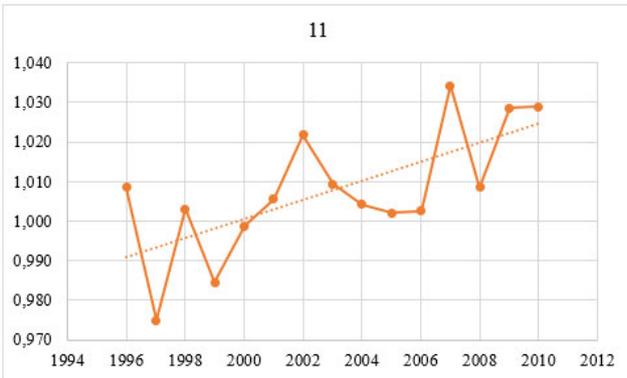
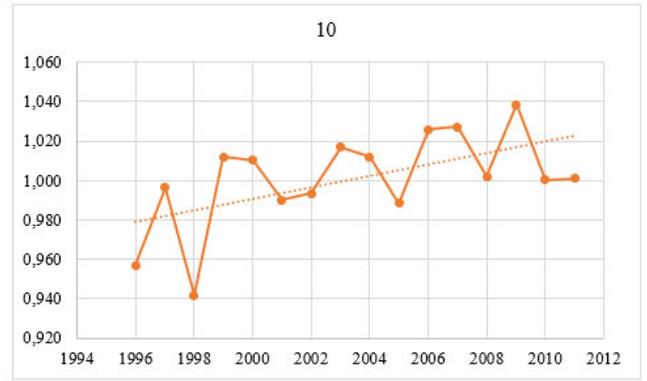
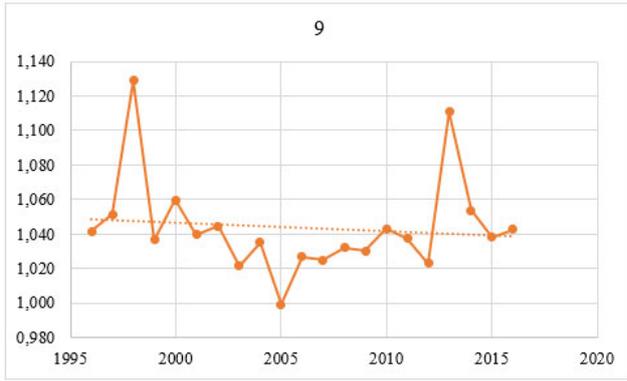


