

**Mémoire présenté devant l'ENSAE  
pour l'obtention du diplôme d'Actuaire ENSAE  
et l'admission à l'Institut des Actuaire**

**le 04/11/2015**

Par : Clothilde DAVESNE

Titre: Etude du risque de modèle dans le cadre d'un modèle interne

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

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

Membre présent du jury de l'Institut  
des Actuaire :

Entreprise :

Nom : SCOR SE

Signature :

Membres présents du jury de l'ENSAE :

Directeur de mémoire en entreprise :

Nom : Ecaterina NISIPASU

Signature :

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

Signature du responsable entreprise :

Secrétariat :

Bibliothèque :

Signature du candidat :

# Résumé

Dans le cadre de la réglementation Solvabilité II, le régulateur demande aux compagnies ayant mis en place un modèle interne de s'intéresser à un risque peu étudié jusqu'à présent, le risque de modèle. L'erreur de modèle est constituée de multiples sources d'erreurs présentes à différentes étapes de la modélisation. Cette étude se concentre essentiellement sur l'erreur de spécification en considérant les autres sources d'erreurs comme déjà intégrées par ailleurs. Dans une première partie, cette étude vise à reprendre et améliorer des travaux précédents qui s'intéressaient à l'élaboration d'une méthode d'évaluation du risque de modèle sur des risques vus individuellement. Deux approches sont donc étudiées et comparées pour mesurer ce risque, l'approche *worst case* (pire des cas) et l'approche bayésienne (par pondération). L'approche par pondération a l'avantage de prendre en compte la pertinence des modèles lors de l'appréciation de l'erreur. Dans une seconde partie, ces approches sont appliquées sur des modèles d'agrégation de deux risques, ce qui représente la nouveauté de cette étude. L'objectif est d'évaluer conjointement l'erreur liée au choix des marginales et celle liée au choix de la copule. Les approches sont mises en œuvre sur des modèles d'actifs et des modèles de provisionnement non-vie. Sur les risques étudiés, l'erreur de modèle est non significative pour les modèles d'actifs contrairement aux modèles de provisionnement. Les résultats montrent par ailleurs l'importance du choix des modèles considérés et de l'approche d'évaluation de l'erreur. Ce papier présente donc les différentes méthodologies d'évaluation de l'erreur proposées dans la littérature et en évoque les limites. Il donne, par ailleurs, une première approche pour l'estimation de l'erreur d'agrégation de deux risques qui pourrait ensuite être élargie à des modèles de dépendance comportant davantage de risques tels que le modèle interne.

Mots clés : Risque de modèle, Erreur de modèle, Mesure de risque, Value-at-Risk, Solvabilité II, Approche *worst case*, Approche bayésienne, Agrégation, Copules, Risques non-vie, Risque action

# Abstract

Under the Solvency II Directive, the regulator requires companies that have implemented an internal model to focus on a risk little studied so far, the model risk. Multiple sources of error present at different stages of the modeling comprise the model error. This study focuses primarily on the specification error considering other sources of error as already integrated elsewhere. The first part of this study aims to resume and improve previous work which aimed to develop a method for assessing the model errors on risks seen individually. Two approaches will therefore be studied and compared to measure this risk, the worst case approach and the Bayesian approach (or weighting approach). The weighting approach has the advantage of taking into account the model's relevance in the assessment of error. In the second part, these approaches are applied to two risks aggregation models, representing the novelty of this study. The goal is to jointly assess the error related to the choice of marginal and that related to the choice of the copula. The approaches are implemented on asset models and non-life reserving models. On the relevant segments, the model error is not significant on asset models unlike for reserving models. The results also show the importance of the choice of the models considered and the assessment error approach. This paper therefore presents different error assessment methodologies proposed in existing work and discusses their limitations. Moreover, it provides a first approach estimating the two risk aggregation error which could then be extended to dependency models with more risks such as an internal model.

Key words : Model risk, Model error, Measure of risk, Value-at-Risk, Solvency II Directive, Worst case approach, Bayesian approach, Aggregation, Copulas, Non-life risks, Equity risks

# Note de synthèse

## Contexte et objectifs de l'étude

Cette étude a été effectuée dans l'entreprise de réassurance SCOR. Elle s'inscrit dans le cadre de la réglementation Solvabilité II qui entre en vigueur en Janvier 2016. Ce nouveau dispositif normatif s'articule autour de trois piliers ; nous nous concentrons uniquement sur le premier et les exigences quantitatives qu'il impose. Un assureur ou réassureur doit être capable de démontrer qu'il dispose de suffisamment de fonds propres pour éviter la ruine à horizon 1 an avec une probabilité de 99.5%. Le régulateur laisse le choix du mode de calcul de l'exigence en capital, soit par une formule standard, soit par un modèle interne. Le développement d'un modèle interne, mode d'évaluation privilégié par la SCOR, est à l'origine de cette étude. En effet, une compagnie ayant son propre modèle interne doit être en mesure d'établir une cartographie précise de ses risques afin de les intégrer dans son modèle. Un risque qui n'est pas toujours bien évalué est le risque de modèle.

Le risque de modèle se définit comme l'ensemble des risques engendrés par le développement d'un modèle mathématique ou statistique. Ainsi, il existe de multiples sources d'erreur de modèle, parmi lesquelles les erreurs de données, les erreurs d'estimation, les erreurs d'implémentation ou encore les erreurs opérationnelles. Cette étude se concentre essentiellement sur l'erreur de spécification du modèle, c-à-d. le choix du modèle en lui-même.

Il n'y a pas à l'heure actuelle de standards concernant le risque de modèle en assurance et en réassurance. Les régulateurs appellent à la prudence, à la vérification et validation des modèles par des entités indépendantes, et à la mise en place de gouvernance interne aux institutions. Bâle II a formalisé un certain nombre de critères devant être respectés lors de l'élaboration d'un modèle et incite à provisionner des provisions pour prendre en compte l'incertitude de modèle sans proposer de méthodes pour les calculer.

Cette étude vise donc à présenter et mettre en application deux approches proposées dans la littérature pour l'appréciation de l'erreur de modèle. Tout d'abord, nous nous focalisons sur l'évaluation de l'erreur dans le cas d'un risque isolé afin de développer la méthodologie, en comprendre les limites et en interpréter les résultats pour des risques différents. Par la suite, nous ne nous plaçons plus dans le cadre d'un actif isolé mais dans le cadre d'un portefeuille de risques afin de prendre en compte l'erreur de dépendance. Etant donné que l'erreur de modèle d'agrégation n'a pas fait l'objet de littérature extensive, nous nous focalisons sur la dépendance de deux risques et sur la méthodologie de calcul de l'erreur pour ces modèles.

## Méthodologie

L'objectif d'un modèle statistique  $M$  est de prédire ou d'inférer sur une variable d'intérêt  $y$  (ou risque) sachant une quantité d'information connue (données historiques, avis d'experts...). Deux approches d'évaluation de l'erreur de modèle, reposant toutes les deux sur un certain nombre d'hypothèses communes, ont été proposées dans la littérature. Ces hypothèses correspondent à des choix à effectuer préalablement :

Choix d'un modèle de référence  $M^*$  : C'est le modèle utilisé en pratique pour modéliser le risque considéré, généralement celui qui représente le meilleur compromis entre adéquation et simplicité d'implémentation.

Choix d'un ensemble de modèles  $M = (M_1, \dots, M_n)$  : Cet ensemble regroupe tous les modèles ayant prouvé leur adéquation à modéliser ce risque et se rapprochant du 'vrai' modèle. Le 'vrai' modèle, par définition inconnu, est celui qui nous permettrait d'obtenir la vraie distribution de la variable d'intérêt.

Choix de la mesure de risque pour évaluer les exigences en capital d'un modèle, notée  $\mathcal{R}(M^*)$ . Dans l'ensemble de l'étude, nous avons retenu la Value-at-Risk à 99.5% comme mesure de risque pour être en adéquation avec le cadre réglementaire Solvabilité II.

Le capital réglementaire total,  $\rho_{total}$ , est alors constitué du capital réglementaire pour le modèle de référence,  $\mathcal{R}(M^*)$ , et du capital pour tenir compte de l'erreur de modèle,  $\rho(M^*)$ .

$$\rho_{total} = \mathcal{R}(M^*) + \rho(M^*)$$

### Approche *worst case*

L'approche *worst case* ou du pire des cas est une approche prudente qui évalue la différence entre le capital réglementaire obtenu avec le modèle de référence et celui obtenu avec le pire modèle, le modèle nécessitant le plus de capital réglementaire. Cette approche est appréciée par sa simplicité de mise en œuvre mais a l'inconvénient d'être trop prudente puisqu'elle évalue le capital maximal nécessaire pour tenir compte de l'erreur de modèle.

Plusieurs mesures d'évaluation relatives au pire des cas ont été proposées, nous nous focalisons sur la mesure absolue :

$$AM = \rho(M^*) = \frac{\max_{M_i \in \mathcal{M}} \mathcal{R}(M_i)}{\mathcal{R}(M^*)} - 1$$

Cette mesure quantifie à quel point le pire scénario est mauvais par rapport au modèle de référence. Elle est nulle si le modèle de référence est le pire modèle et est d'autant plus faible que le modèle de référence est prudent (nécessite de capital).

### Approche bayésienne

L'approche bayésienne est une approche par pondération. Cette méthode nécessite la connaissance de la vraisemblance de chaque modèle, notée  $L_M$ , qui n'est pas toujours disponible.

La pondération *a posteriori* (bayésienne) est évaluée à partir d'un *a priori* pour chaque modèle, noté  $p(M_j)$ , telle que :

$$p_{post}(M_i) = \frac{L_{M_i} p(M_i)}{\sum_j L_{M_j} p(M_j)}$$

En l'absence de la vraisemblance, la pondération est évaluée à dire d'experts et au vu d'éléments statistiques (tests d'adéquation, aspects qualitatifs...).

Pour mesurer l'erreur, des simulations par bootstrap sont effectuées. Le principe est de simuler une distribution de la variable d'intérêt projetée à horizon 1 an (pour être en adéquation avec le cadre réglementaire). Chaque observation de la distribution est issue d'un des modèles de l'ensemble qui a été préalablement tiré avec remise à l'aide des pondérations (bayésiennes ou non). Le capital réglementaire  $\mathcal{R}_{Bootstrap}(M)$  est évalué sur cette distribution obtenue par bootstrap des modèles et l'erreur de modèle par cette approche est définie par :

$$\rho(M) = \frac{\mathcal{R}_{Bootstrap}(M)}{\mathcal{R}(M^*)} - 1$$

L'approche bayésienne est développée sur les mêmes hypothèses fortes que l'approche *worst case* mais donne une meilleure précision dans le calcul de l'erreur. Par conséquent, le capital requis pour couvrir cette erreur est moindre. La difficulté réside essentiellement dans le calcul de la vraisemblance lorsque cela est possible et des pondérations.