$(C_{i,16}, C_{i,17})$

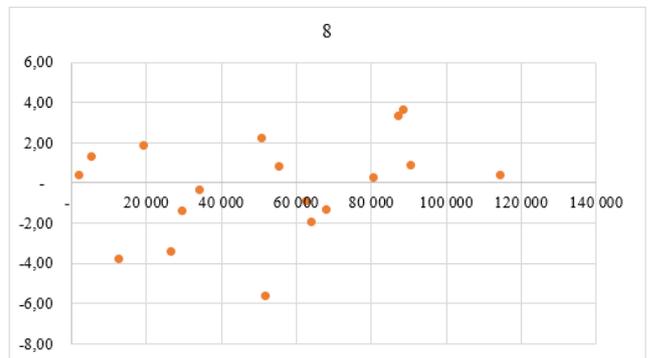
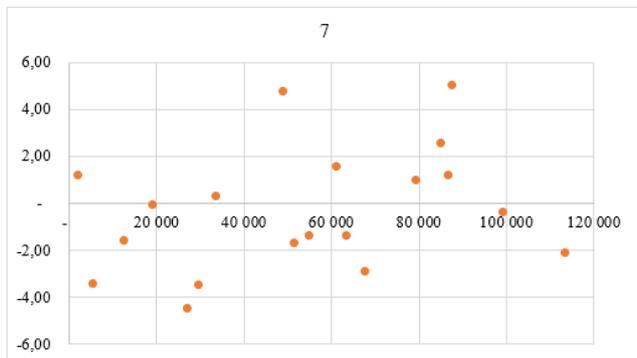
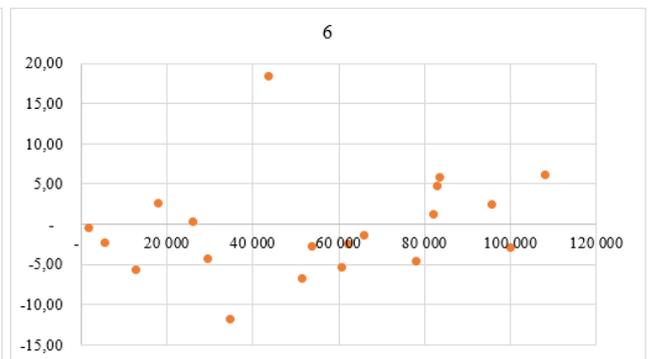
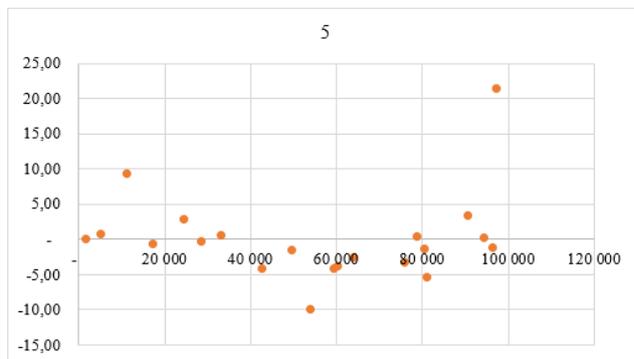
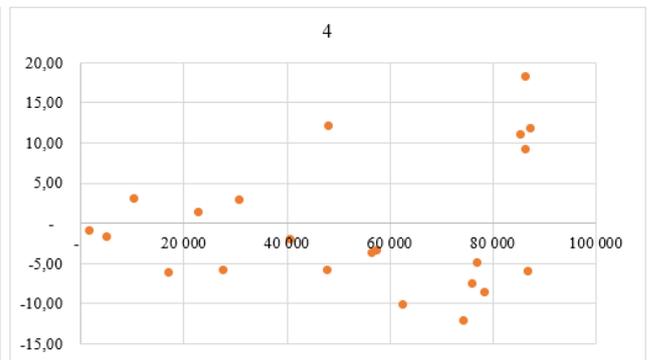
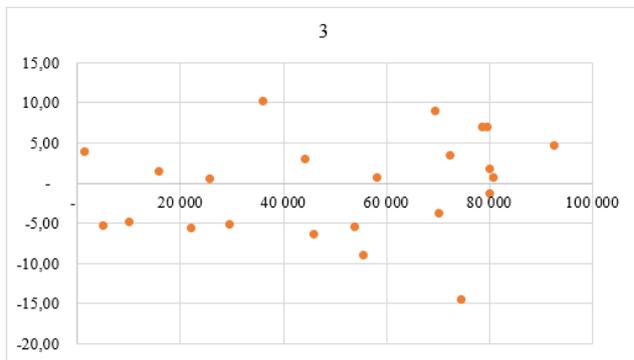
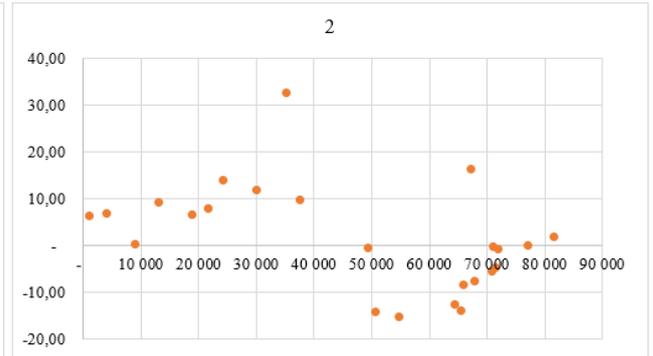
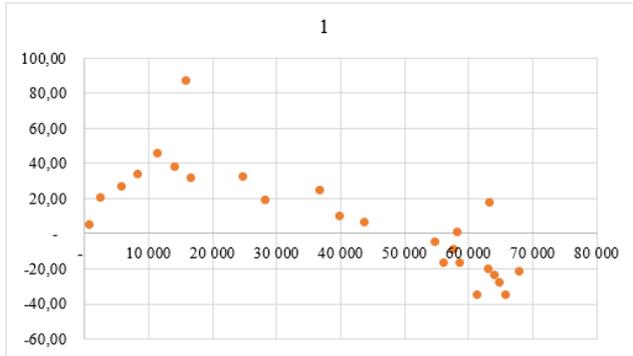