## Principaux Résultats

Ces deux approches ont été mises en application sur des risques actions britannique et américain et sur des risques de réassurance non-vie transport maritime et automobile. Les risques d'action sont la baisse significative de cours boursiers journaliers mesurant le prix moyen brut d'un ensemble d'actifs. L'erreur de modèle est évaluée, à partir d'un historique de 2001 à 2013, sur la distribution prédite du cours à fin 2014. Ces risques ont donc un historique relativement long nous permettant de mettre en place les méthodes statistiques décrites.

Les risques non-vie correspondent aux sinistres des segments transport maritime et automobile. Le risque de modèle est estimé sur des sinistres agrégés payés sur la période 1999-2013. Le développement des sinistres est donc moyen.

### Etude d'un risque seul

Tout d'abord, les deux approches sont appliquées sur des modèles à une dimension, c-à-d. qui ne permettent de modéliser qu'un risque.

#### *Modèle d'actifs*

La première partie de l'étude consiste à choisir les modèles permettant de prédire le cours boursier à un an et à effectuer la validation de chacune de leurs hypothèses afin de vérifier leur adéquation à nos données. Pour modéliser les actifs, nous avons gardé quatre modèles : le modèle de Black-Scholes, le modèle de Merton, le modèle de Garch et le modèle de Black-Scholes à changements de régime.

Au vu des tests de validation, le modèle de Garch semble être, dans notre étude, le plus adéquat. Notre choix se porte donc sur ce modèle comme référence. Le modèle de Black-Scholes possède une bonne adéquation malgré le fait que la limitation à deux régimes (par simplicité) soit probablement trop restrictive. Les deux autres modèles ne sont pas très adaptés à notre étude. Nous décidons de les garder dans la mesure où ces modèles sont souvent utilisés en pratique malgré les critiques à leur encontre.

L'erreur de modèle  $a$ , en premier lieu, été évaluée par l'approche bayésienne. Les pondérations bayésiennes ont pu être évaluées puisque la vraisemblance de chacun des modèles était disponible. Le choix de la distribution *a priori* n'étant pas évident, nous avons choisi de donner un poids légèrement plus élevé au modèle de Garch. Par ailleurs, nous avons réalisé une analyse de sensibilité sur l'*a priori* pour tester l'impact de pondérations plus extrêmes (par exemple, en ne donnant aucun poids au modèle de Black-Scholes qui semble moins adapté). Dans tous les cas, la distribution *a posteriori* donne un grand poids au modèle de Garch, un poids plus faible au modèle de Black-Scholes à changements de régime et des poids non significatifs aux autres modèles.

Finalement, l'erreur de modèle est très proche quel que soit l'*a priori* choisi et est négative. Cela s'explique par le fait que le modèle de Garch est le plus prudent ; le capital réglementaire requis avec ce modèle est supérieur à celui obtenu par bootstrap. Une erreur de modèle négative n'impliquera pas une reprise du capital réglementaire pour des questions de prudence, l'erreur est donc supposée nulle.

Le risque de modèle a été évaluée, en deuxième lieu, par l'approche *worst case*. L'erreur est nulle dans ce cas puisque le modèle de Garch est le plus prudent.

#### *Modèle de provisionnement*

Concernant les modèles de provisionnement, nous avons retenu trois classes de modèle pour simuler la distribution des provisions estimées : la classe bootstrap Mack, la classe bootstrap ODP et la classe MCMC GLM. Chaque classe contient plusieurs modèles qui diffèrent en terme de distribution de l'erreur (par exemple, l'erreur du bootstrap Mack peut être modélisée par la loi gamma ou par rééchantillonnage...). Les hypothèses des modèles étant majoritairement validées, nous gardons chacune des classes et chacun de leurs modèles. Le modèle de référence est le modèle 'log-normal'. Ce

modèle simule la distribution des provisions estimées par une loi log-normale dont les paramètres sont évalués à partir du *best estimate* et de la variance du modèle de Mack (avant bootstrap). L'erreur est d'abord estimée intra-classe, c-à-d. entre les modèles d'une classe. L'évaluation de l'erreur extra-classe est établie en pondérant chaque classe et en agrégeant les erreurs de chaque classe (erreur intra-classe) à l'aide de ces pondérations. Les poids sont choisis équiprobables dans la mesure où nous ne sommes pas capable de différencier les classes en termes d'adéquation. Pour ce faire, il faudrait, par exemple, des avis d'experts.

Excepté pour les modèles MCMC GLM, la vraisemblance des modèles n'est pas évaluable. Par ailleurs, le manque de données (du fait de l'agrégation des sinistres dans les triangles de provisionnement) rend l'évaluation des modèles plus difficile. La pondération a été choisie équiprobable au sein des modèles d'une classe (excepté pour la classe MCMC GLM) dans la mesure où les résultats étaient assez similaires pour chaque modèle et que nous avons pas d'*a priori* sur l'adéquation des modèles. Pour la classe MCMC GLM, nous utilisons l'approche bayésienne à partir de la vraisemblance. La pondération *a posteriori* met quasiment l'ensemble du poids au modèle GLM à erreur Poisson. Le modèle GLM à erreur Gamma étant le plus conservateur, nous préférons, par prudence, modifier la pondération en ne lui attribuant que 2/3 du poids et le reste sur le modèle GLM à erreur gamma.

L'erreur globale est de 1.94% par l'approche par pondération et de 2.94% par l'approche par le pire des cas. Elle n'est donc pas très élevée et peut être considérée comme négligeable. Cependant, les résultats sont très variables en fonction de la pondération des classes et de celle des modèles intra-classes. L'erreur maximale est de 6.9% si la classe MCMC GLM a tout le poids et que l'erreur intra-classe est évaluée par l'approche *worst case*. L'erreur maximale est significative (elle dépasse 5%) mais est issue de modèles qui ne semblent pas adéquats.

## Etude de dépendance

La dépendance a été modélisée par des copules connues (gaussienne, Student, Clayton, Gumbel et Frank) afin de prendre en compte des dépendances non-linéaires. Au même titre que les modèles des marges (modèles d'actifs ou de provisionnement spécifiés au préalable), les copules et leur paramétrisation ont fait l'objet d'une validation en effectuant des tests d'adéquation par exemple. Pour évaluer l'erreur sur des modèles de dépendance, nous avons procédé en deux étapes. D'une part, nous avons évalué l'erreur de dépendance pure, c-à-d. l'erreur due au choix de la copule. Pour cela nous avons fixé les marges (modélisées par le modèle de référence) et fait varier uniquement la copule. L'erreur de copule est évaluée par les deux approches. Pour l'approche par pondération, les poids sont évalués à l'aide de la vraisemblance lorsque nous pouvons l'évaluer et à partir d'informations qualitatives, telles que les avis d'experts, sinon. D'autre part, nous avons évalué l'erreur totale de dépendance, incluant l'erreur liée au choix de la copule et celle liée au choix du modèle des marges. Les modèles référents pour simuler les marges sont les mêmes que dans le cas d'un risque seul.

### Modèle d'actifs

Pour les modèles d'actifs, la copule de référence est la copule de Student à 12 degrés de liberté, la copule ayant la meilleure adéquation. Etant donné que les données sont suffisantes pour évaluer la vraisemblance des copules, l'approche par pondération est une approche bayésienne.

Sur nos données, l'erreur de dépendance pure est non significative par l'approche bayésienne et relativement élevée par l'approche *worst case*. En effet, la copule de Frank est la plus prudente mais aussi la moins vraisemblable. L'approche par pondération est justifiée puisqu'elle ne donne quasiment pas de poids à cette copule, l'erreur est donc plus réaliste.

Concernant la dépendance globale, les résultats sont assez similaires avec une erreur de 6% par l'approche *worst case* mais négative (-1.6%) par l'approche bayésienne. Ces résultats s'expliquent aussi par l'impact de la copule de Frank lors du calcul de l'erreur. L'approche par pondération semble plus cohérente, c'est pourquoi l'erreur de modèle est considérée nulle pour cette étude.

## Modèle de provisionnement

Pour les modèles de provisionnement, la copule de référence est la copule de survie de Clayton (copule pivotée à 180 degrés). La paramétrisation des copules est issue principalement d'avis d'experts du fait du manque de données. Les vraisemblances des copules n'étant donc pas calculables, les pondérations sont choisies équiprobables ou favorisant la copule de survie de Clayton. Dans les deux cas, les résultats sont identiques.

Sur nos données, l'erreur de dépendance pure est non significative quelle que soit l'approche considérée.

L'erreur de dépendance globale est quant à elle très proche de l'erreur obtenue dans le cas d'un risque seul. Elle est donc assez élevée notamment lorsque le poids est fort sur les modèles MCMC GLM.

## Conclusions et Limites

L'erreur de modèle évaluée est globalement non significative sur les risques étudiés. Seule l'erreur sur les modèles de provisionnement est positive et importante pour certains modèles. Les résultats sont à prendre avec prudence sur les modèles de provisionnement du fait du manque de données. Ainsi, bien que les GLM ne semblent pas toujours adéquats, nous préférons les garder pour ne pas risquer de sous-estimer l'erreur, n'ayant pas d'arguments quantitatifs permettant de les invalider.

Cette étude ne permet pas de généraliser les résultats à l'ensemble des segments en particulier pour les risques non-vie où il serait nécessaire d'élargir l'étude à d'autres sinistres non-vie.

Les deux approches présentent des limites en particulier le coût que représente l'implémentation de chacun des modèles et le temps de calcul. Elles font toutes deux des hypothèses très fortes concernant le choix de l'ensemble des modèles et du modèle de référence. En modifiant le modèle de référence, l'erreur peut considérablement augmenter (par ex, 16% d'erreur dans le pire des cas pour les modèles de dépendance d'actifs en prenant la copule de Frank comme référence). Par ailleurs, on suppose que le 'vrai' modèle appartient à l'ensemble, cette hypothèse n'est en théorie pas vérifiable. Cependant, en pratique, des actuaires expérimentés ont une bonne connaissance de ces modèles ; ainsi, l'ensemble des modèles retenus contient très probablement le 'vrai' modèle.

L'approche *worst case* est trop prudente dans la mesure où elle ne tient pas compte de la vraisemblance des modèles. Ainsi, nous avons constaté qu'en prenant trop de modèles dans l'ensemble, y compris des modèles peu adéquats, l'erreur dans le pire des cas pouvait être biaisée par ces modèles. Cependant, elle donne une idée de la prudence du modèle de référence par rapport aux autres modèles.

L'approche par pondération est quant à elle plus réaliste mais la grande difficulté de cette approche réside dans le choix des pondérations notamment dans les cas où les vraisemblances ne sont pas disponibles. Les résultats peuvent être très variables en fonction des pondérations.

Afin d'être en mesure de généraliser cette étude, il serait nécessaire de l'appliquer à d'autres segments pour vérifier que les résultats sont comparables. Par ailleurs, un travail plus approfondi sur le choix des pondérations devrait être réalisé dans le cas où les vraisemblances ne sont pas mesurables. Par exemple, des avis d'experts sur les modèles pourraient être demandés et agrégés. Pour finir, nous pourrions augmenter la taille du portefeuille de risques afin d'évaluer l'erreur sur des modèles d'agrégation de plusieurs risques (plus de deux). La prise en compte de la dépendance dans l'évaluation de l'erreur du modèle permet de se rapprocher du cadre concret d'un modèle interne.

# Executive summary

## Context and purpose of the study

This study was conducted in the reinsurance company SCOR and concerns a problem arising from the Solvency II Directive that comes into force in January 2016. This new regulatory system is based on three pillars. We focus only on quantitative requirements imposed in the first pillar. The SCR (Solvency Capital Requirement) is the capital required to ensure that the (re)insurance company will be able to meet its obligations over the next 12 months with a probability of at least 99.5%. The capital requirement can be calculated by a standard formula or by an internal model. Scorer has chosen to develop an internal model, which is why this study focuses on this valuation technique. Indeed, a company with its own internal model must be able to establish a precise mapping of its risks in order to integrate them into its model. A risk that is not always clearly identified is the model risk.

Model risk is defined as all risks arising from the development of a mathematical or statistical model. Thus, there are a multitude of model error sources including data errors, estimation errors, implementation errors or operational errors. This study focuses primarily on the error arising from model misspecification, ie the choice of the model itself.

There are currently no formal regulations on the model risk. Regulators urge caution, verification and validation of models by independent entities and establishment within institutions of internal governance. Basel II has formalized a number of criteria to be respected during the development of a model. It encourages provision of reserves to take into account the model uncertainty without proposing methods to calculate them.

This study aims to present and apply two approaches proposed in the literature for model error assessment. First, we focus on the evaluation of the error in the case of an individual risk to develop the methodology, understand the limitations and interpret the results for different risks. Subsequently, dependency models are studied in order to get closer to models used to simulate risk portfolios. Since the aggregation model error is not the subject of extensive literature, we focus primarily on the dependence of two risks and the methodology for calculating the error for these models.

## Methodology

The aim of a statistical model  $M$  is to infer or predict a variable of interest  $y$  (or risk) conditionally on a known quantity of information (historical data, expert judgement...). Two approaches to assess the model error, based both on a number of common assumptions, have been proposed in the previous work. These assumptions correspond to choices to be made :

Choice of a reference model  $M^*$  : The reference model is the model used in practice to model the risk in question. It is usually the one that gives the best compromise between suitability and ease of implementation.

Choice of a set of models  $M = (M_1, \dots, M_n)$  : This set includes all models that have proven their suitability to model this risk and approaching the 'true' model. The 'true' model is the model that would allow us to obtain the true distribution of the variable of interest. It is unknown by definition.

Choice of risk measure to assess capital requirements of a model, noted  $\mathcal{R}(M^*)$ . Throughout the study, we selected the Value-at-Risk at 99.5% as a measure of risk to be in line with the Solvency II Directive.

The total regulatory capital is then composed of regulatory capital for the reference model,  $\mathcal{R}(M^*)$ , and the capital to account for model error,  $\rho(M^*)$ .

$$\rho_{total} = \mathcal{R}(M^*) + \rho(M^*)$$

## Worst case approach

The worst case approach is a conservative approach that evaluates the difference between the regulatory capital obtained with the reference model and that obtained with the worst model, the model requiring the most regulatory capital. This approach is valued for its simplicity of implementation but has the disadvantage of being too cautious as it evaluates the maximum capital provision to account for model error.

Several risk assessment measures have been proposed in the worst case approach. We focus on the absolute measure which is always positive or zero :

$$AM = \rho(M^*) = \frac{\max_{M_i \in M} \mathcal{R}(M_i)}{\mathcal{R}(M^*)} - 1$$

This measure quantifies how bad is the worse scenario compared to the reference model. There is no error if the reference model. The more prudent the reference model, the lower is the error.

## Bayesian approach

The Bayesian approach is a weighting approach. This method requires the knowledge of the likelihood of the model, noted  $L_M$ , that is not always available. *a posteriori* weighting (Bayesian weighting) is estimated from an *a priori* for each model such as :

$$p_{post}(M_i) = \frac{L_{M_i} p(M_i)}{\sum_j L_{M_j} p(M_j)}$$

In the absence of the likelihood, weighting is evaluated by experts and in the light of statistical data (fit testing, qualitative measure ...).

To assess the error, bootstrap simulations are performed. The idea is to simulate a distribution of the variable of interest projected to a one-year horizon (to be consistent with the regulatory framework). Each observation of the distribution is simulated by one of the models that has been previously drawn using weights. Regulatory capital  $\mathcal{R}_{Bootstrap}(M)$  is evaluated on that distribution. The model error in this approach can be negative and is defined by :

$$\rho(M) = \frac{\mathcal{R}_{Bootstrap}(M)}{\mathcal{R}(M^*)} - 1$$

The Bayesian approach gives better accuracy in the calculation of the error. Therefore, the capital required to account for this error is smaller. The difficulty mainly lies in the computation of the likelihood and the weights.

## Main results

Both approaches have been implemented equity risk and non-life risks. The risks related to equity is a significant reduction in British and American daily stock prices. The model error is estimated on the predicted distribution of prices at the end of 2014 based on the data from 2001 to 2013.

Non-life claims are related to marine transport and motor. We therefore work on two run-off triangles of paid claims over the period 1999-2013.

## Study of a single risk

Firstly, the two approaches are applied to one-dimensional models. We evaluate the error of the asset model on UK data and the error of reserving models on marine transport data.

## Asset models

The first part of the study consists of choosing the models to predict the equity price. Next, we performed validation of each of their hypotheses to check their suitability. We kept four models to model assets : the Black-Scholes model, Merton model, Garch model and the regime-switching Black-Scholes.

In the light of the validation tests, the Garch model seems most appropriate and was chosen as the referent model. The Black-Scholes model has a good fit despite the fact that its limitation to two regimes (for simplicity) is probably too restrictive. The other models are not very suitable for our study. We decide to keep these models since they are often used in practice despite the criticism against them.

The model error was first evaluated by the Bayesian approach. Bayesian weights could be assessed because we were able to measure the likelihood of each model. The choice of *a priori* distribution is not obvious. We therefore chose to give a slightly higher weight to the Garch model. In addition, we performed a sensitivity analysis on the *a priori* to test the impact of more extreme weights (for example, by giving no weight to the Black-Scholes model that seems less suitable). In all cases, the *a posteriori* distribution gives a lot of weight to the Garch model, lower weight to the regime-switching Black-Scholes model and insignificant weight to other models.

Ultimately, the model error is very similar regardless of the selected *a priori* and is negative. This is explained by the fact that the Garch model is the most prudent ; the regulatory capital required with this model is higher than with the bootstrap. A negative model error will not involve a resumption of regulatory capital for prudential issues, the error is assumed to be zero.

Model risk was then assessed by the worst case approach. The error is zero in this case since the model Garch is more cautious.

## Reserving models

Regarding reserving models, we selected three model classes to simulate the distribution of estimated reserves : the bootstrap Mack class, the bootstrap ODP class and MCMC GLM class. Each class contains several models that differ in terms of error distribution (for example, the bootstrap Mack error can be modeled by Gamma distribution or by resampling ...). The model assumptions are validated for the majority. We keep each class and each of their models. The reference model is 'log-normal' the model. This model simulates the distribution of reserves by lognormal distribution whose parameters are estimated from the best estimate and the variance of the Mack model (before bootstrap). The error is first estimated intraclass, ie between models belonging to the same class. The evaluation of the extra-class error is established by weighting each class and by aggregating the errors of each class (intra-class error) using these weights. The weights are chosen equiprobable since we are not able to differentiate the classes in terms of adequacy.

Except for the MCMC GLM models, model likelihood is not assessable. Furthermore, the lack of data (due to the aggregation of claims in run-off triangles) makes the assessment of models more difficult. Weighting was chosen equiprobable within a class of models (except for class MCMC GLM). Indeed, the results were quite similar for each model and we do not have any *a priori* on models adequacy. For the MCMC GLM class, we use the Bayesian approach using likelihood. *a posteriori* weighting almost puts the entire weight on the GLM model with Poisson error. While this model seems better, we prefer to give it only 2/3 of the entire weight and the rest to the GLM model with Gamma error in order to be cautious since the GLM model with Gamma error is the most prudent.

The overall error is 1.94% by weighting approach and 2.94 % worst case approach. It is therefore not very high and may be considered negligible. However, the results are highly variable depending on the weighting of classes and that of the intra-class models. The maximum error is 6.9% if the MCMC GLM class has all the weight and the error intraclass is evaluated by the worst case approach. The maximum error is significant (it exceeds 5 %) but comes from models that do not seem adequate.

## Dependency study

Dependence was modeled by known copulas (Gaussian, Student, Clayton, Gumbel and Frank) to take into account non-linear dependencies. As for the margins models (asset or reserving models specified before), copulas and their parameterization were validated. To evaluate the error on dependency models, we proceeded in two stages. First, we evaluated the error of pure dependence, ie the error due to the choice of the copula. To do this we set the margins (modeled by the reference model) and only vary the copulas. The copula error is evaluated by both approaches (worst case and Bayesian). For the weighting approach, the weights are evaluated using the likelihood when we can evaluate it and from qualitative information otherwise. Secondly, we assessed the total dependency error including the error related to the choice of the copula and that related to the choice of the margin model. Reference models to simulate margins are the same as in the case of a single risk.

### *Asset models*

For asset models, the reference copula is the Student copula with 12 degrees of freedom. It is the copula with the best fit. Since data are sufficient to assess the copulas likelihood, the weighting approach is a Bayesian approach.

On our data, the error of pure dependence is not significant in the Bayesian approach but is relatively high in the worst case approach. Indeed, the Frank copula is the most conservative but also the least likely. The weighting approach is justified because it hardly gives weight to this copula, so the error is more realistic.