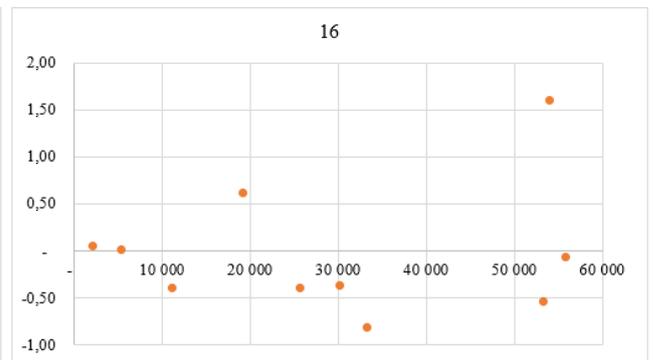
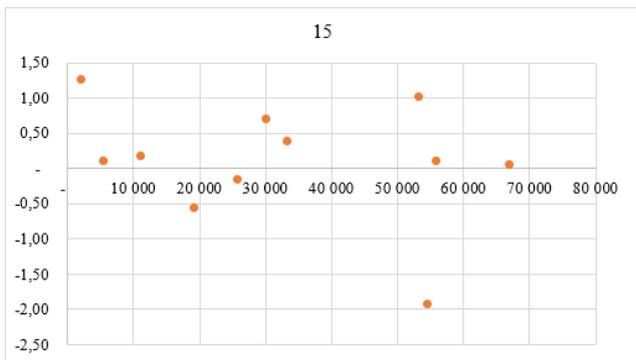
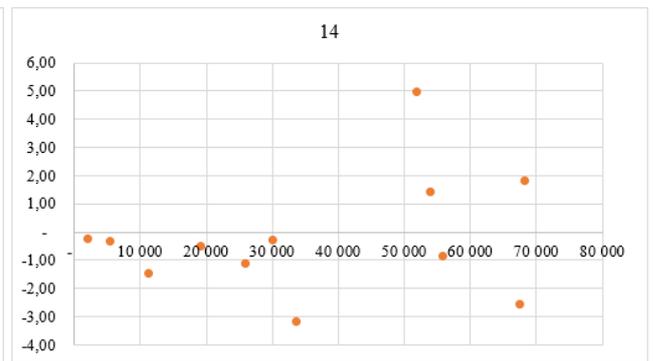
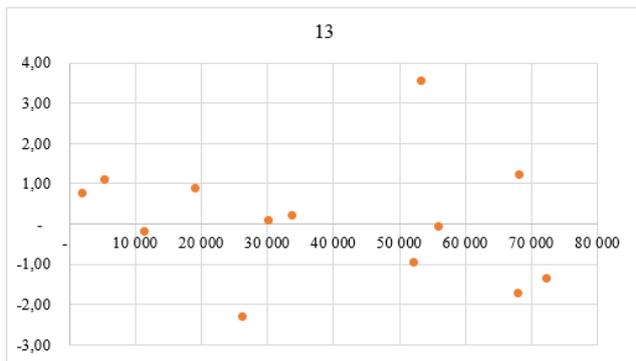
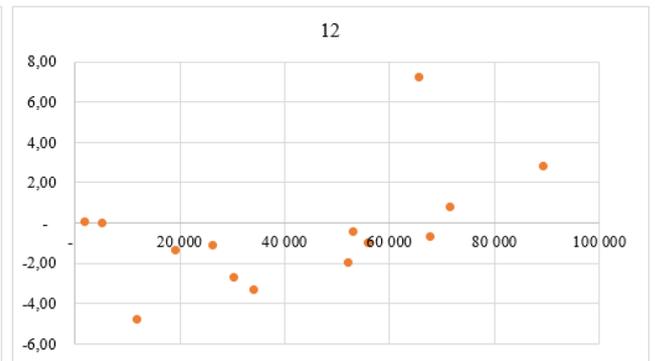
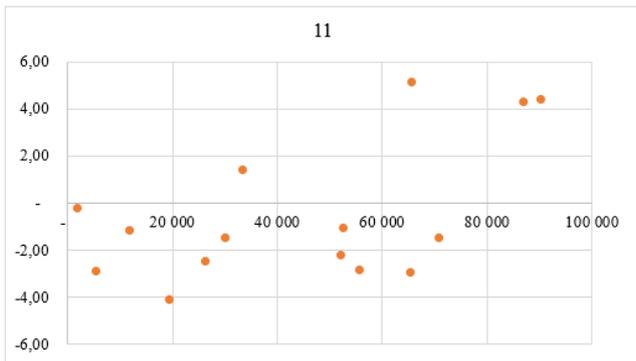
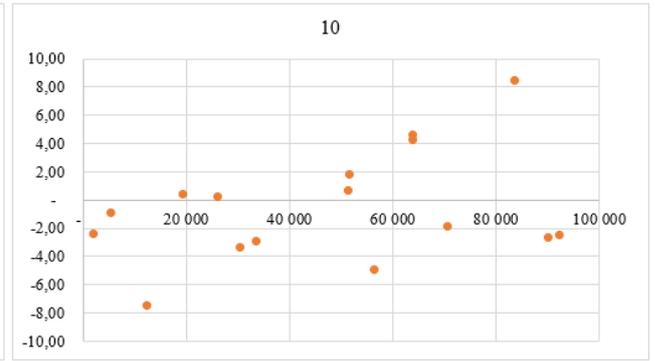
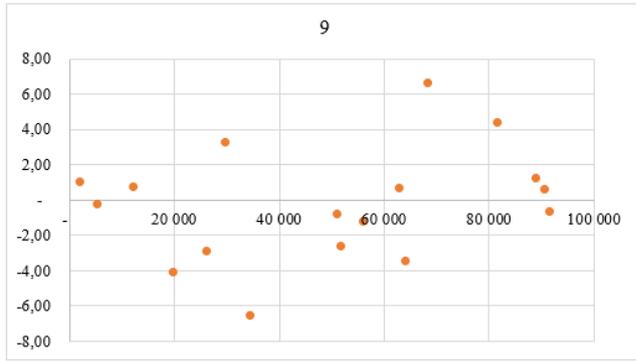
Annexe 4 : Graphiques test d'indépendance vision Survenance × Développement