Concerning the global dependency, the results are quite similar with an error of 6% in the worst case approach but negative (-1.6%) in the Bayesian approach. The explanation for these results is the same. The weighting approach seems more consistent, which is why the model error is considered to be nil for this study.

### *Reserving models*

For reserving models, the reference copula is the survival Clayton copula (copula rotated 180 degrees). The copula's parameterization is derived mainly from expert judgement due to lack of data. Copulas likelihood can not be calculated for the same reason. Consequently, weights are equiprobable or promote survival Clayton copula. In both cases, the results are identical.

On our data, the error of pure dependence is not significant whatever the approach considered.

The overall dependency error is very close to the error obtained in the case of a single risk. It is therefore quite high especially when the weight is large on the MCMC GLM models.

## Conclusions et Limitations

The model error is globally insignificant in our application. Only the error resulting from reserving models is positive and significant for some models. It is also difficult to assess the adequacy of these models due to the lack of data. Although GLM do not always seem right, we nevertheless prefer to keep them to avoid underestimating the error since we do not have any quantitative arguments to invalidate them.

This study does not allow generalization of the results to all segments especially for non-life risks where some segments may be much less stable than those studied.

Both approaches have limits, in particular the cost of implementation of each model and the calculation time. Furthermore, they both make very strong assumptions about the choice of the model set and the reference model. By changing the reference model, the error can be greatly increased (eg 16% error in the worst case approach when taking the Frank copula as a reference for asset dependency models). Furthermore, we assume that the 'true' model belongs to the set, this hypothesis is unverifiable in practice.

The worst case approach is too conservative since it does not take into account the likelihood of the models. Thus, the error in the worst case scenario could be biased by taking too many models in the set, including less suitable models. However, it gives an idea of the reference model's prudence over other models.

The weighting approach is more realistic but the difficulty arising from this approach is the choice of weights in particular in cases where the likelihoods are not available. The results can vary greatly depending on weightings.

In order to be able to generalize this study, it is necessary to apply it to other segments. Indeed, we need to check whether the results are comparable. In addition, further work on the choice of weights should be performed in the case where the likelihoods are not measurable. For example, expert judgement on models could be requested and aggregated. Finally, we could increase the portfolio size to evaluate the error on models aggregating more risks. With a more complex model, we would get closer to the concrete framework of the internal model.

## Remerciements

Tout d'abord, je tiens à remercier la SCOR pour m'avoir offert la possibilité de réaliser mon mémoire de fin d'études dans les meilleures conditions et dans un très bon environnement de travail.

Je remercie toute l'équipe du Département Modélisation Actuarielle pour m'avoir guidée tout au long de ce stage et aidée à l'élaboration de ce mémoire. En premier lieu, je suis très reconnaissante envers ma tutrice, Ecaterina NISIPASU, pour avoir permis la réussite de ce travail. Je tiens également à remercier, tout particulièrement, Théophile DOMERGUE, Mihaela TOPUZU MEDECIN, Laila ELBAHTOURI et Andres CUBEROS pour s'être rendus disponible et avoir répondu à mes interrogations.

Plus généralement, je remercie les équipes de la Direction Actuariat Groupe pour m'avoir accueillie et pour les nombreux repas partagés.

Je tiens aussi spécialement à remercier les stagiaires et apprentis de la SCOR pour leur soutien au quotidien et tous les moments partagés ensemble pendant ces quelques mois. Je pense, en particulier, à Clémence ABRAHAMIAN, Morgane BARDON, Johan FAUCHEUX, Cyrine GARNAOUI, Sara MARZAC et Julie PLOUTON.

Je veux également remercier les enseignants de l'ENSAE pour leur accompagnement durant ce stage mais aussi durant ma scolarité.

Je remercie énormément Maxence PHILIPPE et Amy WRIGHT qui ont gentiment accepté de lire et corriger des parties de mon mémoire, je leur en suis extrêmement reconnaissante.

Pour finir, mes remerciements vont à mes parents, à mon frère et à ma marraine ainsi qu'à tous mes amis et ma famille qui m'ont soutenue, encouragée et permis d'en arriver là.

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>I Présentation de l'étude</b>	<b>2</b>
<b>1 Contexte et cadre réglementaire</b>	<b>2</b>
1.1 La réassurance et l'entreprise SCOR . . . . .	2
1.1.1 La réassurance . . . . .	2
1.1.2 La SCOR . . . . .	3
1.2 La réglementation et le dispositif Sovalbilité II . . . . .	3
<b>2 Le risque lié à l'erreur de modèle</b>	<b>5</b>
2.1 Présentation . . . . .	5
2.1.1 Définitions . . . . .	5
2.1.2 Cartographie des risques . . . . .	6
2.2 Le risque de modèle dans la réglementation . . . . .	10
2.2.1 Les crises financières et le risque de modèle . . . . .	10
2.2.2 La réglementation bancaire et assurantielle . . . . .	11
<b>II Méthodologies</b>	<b>12</b>
<b>3 Quantification de l'erreur de modèle</b>	<b>12</b>
3.1 Cadre théorique . . . . .	12
3.1.1 Le modèle . . . . .	12
3.1.2 Mesures de risque . . . . .	12
3.2 L'approche <i>worst case</i> . . . . .	14
3.2.1 Description . . . . .	14
3.2.2 Critiques de l'approche . . . . .	15
3.3 L'approche bayésienne . . . . .	15
3.3.1 Principe . . . . .	15
3.3.2 Critiques de l'approche . . . . .	17
<b>4 L'erreur de modèle sur des structures de dépendance</b>	<b>19</b>
4.1 Rappels et copules classiques . . . . .	19
4.2 Estimation du paramètre de copule . . . . .	21
4.3 Proposition de méthodologie de quantification de l'erreur de modèle . . . . .	22
<b>III Etude de modèles d'actifs</b>	<b>24</b>
<b>5 Application à un risque isolé</b>	<b>24</b>
5.1 Présentation des modèles et des données . . . . .	24
5.1.1 Les données . . . . .	24
5.1.2 Les modèles . . . . .	25
5.2 Validation des Modèles . . . . .	31
5.2.1 Vérification des hypothèses . . . . .	31
5.2.2 Probability Integral Transform (PIT) . . . . .	36
5.2.3 Procédure . . . . .	40
5.2.4 Résultats . . . . .	41
5.2.5 Tests de sensibilité . . . . .	43
5.3 Comparaison des approches Bayésienne et <i>worst case</i> . . . . .	44

5.3.1	Approche <i>worst case</i> . . . . .	44
5.3.2	Analyse comparative . . . . .	45
5.4	Etude comparative sur rendements journaliers . . . . .	45
5.5	Conclusion . . . . .	46
<b>6</b>	<b>Application à des structures de dépendance</b>	<b>48</b>
6.1	Choix de la copule et ses paramètres . . . . .	48
6.1.1	Choix de la copule . . . . .	48
6.1.2	Choix des paramètres de copule . . . . .	49
6.2	L'erreur de dépendance pure . . . . .	51
6.2.1	Résultats sur données mensuelles . . . . .	51
6.2.2	Comparaison avec les résultats sur rendements journaliers . . . . .	55
6.3	L'erreur de modèle d'actifs . . . . .	56
6.4	Conclusion . . . . .	57
<b>7</b>	<b>Analyse et interprétations des résultats sur les modèles d'actifs</b>	<b>58</b>
<b>IV</b>	<b>Etude des modèles de provisionnement non-vie</b>	<b>59</b>
<b>8</b>	<b>Application à un risque isolé</b>	<b>59</b>
8.1	Données, problématiques et modèles . . . . .	59
8.1.1	Les données . . . . .	59
8.1.2	Notation et introduction à la problématique du provisionnement . . . . .	60
8.1.3	Le modèle de Chain-Ladder . . . . .	60
8.1.4	Le modèle de Mack . . . . .	63
8.1.5	Les Modèles Linéaires Généralisés (GLM) . . . . .	64
8.2	Validation des modèles . . . . .	67
8.2.1	Chain-Ladder et Mack . . . . .	67
8.2.2	GLM . . . . .	70
8.3	Evaluation de l'erreur de modèle . . . . .	74
8.3.1	Modèle de référence . . . . .	75
8.3.2	L'approche mixte . . . . .	76
8.3.3	L'approche par pondération . . . . .	77
8.4	Conclusion . . . . .	79
<b>9</b>	<b>Application à des structures de dépendance</b>	<b>80</b>
9.1	La non adéquation du modèle GLM avec erreur gamma . . . . .	80
9.2	Choix de la copule et ses paramètres . . . . .	81
9.2.1	Choix de la copule . . . . .	81
9.2.2	Choix des paramètres de copule . . . . .	81
9.3	Erreur de dépendance pure . . . . .	84
9.4	Erreur de dépendance globale . . . . .	85
9.5	Conclusion . . . . .	87
<b>10</b>	<b>Analyse et interprétations des résultats sur les modèles de provisionnement</b>	<b>88</b>
	<b>Conclusion</b>	<b>89</b>
	<b>Bibliographie</b>	<b>91</b>
	<b>Appendices</b>	<b>95</b>

# Introduction

Les sociétés d'assurance et de réassurance sont soumises à un cadre réglementaire afin de garantir l'intérêt de leurs clients. La réglementation Solvabilité II, qui entre en vigueur en Janvier 2016, impose aux compagnies de détenir un niveau de fonds propres suffisant pour faire face aux risques auxquels elles s'exposent. D'un point de vue quantitatif, un assureur ou un réassureur doit démontrer que le niveau de ses fonds propres couvre son exigence en capital, le SCR (solvency capital requirement). En d'autres termes, les fonds propres détenus doivent permettre de limiter la probabilité de ruine à un an à 0.5%, ce qui correspond à l'occurrence d'un évènement bicentenaire.

Le régulateur laisse le choix de la méthode retenue pour le calcul du SCR. Cette étude ne s'intéresse qu'au cas de l'adoption d'un modèle interne, élaboré par l'assureur (ou le réassureur) et approuvé par le régulateur, pour l'évaluation du SCR. Pour le développement de son modèle interne, la compagnie doit être en mesure d'identifier et de modéliser son profil de risque. Un risque, peu étudié jusqu'à présent, mais sur lequel le régulateur demande de porter une attention particulière, est le risque de modèle. Ce risque provient de l'utilisation d'un modèle de prévision erroné ou mal spécifié. Aucun modèle n'est parfait puisqu'il existe toujours des incertitudes quant aux évolutions futures des paramètres du modèle. L'erreur de modèle est donc la différence entre la valeur prédite issue du 'vrai' modèle, par définition inconnu, et la valeur prédite par le modèle utilisé en pratique. Nous présentons une cartographie détaillée des différentes sources d'erreur qui composent le risque de modèle. Cette étude se concentre ensuite uniquement sur l'erreur de spécification du modèle, i.e. le choix du modèle en lui-même et non les erreurs d'estimation ou de données par exemple.