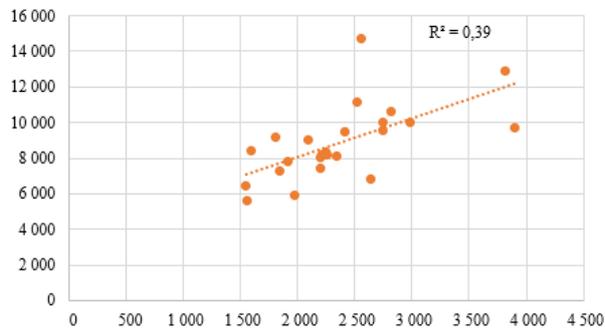
Annexe 5 : Graphiques test de variance vision Survenance × Développement



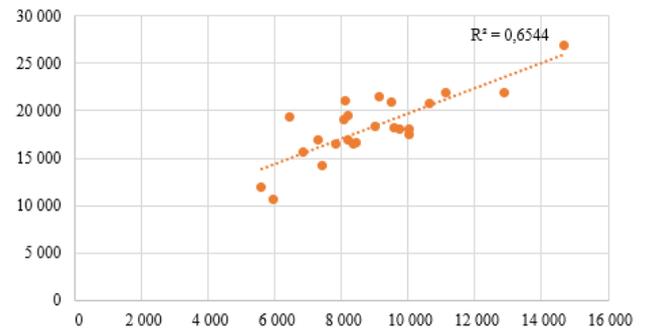


Annexe 6 : Graphiques test de linéarité vision DOC × Survenance

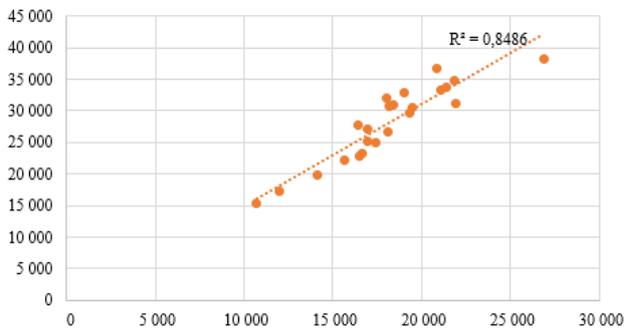
$(C_{i,1}, C_{i,2})$



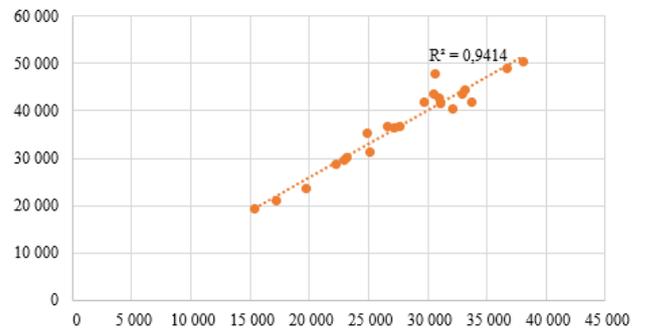
$(C_{i,2}, C_{i,3})$



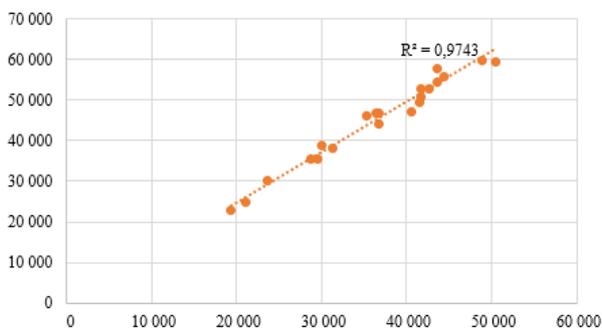
$(C_{i,3}, C_{i,4})$



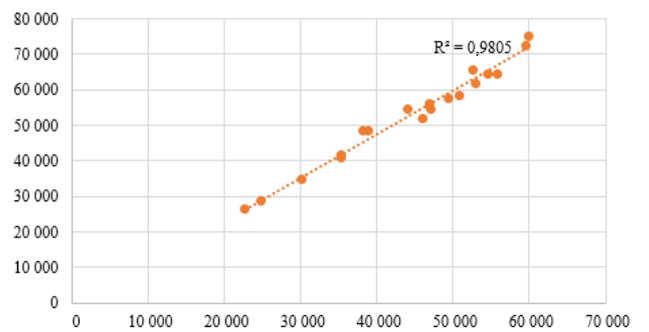