Dans le cadre de l'élaboration d'un modèle interne, la compagnie se doit d'avoir un modèle qui valorise au mieux ses fonds propres requis afin de ne pas risquer de les sous-estimer ou surestimer. Une sous-estimation des fonds propres requis va donner l'impression que la compagnie est plus solvable qu'elle ne l'est réellement et une surestimation requière le blocage de plus de capital que nécessaire, l'allocation du capital n'est alors pas optimale. L'évaluation de l'erreur de modèle s'avère donc importante dans le processus de gestion des risques.

L'enjeu de ce mémoire est d'évaluer l'erreur de modèle sur des variables d'intérêt vues individuellement dans un premier temps et agrégées dans un second temps. L'objectif est, d'une part, de reprendre et d'améliorer les travaux effectués dans deux mémoires récents d'actuariat (cf [15] et [27]) et d'autre part, de continuer ces travaux sur des modèles d'agrégation de deux risques. Deux approches d'évaluation ont été proposées, l'approche du pire des cas (*worst case*) et l'approche par pondération (notamment la pondération bayésienne).

Dans la première partie, nous présentons et définissons l'erreur de modèle ainsi que la cartographie des risques sous-jacents.

Dans la seconde partie, nous exposons les approches d'évaluation de l'erreur de modèle et les mesures de risques associées. Nous développons la méthodologie mise en place dans les deux dernières parties applicatives.

Dans la troisième, nous appliquons les méthodes d'évaluation de l'erreur de modèle sur des modèles d'actifs et comparons les résultats obtenus dans les deux approches. Tout d'abord, un actif seul est modélisé. Nous introduisons ensuite une nouveauté dans l'évaluation de l'erreur de modèle, à savoir l'agrégation. Nous évaluons donc le risque lié à la dépendance de deux actifs, tout d'abord en ne nous intéressant qu'à l'erreur de dépendance pure (liée uniquement au choix de la copule) puis en évaluant l'erreur globale (liée au choix de la copule et des marginales).

Dans la quatrième partie, les mêmes méthodologies que dans la troisième partie sont appliquées sur des modèles de provisionnement non-vie.

## Partie I

# Présentation de l'étude

## 1 Contexte et cadre réglementaire

### 1.1 La réassurance et l'entreprise SCOR

Le risque est présent dans la quasi-totalité des activités humaines, dès qu'il y a un aléa, il y a nécessairement un risque. L'aversion au risque des agents a motivé le développement du secteur assurantiel. En effet, un certain nombre d'aléas peuvent venir grever significativement le patrimoine des agents économiques qui cherchent donc à diminuer leur risque. Le principe de l'assurance est de couvrir une partie des risques des agents dont le montant est aléatoire (fonction de la sinistralité future) en échange de prime dont le montant est fixé (par avance). Le cycle de production d'un assureur est donc inversé ; ce dernier cherche, par conséquent, à tarifier la prime au plus juste afin d'être en mesure de couvrir les sinistres futurs.

#### 1.1.1 La réassurance

**Définition 1.1.** La réassurance est une assurance secondaire. Les réassureurs ne traitent pas directement avec les agents mais avec les assureurs. Le réassureur est décrit dans le code de commerce allemand comme "l'assureur de l'assureur". L'assureur (ou "la cédante") cède une partie de son risque/aléa à un réassureur en contrepartie du versement d'une prime définie par le contrat de réassurance.

#### Bref historique de la réassurance

La réassurance est apparue en Italie, à la Renaissance pour la couverture des bateaux de commerce. La pratique de la réassurance était à l'époque uniquement spéculative et non basée sur des statistiques élaborées à l'aide de données comme c'est le cas de nos jours. En Angleterre, l'assurance et la réassurance se sont vraiment développées à la suite de l'incendie de 1666 qui a détruit une grande partie de Londres. La Lloyd's, un marché de l'assurance sur lequel les assureurs se réunissent pour mutualiser et répartir les risques, a été créé à la suite de cet événement par les marchands et les assureurs. Cependant, il faut attendre la révolution industrielle pour percevoir un réel changement dans l'activité assurantielle. L'expansion de la réassurance, en cette période de métamorphose du paysage économique, est expliquée par le besoin croissant d'assurance incendie. A la fin du 20ème siècle, la réassurance est devenue davantage une nécessité du fait de la mise en place d'un cadre réglementaire, de la taxation, du développement de nouveaux produits et nouvelles méthodes de tarification, et de la compétitivité du marché de l'assurance.

#### La réassurance, une valeur ajoutée

La réassurance n'est plus uniquement un moyen de gestion des risques mais aussi une réelle valeur ajoutée pour l'activité d'une assurance.

D'une part, la réassurance permet une diversification des risques et ainsi de diminuer le risque qui porte sur l'assureur. En effet, une trop grande concentration des risques peut entraîner de grandes pertes si un sinistre affecte plusieurs polices d'assurances simultanément. Elle permet également de couvrir des risques tels que les catastrophes qui correspondent soit à un événement rare ayant un coût très élevé (ex : les catastrophes naturelles) soit à un événement à forte fréquence mais faible gravité. Ainsi, la réassurance a plusieurs fonctions et notamment celle de garantir la stabilité financière de l'assureur. Cela lui permet effectivement d'améliorer sa gestion actif/passif (allocation du capital, gestion de portefeuille), de limiter les périodes d'exposition et les variations dans les paiements des sinistres. C'est aussi un moyen de faciliter les fusions et d'augmenter la taille du marché, en assurant de nouveaux risques par exemple.

D'autre part, le réassureur, par sa position sur le marché et la taille de ses portefeuilles, a accès à un grand nombre d'informations et possède une grande expertise qu'il met à profit des cédantes. Les données collectées sont des outils précieux pour les analyses comparatives, les analyses des tendances ou le développement de nouveaux produits.

### 1.1.2 La SCOR

SCOR SE (Société COmmerciale de Réassurance) est un groupe français de réassurance créé en 1970 et ayant le statut de Société Européenne. C'est l'un des principaux réassureurs mondiaux et la seule société de réassurance ayant vu sa note, relative à sa solidité financière, relevée cette année par deux agences de notation (S&P et Fitch) confirmant ainsi son statut de réassureur global de premier rang.

La SCOR se divise en trois composantes : les activités commerciales séparées entre réassurance dommages et réassurance vie et l'activité de gestion d'actifs. Ce mémoire s'est déroulé dans le Département Modélisation Actuarielle, un des trois départements de la Direction Actuariat Groupe composé également des Départements de Provisionnement Vie et non-vie. La SCOR a choisi, dans le cadre de la réglementation Solvabilité II, de mettre en place un modèle interne (voir section suivante) ; l'équipe du Département de Modélisation Actuarielle est notamment en charge de la validation annuelle de ce modèle. Par ailleurs, ses principales missions sont de mener des études permettant l'implémentation des nouvelles méthodologies ou normes réglementaires. C'est dans ce cadre que s'inscrit ce mémoire en tentant de répondre à une nouvelle problématique émanant du régulateur : l'évaluation du risque de modèle.

## 1.2 La réglementation et le dispositif Solvabilité II

La réglementation Solvabilité II, qui entre en vigueur en Janvier 2016, repose sur trois piliers formulant chacun des exigences. Le pilier I se concentre sur les exigences quantitatives, "c'est-à-dire les règles de valorisation des actifs et des passifs, ainsi que les exigences de capital et leur mode de calcul." <sup>1</sup>. Le pilier II regroupe les exigences qualitatives et le pilier III, les exigences de transparence et de diffusion de l'information. Cette étude s'intéresse essentiellement à l'aspect quantitatif imposé par le pilier I.

Le pilier I se traduit principalement par une exigence de fonds propres, appelée SCR (Solvency Capital Requirement). L'assureur doit être en mesure de démontrer que les fonds propres qu'il détient sont suffisants pour couvrir ses différents risques. Plus formellement, le SCR correspond au montant de capital minimum dont doit disposer l'assureur pour faire face à ses engagements (absorber les pertes potentielles) à horizon 1 an avec une probabilité de 99.5%.

Selon la réglementation, le SCR peut être évalué par une formule standard ou un modèle interne complet ou partiel (la formule standard est utilisée pour couvrir certains risques). Dans tous les cas, le SCR doit couvrir quasiment l'ensemble des risques auxquels la compagnie est exposée que ce soit les risques Vie, non-vie, de marché, opérationnels... Les risques n'étant pas couverts par le SCR comme le risque de réputation et les risques stratégiques sont pris en compte, par ailleurs, dans la gestion des risques de l'entreprise et plus particulièrement dans l'ORSA (Own Risk and Solvency Assessment). Bien que plus coûteux en ressources pour la mise en place et la veille, le modèle interne permet à l'entreprise d'évaluer son exigence en capital en fonction de son profil de risque et ainsi d'en tirer un bénéfice, en permettant, par exemple, une meilleure gestion de ses risques et une optimisation de l'allocation de son capital. Comme précisé précédemment, la SCOR a choisi de mettre en place un modèle interne pour le calcul du SCR, ainsi nous nous focaliserons uniquement sur ce mode de calcul de l'exigence en capital dans la suite de l'étude.

1. Disponible sur : <http://acpr.banque-france.fr/solvabilite2/pilier-1-les-exigences-quantitatives/les-exigences-quantitatives.html>

C'est dans ce contexte réglementaire que nous nous intéressons à un risque peu étudié à présent, le risque de modèle. Le régulateur français demande, aux compagnies ayant développé leur modèle interne, de mesurer ce risque et son impact sur le SCR. L'idée sous-jacente est que l'utilisation d'un modèle erroné, notamment si ce dernier sous-estime son SCR, pourrait impacter significativement le capital de la compagnie. Une erreur de modèle positive et importante pourrait rendre l'assureur ou le réassureur incapable de faire face à ses engagements. Une erreur significativement négative pourrait à l'inverse rendre l'assureur ou le réassureur moins compétitif sur le marché du fait d'une surestimation de l'exigence en capital. Ainsi, il est nécessaire de développer des méthodes afin d'être en mesure d'évaluer l'erreur de modèle.