$(C_{i,4}, C_{i,5})$



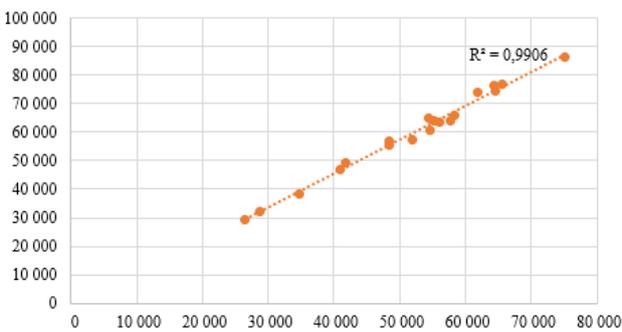
$(C_{i,5}, C_{i,6})$



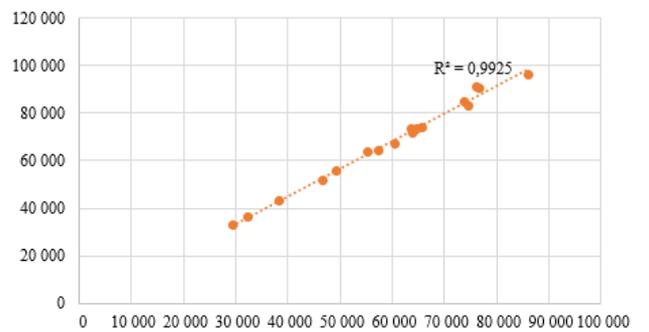
$(C_{i,6}, C_{i,7})$

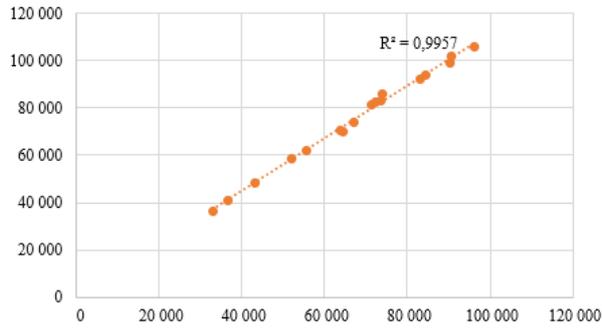
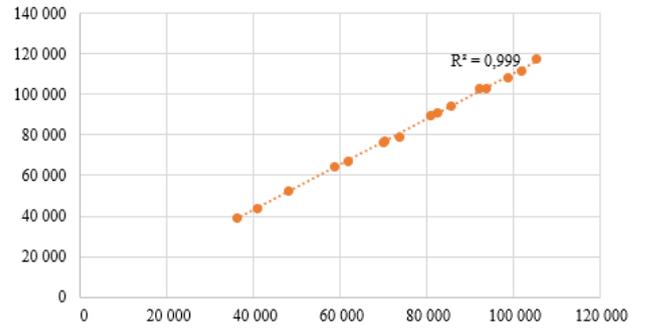
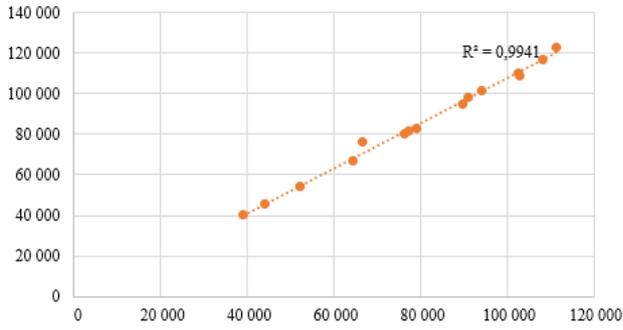
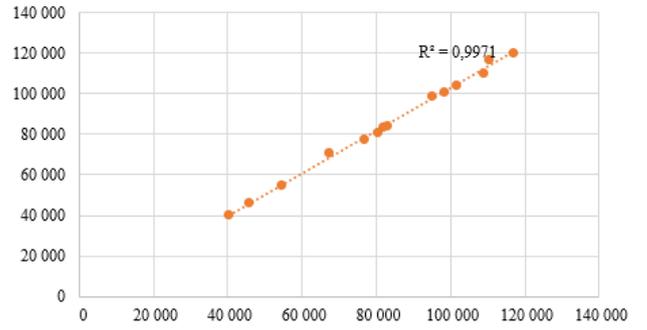
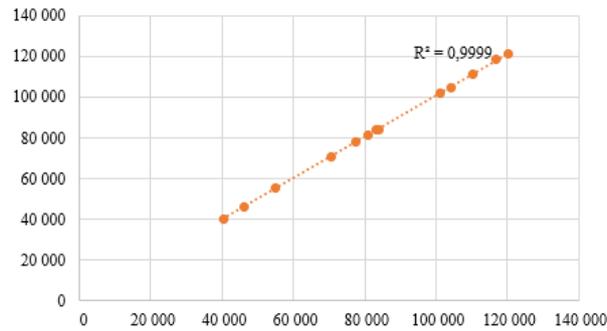
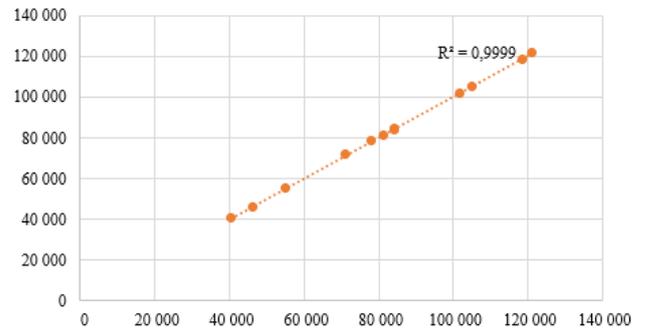
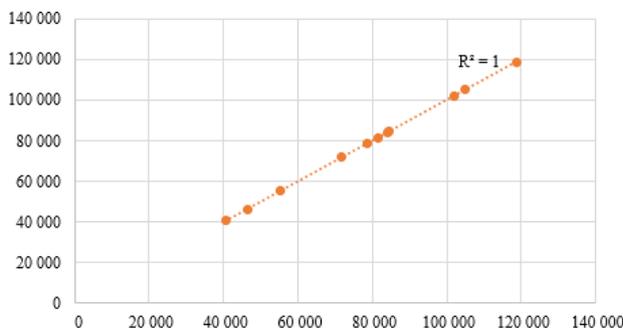
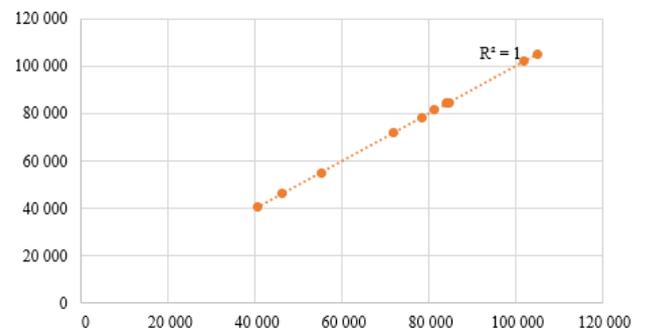


$(C_{i,7}, C_{i,8})$

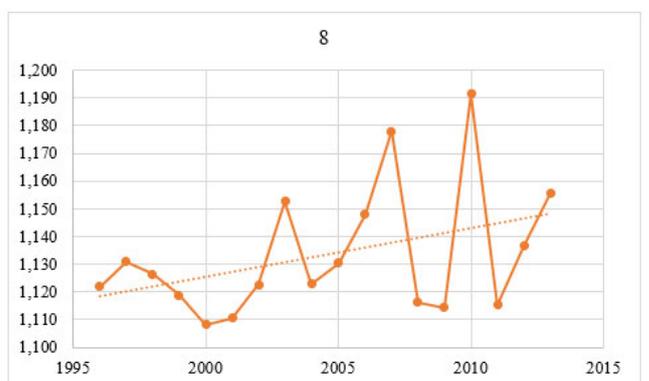
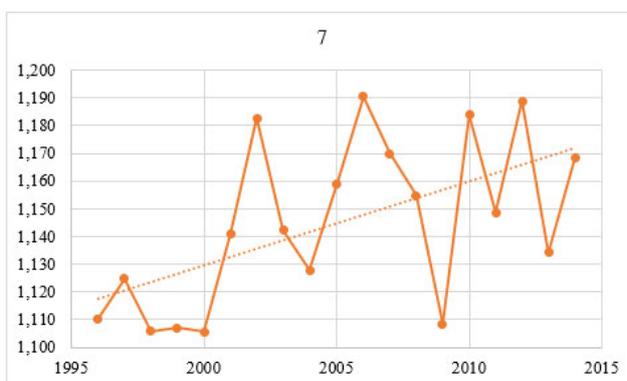
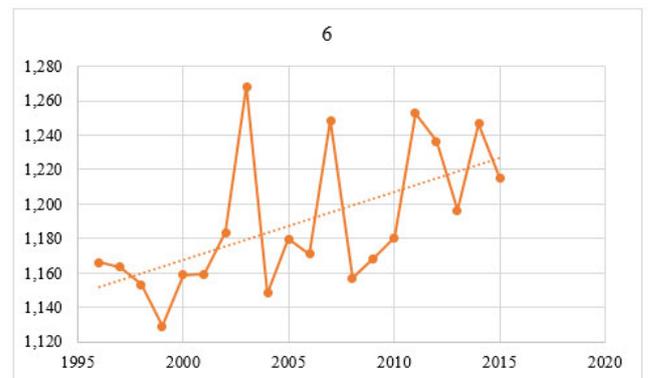
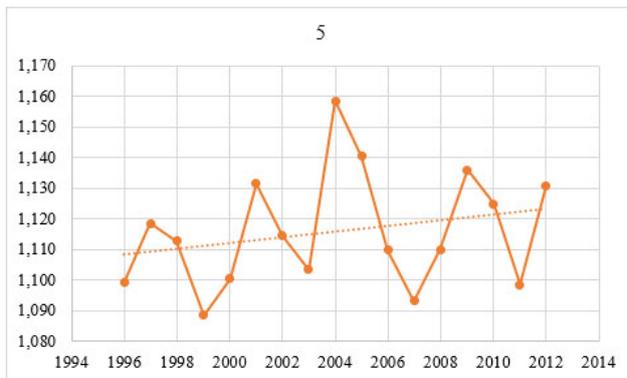
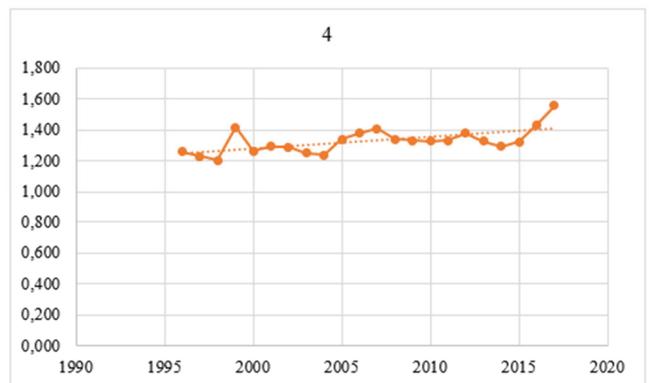
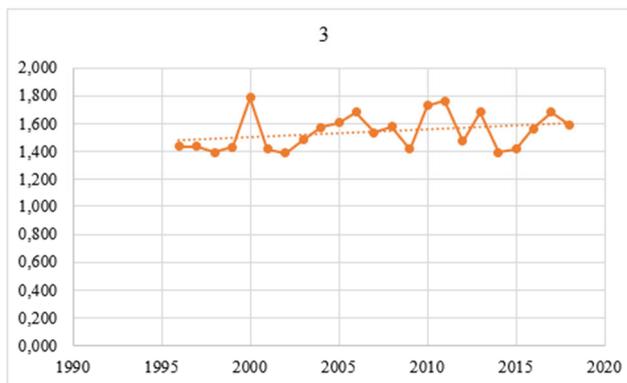
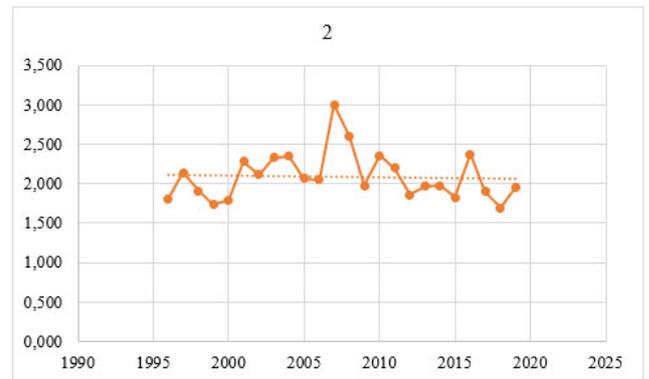
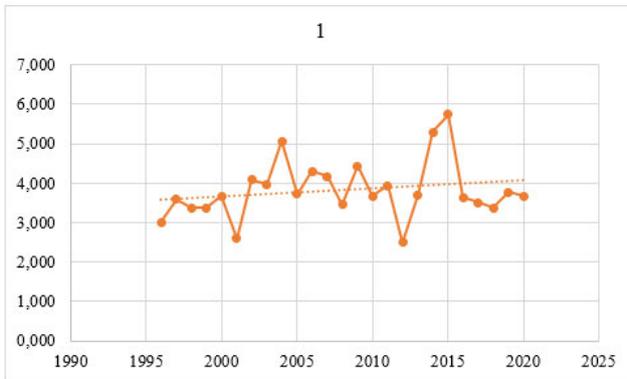