## 2 Le risque lié à l'erreur de modèle

*"A des moments, nous pouvons perdre de vue le but ultime d'un modèle quand les mathématiques deviennent trop intéressantes. Les mathématiques des modèles financiers peuvent être appliquées précisément, mais les modèles ne sont pas du tout précis dans leur application au monde réel. Leur fidélité en tant qu'approximation de la réalité varie significativement en fonction du moment et du lieu."*  
 R.Merton, discours lors de la réception du prix Nobel d'économie en 1997.

### 2.1 Présentation

Un modèle est une méthode quantitative, un système ou une approche qui applique les théories, techniques et hypothèses mathématiques, statistiques, économiques et financières pour transformer des données en entrée en estimations quantitatives. Un modèle mathématique est un outil d'aide à la décision qui est construit dans un but précis (prédiction d'un cours, description de données...). Un modèle n'est jamais parfaitement réaliste. Le choix du modèle et des paramètres afférents peuvent amener à des résultats différents bien que modélisant la même quantité. Dans cette section, nous tenterons de définir le risque de modèle et ses sources.

#### 2.1.1 Définitions

Il y a plusieurs sens à l'expression "risque de modèle". Communément, on définit le risque de modèle comme la différence de prix que procure plusieurs modèles de valorisation d'un instrument ayant les mêmes données en entrée (hypothèses sur la quantité d'intérêt, données historiques...). En effet, nous supposons qu'il existe un unique "vrai modèle" et l'erreur de modèle vient du risque auquel est exposé, par exemple, un trader en utilisant un modèle erroné ou mal spécifié.

Il existe deux approches différentes pour traiter du problème de la validation d'un modèle :

#### L'approche par valeur

La première, appelée l'approche par valeur, est proposée par Derman [12]. Derman définit le risque de modèle comme "le risque que le modèle ne soit pas une description réaliste (ou au moins plausible) des facteurs qui affectent la **valeur** du dérivé". Il est le premier à évoquer le terme de risque de modèle dans le domaine de la finance. Pour lui, il est évident "qu'un modèle n'est toujours qu'une tentative de simplification" d'une réalité bien plus complexe. En partant de cette constatation, il est clair qu'il n'existe pas de modèles parfaitement réalistes. De plus, dans sa définition, seule la modélisation des facteurs qui impactent la valeur de l'instrument importe. Selon lui, l'erreur de modèle peut donc provenir de différentes sources toutes relatives aux facteurs utilisés dans l'évaluation de l'instrument ; les sources principales étant les erreurs dans l'approximation de variables stochastiques par des variables déterministes, les erreurs sur les hypothèses de dépendance ou de dynamique des variables.

#### L'approche par prix

La deuxième approche, appelée l'approche par le prix, prend le problème d'une toute autre perspective. Ainsi, Rebonato [38] définit le risque de modèle comme "le risque d'occurrence à un instant donné d'une différence significative entre la valeur Mark-to-Model d'un instrument complexe et/ou illiquide détenu dans le bilan ou hors bilan d'une entreprise financière et le **prix** auquel le même instrument sera vendu sur le marché". Cette définition reste assez restreinte et ne s'intéresse qu'aux risques liés à la modélisation d'un instrument financier et non à la modélisation d'un portefeuille. Cette approche s'oppose à la précédente puisqu'elle suppose que les pertes liées à l'erreur de modèle ne sont pas issues "d'un écart entre la valeur donnée par le modèle et la vraie valeur de l'instrument" mais plutôt d'un écart entre la valeur donnée par le modèle et le prix de marché de l'instrument. Cela suppose donc que les pertes proviennent du processus d'évaluation du prix Mark-to-Market et que la vraie valeur est la valeur de marché, ce qui est cas dans le cadre de référence de Solvabilité II.

D'une manière plus générale, Crouhy et al.[10] définissent le risque de modèle comme étant "tous risques induits par l'application d'un modèle statistique". Cette définition concerne aussi bien les modèles stochastiques que les modèles économétriques et comprend toute les sources du risque. Le risque de modèle est aussi défini de manière plus restreinte, en excluant le risque opérationnel, comme "tous risques induits par le choix, la spécification et l'estimation d'un modèle statistique".

Nous retiendrons la définition très générale de Kherkof, le risque de modèle est l'aléa engendré par le fait de "travailler avec un modèle potentiellement incorrect".

### 2.1.2 Cartographie des risques

Wang et al. [42] ont repris la classification des différentes sources de risque de modèle présentée dans différents articles et notamment celui de Crouhy.

Crouhy définit 6 types de risque de modèle :

- Le modèle est hors sujet
- Le modèle est incorrect (mal spécifié)
- L'implémentation du modèle est erronée
- Le modèle est mal calibré
- Les données sont mal traitées
- La mise en application est inadaptée

Nous allons, par la suite, développer chacune de ces sources de risques et tenter d'élaborer une cartographie des sources du modèle pour avoir une vision plus claire et synthétique du risque de modèle.

#### Modèle sorti du contexte

Dans ce cas le modèle ne répond tout simplement pas à la problématique de départ. Cela peut se produire si le modèle est basé sur des données n'étant pas appropriées ou sur des théories économiques sans fondement. Nous considérerons que ce cas est relativement marginal puisqu'en pratique, dans les grandes institutions, les modèles font l'objet de revues, de back-testing et d'audit. Par conséquent, cette source de risque sera négligée.

#### Modèle incorrect

##### **Solutions analytiques fausses**

La solution analytique résultant du modèle est fausse. Un exemple clé est la recherche de maximum global qui peut dans certain cas ne résulter qu'en un maximum local. Ce risque appartient au risque de spécification du modèle.

##### **Hypothèses sur le processus stochastique fausses**

Dans le cas où le modèle fait appel à des processus stochastiques, certaines hypothèses peuvent être faites à tort, par exemple, en faisant des approximations par un processus déterministe. En pratique, il est difficile de modéliser un processus stochastique car nous ne pouvons nous fier qu'aux données observées pour émettre des hypothèses. Ce risque sera considéré comme le risque stochastique par la suite.

##### **Mauvais choix de variables**

Ce risque peut survenir dans des modèles à multiples variables lorsque les variables sont malencontreusement échangées, mal identifiées ou manquantes. Ce risque appartient au risque de spécification du modèle puisqu'il est lié aux variables afférentes au modèle.

### **Mauvaises hypothèses sur la distribution des paramètres**

Si la distribution des paramètres est mal choisie, cela peut conduire à de grandes déviations des résultats. Par exemple, si on fait l'hypothèse qu'un paramètre est gaussien alors que ce dernier possède une distribution à queue épaisse. Ce risque appartient au risque de spécification.

### **Mauvaises hypothèses sur la nature des marchés**

Les marchés sont très influencés par l'environnement économique, géographique et politique. Tous ces facteurs doivent être pris en compte de la manière la plus plausible possible afin de se rapprocher de la réalité. Cependant, certains marchés sont très complexes à modéliser et cela peut entraîner une nouvelle fois une erreur de spécification.

### **La non prise en compte des coûts de trading**

Les coûts de trading sont bien souvent ignorés des modèles. Cependant, ces coûts fluctuent et peuvent être modélisés par un processus stochastique. Par conséquent, dans des cas d'extrême fluctuation, la non prise en compte de ces coûts entraînent une erreur majeure du modèle. Cette erreur est liée à des hypothèses sur le modèle et est donc rattachée aux erreurs de spécification.

### **Mauvaises hypothèses sur les sous-jacents des actifs**

La plus grande erreur à cet égard repose dans la mauvaise prise en compte des dépendances entre les actifs. Ceci est aussi une erreur de spécification.

## **L'implémentation du modèle erronée**

Le développement croissant de modèles de plus en plus complexes a rendu nécessaire l'amélioration de l'implémentation de ces derniers. Plusieurs critères sont indispensables dans l'implémentation des modèles comme le temps de déroulement de l'algorithme, sa stabilité et sa clarté afin qu'elle soit compréhensible et vérifiable par d'autres collaborateurs.

### **Méthodes numériques inappropriées**

Le choix de la méthode numérique peut être inadapté et entraîner des erreurs d'approximation exacerbées en période de crise. Cette possible erreur doit être prise en compte en amont de l'implémentation car les algorithmes étant généralement longs, il est souvent difficile de trouver et traiter l'erreur *a posteriori*. Cette erreur est une erreur d'implémentation.

### **Nombre de simulations insuffisant**

Le nombre de simulations doit être correctement choisi. En effet, si nous effectuons beaucoup de simulations, l'utilisation du modèle peut être fortement ralentie. Cependant, beaucoup d'algorithmes, en particulier ceux utilisant les méthodes de Monte-Carlo ou les méthodes de Bootstrap, requièrent la réalisation d'un nombre important de simulations pour assurer leur convergence. Cette erreur est une erreur d'implémentation.

### **Choix des données**

Le choix des données peut s'avérer assez complexe dans la mesure où le nombre de données disponibles croît énormément et que le choix du type de données, de leurs sources et de leur quantité n'est pas évident. Par exemple, les données utilisées proviennent parfois de plusieurs sources et ne sont par conséquent pas identiques. De plus, la prolifération des données a encouragé l'utilisation d'historiques beaucoup plus longs pour la calibration des modèles bien que cela entraîne des erreurs (données trop anciennes et incohérentes, période de crise mal gérée...). Cette erreur est une erreur de données.

## **Le modèle mal calibré**

Cette source d'erreur provient de mauvais choix de méthodes statistiques pour les estimations.

### **Choix des méthodes statistiques utilisées**

Il existe une quantité non négligeable de méthodes statistiques pouvant donner des estimations différentes et le choix de l'une par rapport à l'autre peut être assez subjectif.

### **Incertitude des estimations**

Les méthodes statistiques engendrent des erreurs d'estimation dépendant des données d'entrée qui sont parfois elles même estimées. La source d'erreur est à la fois dans les données et l'estimation.

### **Gestion des données aberrantes**

Le retraitement des données est indispensable afin d'avoir les données les plus "propres" possibles. Il faut s'assurer que les données aberrantes soient supprimées et qu'il n'y ait pas de données manquantes ou tronquées pouvant engendrer un biais d'estimation. Ces erreurs sont des erreurs de données qui peuvent être liées à une erreur purement "humaine" ou opérationnelle lorsqu'il s'agit d'erreurs de frappe par exemple.

### **Calibration des modèles**

Chaque modèle doit être régulièrement vérifié, calibré, réévalué et ajusté. Il faut alors définir à quelle fréquence ces vérifications doivent avoir lieu et s'assurer que les hypothèses de départ en particulier concernant les dynamiques des paramètres sont toujours valables. Cette erreur sera qualifiée d'erreur de maintenance des modèles.

### **Jugements subjectifs**

Les modèles sont souvent basés à la fois sur un historique de données et sur des considérations plus subjectives issues d'expériences passées et d'avis d'experts. Les deux entrées présentent des avantages bien qu'il faut être précautionneux dans leur utilisation.

### **Mauvais traitement des données**

La qualité des modèles est intimement liée à la qualité de ces données d'entrée.

### **Horizon de temps et fréquence des données**

Comme précisé précédemment, la taille de l'historique de données doit être mûrement réfléchi afin qu'en sortie, les résultats soient significatifs sans perdre en qualité de prédiction.

### **Erreur dans la base de données**

Les données étant issues de sources diverses, il est souvent difficile de vérifier la rectitude des bases fournies.

### **Mauvaise mise en application**

#### **Hypothèses passées plus valides**

Le modèle ayant fait ses preuves par le passé peut ne plus être justifié dans certains contextes en particulier si les hypothèses ne sont plus valides. Cette erreur est due à un problème d'utilisation du modèle comme pour le cas suivant.

#### **Réutilisation des modèles**

Un modèle ayant correctement fonctionné pour un certain actif peut être utilisé pour un actif similaire mais présentant quelques caractéristiques distinctes.

Nous présentons ci-dessous la cartographie du risque de modèle au vu de la description des différents risques que nous venons d'établir.

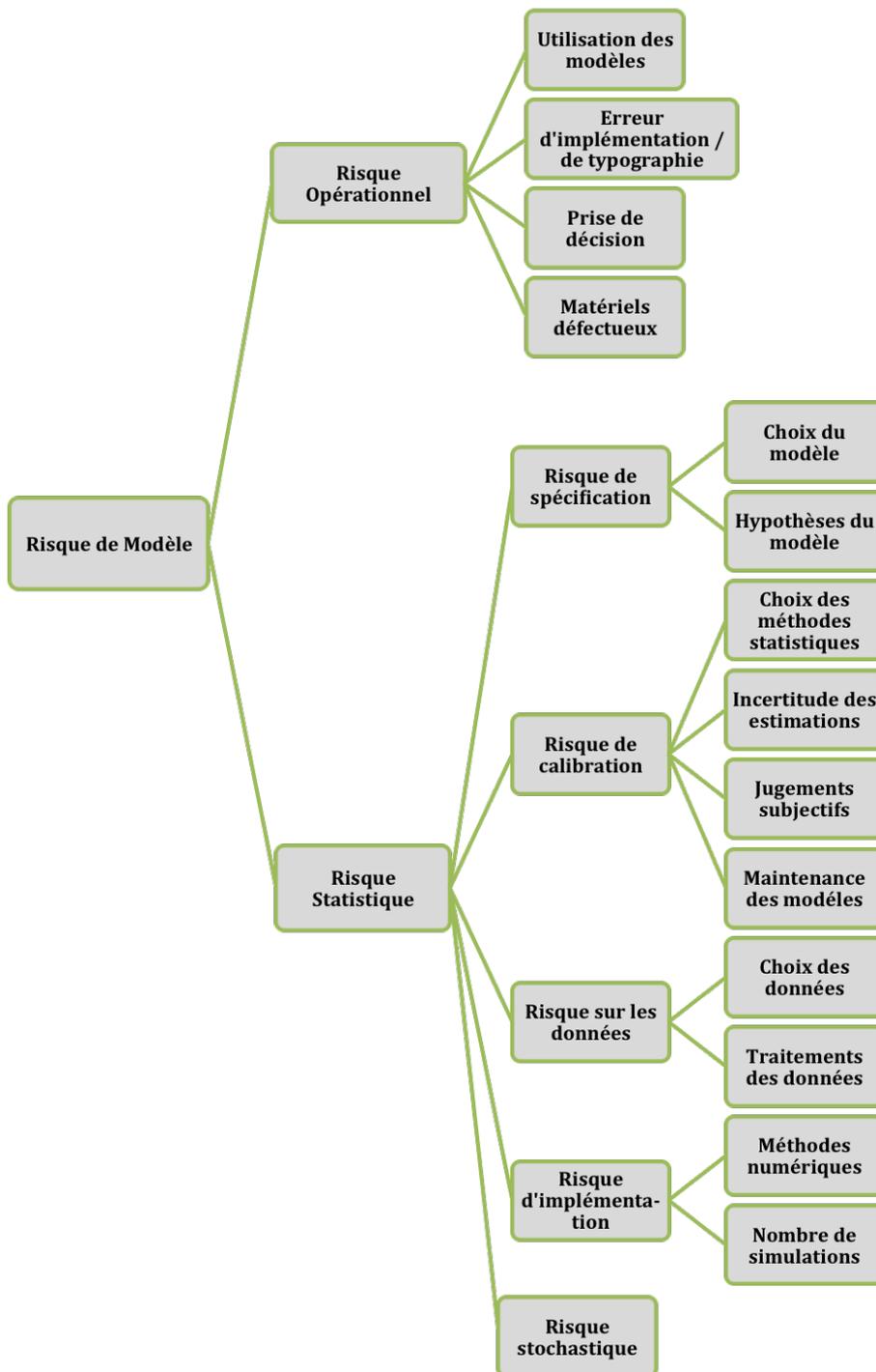


FIGURE 1 – Cartographie du risque de modèle

Tous les risques présentés interviennent dans le processus de création du modèle interne d'une entreprise d'assurance. Dans notre étude, nous nous concentrerons uniquement sur le risque de spécification du modèle et en particulier sur le choix du modèle par soucis de simplification. Cependant, parmi les autres risques définis, beaucoup sont déjà plus ou moins bien pris en compte par les politiques de gestion du risque, notamment le risque opérationnel. Nous considérons donc que ces risques sont déjà traités par ailleurs.

## 2.2 Le risque de modèle dans la réglementation

Le développement des modèles dans le domaine de la finance dans les années 70 avec notamment le modèle de Black-Scholes-Merton ainsi que leur complexification a engendré des questionnements quant aux risques que pouvaient générer ces modèles. Pour faire face à ce nouveau risque, les entreprises et particulièrement les banques ont mis en place des systèmes de gestion du risque de modèle. Nous détaillons par la suite l'historique de la mise en place d'une gestion de ce nouveau risque.

Il est important de noter que ce risque peut, au même titre que les autres risques financiers, entraîner des pertes financières conséquentes. Dans le contexte actuel où des modèles sont créés et utilisés pour la gestion et la prise de décisions stratégiques à chaque niveau des entreprises du secteur bancaire et assurantiel, il est indispensable de ne pas négliger ce risque.

### 2.2.1 Les crises financières et le risque de modèle

Le modèle de Black-Scholes de 1973 d'évaluation d'options marque le début de l'expansion des modèles financiers. Ce modèle toujours largement utilisé est une référence en matière de valorisation d'options aussi bien dans les milieux académiques que professionnels. Ce modèle est particulièrement apprécié de par sa simplicité qui lui vaut aussi d'être l'objet de beaucoup de critiques notamment de la part du mathématicien Mandelbrot [31]. Les hypothèses du modèle, à savoir que les rendements sont gaussiens et de volatilité constante, ainsi que les hypothèses sur le marché comme l'absence d'opportunités d'arbitrage, ne sont pas réalistes. Par conséquent, le choix de ce modèle pour valoriser une option dans le monde réel implique nécessairement une erreur.

Au cours du temps, les produits financiers se sont complexifiés et une gestion des risques à l'aide d'outils formels s'est imposée. En 1996, le comité de Bâle autorise l'utilisation de modèles pour couvrir le risque lié au marché mais n'évoque pas le risque de modèle. Cependant, beaucoup de grandes institutions ont commencé d'elles-mêmes à prendre en compte la possibilité d'erreurs de modèle.

Un des premiers exemples de mauvaise gestion des risques vient de la chute du fond d'investissement Long-Term Capital Management (LTCM) en 1998. Le fond, créé en 1994, dispose rapidement de positions très favorables et représentant un montant colossal pour l'époque. A la suite de la crise asiatique de 1997, le fond parie sur un retour à la normale des taux des obligations, or la crise se propage au marché russe et 1998 voit un nouveau choc se produire sur le marché obligataire. Au bord de la faillite, le fond est sauvé par les grandes banques mondiales afin d'éviter une potentielle crise systémique.

A l'origine de cette quasi-faillite, plusieurs hypothèses erronées utilisées par LTCM ont conduit à une sous-estimation du risque. Par exemple, dans le modèle de LTCM, la volatilité était supposée constante, les rendements étaient considérés comme gaussiens (les queues de distribution étaient donc sous-estimées), les périodes étudiées étaient très courtes et ne prenaient pas suffisamment en compte les épisodes passés de ralentissement économique.

Dans le début des années 2000, la gestion des risques prend un nouveau tournant dans les institutions financières en procédant à la validation des modèles par des entités indépendantes. Dans un rapport de 2000, the Office of the Comptroller of the Currency's (OCC)[36] donne des indications sur la gestion des risques et notamment considère que "le risque de modèle peut être considérablement réduit. Une modélisation solide inclut des procédures rigoureuses pour la validation du modèle.". Ces indications ont par la suite été reprises en 2004 par le comité de Bâle dans la réglementation Bâle II [11].

En réponse à une réglementation de plus en plus contraignante, les banques ont développé des instruments de crédit extrêmement complexes au travers de la titrisation (cf CDS...), pas toujours bien compris par les banques elles-mêmes. Les modèles développés pour la valorisation de ces titrisations et l'évaluation de leur risque donnaient des résultats dramatiquement différents à partir du moment où les hypothèses n'étaient pas complètement vérifiées. Ces modèles ont contribué à la crise des subprimes de 2007.

Le problème majeur était que les modèles étaient fondés sur la copule dite de Li[28] ou la copule gaussienne. Cette copule présente l'avantage d'être simple, bien plus que toutes les autres copules proposées à l'époque pour valoriser ces produits. Beaucoup d'académiciens ont rapidement remis en cause l'utilisation de cette copule qui suppose entre autre que les corrélations sont constantes. Dès 2009, un article de Felix Salmon[40] intitulé "The Formula That Killed Wall Street" et devenu mondialement connu revient sur l'impact néfaste de cette formule sur les marchés financiers et son implication dans la crise mondiale. Déjà en 2005, Li lui-même avait prévenu que "très peu de personnes comprenaient l'essence du modèle". Malgré la connaissance des inconvénients et des faiblesses des modèles basés sur la copule gaussienne pour la valorisation de ces produits, leur utilisation s'est répandue sur toute la place sans prendre de précautions particulières.