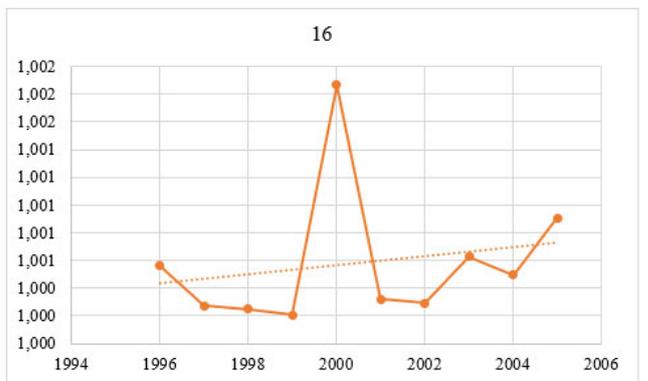
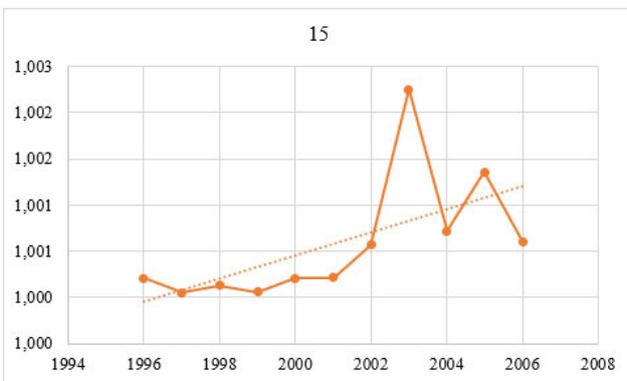
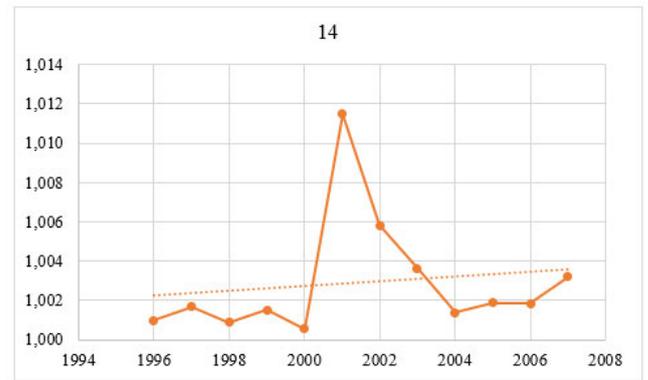
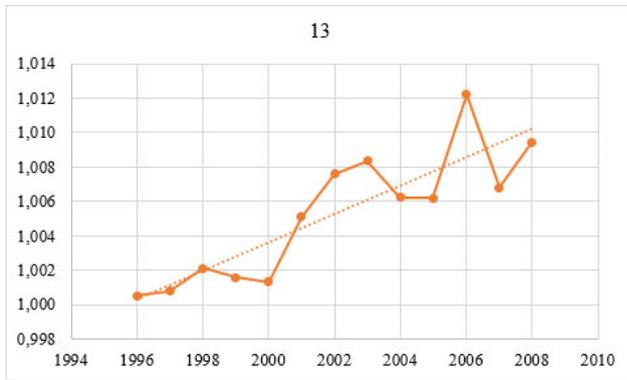
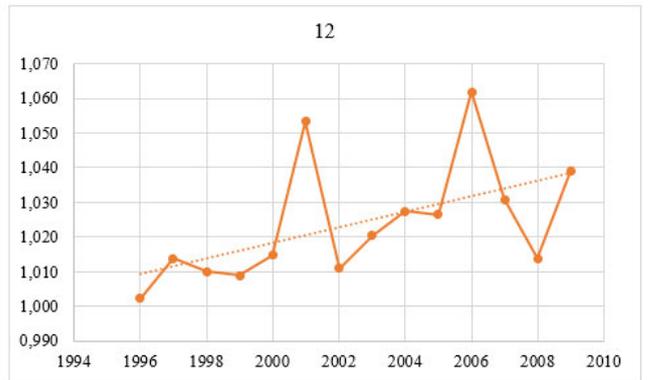
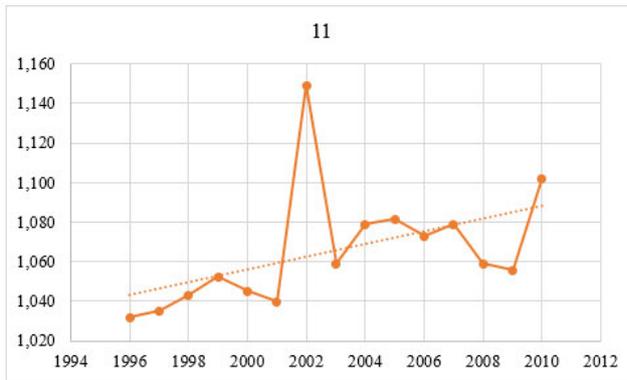
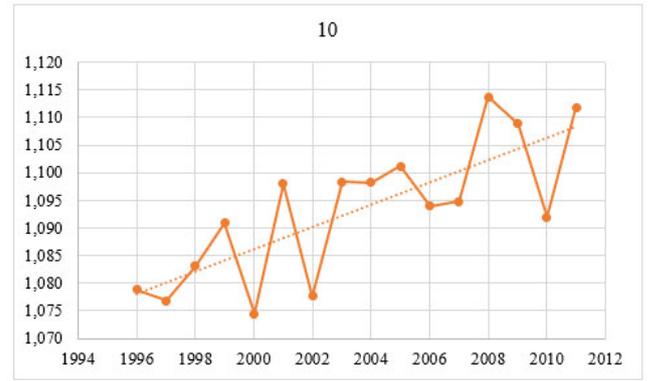
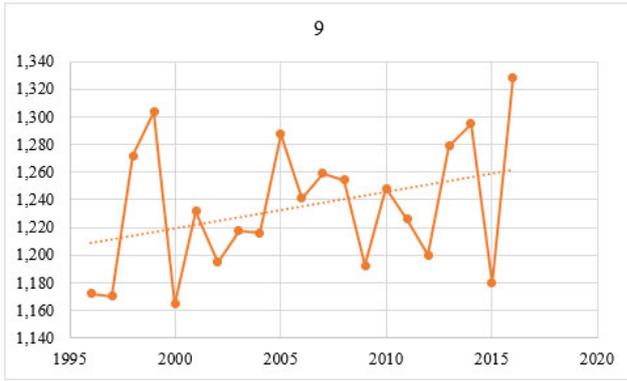
$(C_{i,8}, C_{i,9})$



$(C_{i,9}, C_{i,10})$  $(C_{i,10}, C_{i,11})$  $(C_{i,11}, C_{i,12})$  $(C_{i,12}, C_{i,13})$  $(C_{i,13}, C_{i,14})$  $(C_{i,14}, C_{i,15})$  $(C_{i,15}, C_{i,16})$  $(C_{i,16}, C_{i,17})$ 

Annexe 7 : Graphiques test d'indépendance vision DOC × Survenance





Annexe 8 : Graphiques test de variance vision DOC × Survenance