Ces exemples ne sont qu'une liste non exhaustive des cas répertoriés d'erreur de modèle par le passé. Une source importante dans l'erreur de modèle est l'erreur d'implémentation. Ces erreurs tendent à augmenter dans ces dernières années. On peut notamment penser à des exemples dans le domaine du trading hautes fréquences où les institutions se livrent à de véritables courses à la vitesse. Une erreur d'implémentation peut alors, en l'espace de millisecondes, avoir un impact sur l'ensemble du marché et entraîner des pertes se chiffrant en milliard. C'est ce qui s'est produit en 2010 avec le "Flash Crash" qui aura duré environ 35 minutes où un algorithme dit "spoofing" a perturbé les algorithmes des autres participants au marché en envoyant de faux ordres afin d'induire certaines réactions des marchés.

### 2.2.2 La réglementation bancaire et assurantielle

Il n'y a pas à l'heure actuelle de réglementation formelle concernant le risque de modèle. Les régulateurs appellent à la prudence, à la vérification et validation des modèles par des entités indépendantes et à la mise en place de gouvernance interne aux institutions. Bâle II a formalisé un certain nombre de critères devant être respectés lors de l'élaboration d'un modèle et incite à provisionner des provisions pour prendre en compte l'incertitude de modèle sans proposer de méthodes pour les calculer. Cependant, la réglementation n'impose pas de disposer d'un capital supplémentaire pour couvrir le risque de modèle.

La Réserve Fédérale américaine en collaboration avec l'OCC a émis en 2011 des conseils de surveillance sur la gestion du risque de modèle[35] assez similaires au cadre réglementaire imposé pour les autres grands risques tels que le risque de marché ou de crédit. Le risque est défini comme étant inhérent au modèle. De plus, la modélisation est considérée comme un processus ayant un cycle de vie dont les grandes étapes sont le développement, les tests, la validation, l'implémentation, le contrôle, la mise à jour et pour finir l'arrêt de l'utilisation du modèle. Toutes ces étapes doivent être accompagnées d'une gouvernance claire en terme de gestion des risques de l'entreprise définissant par exemple l'appétit pour le risque de l'entreprise.

Dans Solvabilité II, le Quantitative Impact Study[18] (QIS 5) fournit des conseils quant à la prise en compte du risque de modèle notamment dans le calcul du SCR et des provisions techniques. Comme dans le cadre réglementaire des institutions financières, aucune constitution de capital n'est imposée par Solvabilité II pour couvrir ce risque. Cependant, les obligations en matière de validation de modèles et de mise en place d'une gestion des risques faisant l'objet de contrôle garantissent une certaine fiabilité des modèles et donc une limitation de cette erreur. Un point important de la réforme Solvabilité II est le **principe de proportionnalité** qui est stipulé dans tous les traités de l'Union Européenne. Le principe de proportionnalité impose qu'une entreprise mette en place, pour le calcul du capital réglementaire et des provisions, des méthodes adaptées et proportionnées à la nature, à la taille et à la complexité des risques pris. En d'autres termes, "la transposition des directives doit faire l'objet d'une application au plus juste en intégrant les besoins et contraintes de l'entreprise"<sup>2</sup>. Si les méthodologies utilisées respectent ce principe, alors l'erreur de modèle peut être considérée comme insignifiante.

2. Disponible sur : <http://www.insurancespeaker-solucom.fr/2014/09/proportionnalite-dans-solvabilite-2/>

## Partie II

# Méthodologies

### 3 Quantification de l'erreur de modèle

#### 3.1 Cadre théorique

Dans cette section, nous présentons le modèle théorique de notre étude ainsi que la théorie mathématique afférente aux mesures de risque qui nous permettra par la suite de quantifier l'erreur de modèle.

##### 3.1.1 Le modèle

Dans les problèmes statistiques, nous cherchons à inférer ou prédire une quantité d'intérêt inconnue, notée  $y$ , sachant une quantité  $x$  connue qui représente les données historiques, le contexte et des avis d'experts plus subjectifs. L'objectif est donc de modéliser la probabilité suivante  $\mathbb{P}(y | x)$ , en d'autres termes, de déterminer, au vu de la variable  $x$ , l'incertitude sur  $y$ .

Spécifier cette probabilité peut s'avérer complexe, c'est pourquoi nous nous référons à des modèles afin d'approcher cette distribution de probabilité. Un modèle  $M$  permet de définir une relation hypothétique entre les variables  $x$  et  $y$ . Un modèle comprend donc toutes les hypothèses émises sur la variable  $y$  et est calibré à l'aide de l'historique de données.

Le modèle  $M_\theta$  est représenté par deux éléments,  $(M, \theta)$  où  $M$  correspond aux hypothèses sur la forme structurelle du modèle et  $\theta$  correspond aux paramètres du modèle spécifiques à la structure du modèle. Les paramètres sont estimés après spécification du modèle à partir de l'historique de données  $x$ . La variable d'intérêt, obtenue à partir du modèle  $(M, \hat{\theta})$ , est alors notée  $\hat{y}_{M, \hat{\theta}}$ .

Dans cette étude, nous chercherons à prédire plusieurs variables d'intérêt. Pour chaque variable, nous définirons plusieurs modèles qui nous paraissent cohérents pour la prédiction et l'inférence de cette variable d'intérêt au vu de l'historique de données, de fondements théoriques et des pratiques actuarielles. Comme nous l'avons expliqué précédemment, la définition de l'erreur de modèle est vaste et ce risque est issu de multiples sources. Nous nous concentrerons dans cette étude uniquement sur l'erreur de spécification du modèle, c'est à dire sur l'erreur commise lors du choix de modèle.

Plus précisément, pour chaque quantité d'intérêt considérée, il existe un modèle dit "de référence", noté  $M^*$  qui correspond au modèle utilisé en pratique pour modéliser ce risque sans prendre en compte une éventuelle erreur de modèle. Le "vrai modèle"  $M_{\theta_0}$ , avec  $\theta_0$  le vrai paramètre, correspond au modèle qui permet d'obtenir la vraie distribution inconnue de cette variable.

L'erreur de modèle sera alors évaluée à partir d'une classe de modèles  $M = (M_1, \dots, M_n)$ , appelée aussi ensemble de tolérance, où chaque modèle  $M_i$  est un "bon" candidat pour l'estimation de la variable d'intérêt.

Dans la partie suivante, nous présentons quelques rappels sur les mesures de risques.

##### 3.1.2 Mesures de risque

Cette section est partiellement tirée du livre d'A.Charpentier [8].

**Définition 3.1.** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité. Une mesure de risque  $\mathcal{R}$  est une fonction définie par :

$$\begin{aligned} \mathcal{R} : \Omega &\rightarrow \mathbb{R}_+ \\ X &\mapsto \mathcal{R}(X) \end{aligned}$$

C'est la fonction qui à chaque variable aléatoire  $X$  associe un nombre  $\mathcal{R}(X)$  quantifiant le niveau de danger inhérent au risque.

### Quelques propriétés des mesures de risque :

- **Invariance en loi** :  $X \stackrel{L}{=} Y \Rightarrow \mathcal{R}(X) = \mathcal{R}(Y)$
- **Croissance** :  $X \geq Y \Rightarrow \mathcal{R}(X) \geq \mathcal{R}(Y)$
- **Invariance par translation** :  $\forall k \in \mathbb{R}, \mathcal{R}(X + k) = \mathcal{R}(X) + k$
- **Homogénéité positive** :  $\forall \lambda \in \mathbb{R}_+, \mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$
- **Sous-additivité** :  $\mathcal{R}(X + Y) \geq \mathcal{R}(X) + \mathcal{R}(Y)$
- **Convexité** :  $\forall \beta \in [0, 1], \mathcal{R}(\beta X + (1 - \beta)Y) \geq \beta \mathcal{R}(X) + (1 - \beta) \mathcal{R}(Y)$

Dans le cadre des réglementations des systèmes bancaires et assurantiels, l'utilisation de mesures de risque s'est développée pour le calcul du capital réglementaire. Ces différentes propriétés sont souhaitables et naturelles pour les mesures de risque permettant de calculer ce capital. En effet, la propriété d'invariance en loi garantit que deux risques identiques sont couverts par le même niveau de capital. La propriété de croissance assure qu'un risque plus "dangereux" nécessitera pour le couvrir, une réserve de capital plus importante que pour un risque moins dangereux. La propriété d'invariance par translation implique que si on enlève un montant  $k$  d'une position pour l'investir sans risque alors le capital réglementaire sera réduit du même montant. L'homogénéité de la mesure de risque impose que le niveau de risque soit proportionnel à la taille du portefeuille. Pour finir, les propriétés de sous-additivité et de convexité garantissent que la diversification des risques implique une réduction du risque global et donc du capital réglementaire.

Pour notre étude, nous nous focaliserons sur la mesure de risque Value-at-Risk qui vérifie les quatre premières propriétés mais pas la propriété de sous-additivité.

**Définition 3.2.** La VaR correspond au niveau maximal de perte pour un niveau de confiance donné et un horizon de temps fixé. Soit  $X \in \Omega$  une variable aléatoire, un risque. La VaR au niveau  $\alpha$  de  $X$  est la fonction définie par :

$$VaR_\alpha(X) = \inf\{x \in \mathbb{R}, \mathbb{P}(X \leq x) \geq \alpha\}$$

Si  $X$  est une variable aléatoire continue alors  $VaR_\alpha(X) = q_\alpha(X)$ .

La VaR est une mesure développée dans les années 90 à la suite de plusieurs crises ayant fortement impacté les marchés. Elle a d'abord été popularisée dans le secteur bancaire avant d'être utilisée dans les milieux assurantiels. C'est la mesure retenue dans le cadre réglementaire Solvabilité II. Nous utiliserons donc cette mesure pour définir le montant de capital requis par la réglementation.

Dans la section précédente, nous avons introduit plusieurs notations. En gardant ces notations, nous notons  $\rho$  la mesure du risque lié au choix du modèle :

$$\rho : \mathbb{R} \rightarrow \mathbb{R} \\ \mathcal{R}(\hat{y}_{M, \hat{\theta}}) = \mathcal{R}(M) \mapsto \rho(M) = \rho(\mathcal{R}(\hat{y}_{M, \hat{\theta}}))$$

Le capital réglementaire total à constituer est alors :

$$\rho_{total} = \mathcal{R}(\hat{y}_{M^*, \hat{\theta}}) + \rho(M^*)$$

où

- $\rho_{total}$  : capital total pour faire face à tous les risques considérés
- $\mathcal{R}(\hat{y}_{M^*, \hat{\theta}})$  : capital réglementaire pour le modèle de référence  $M^*$
- $\rho(M^*) = \rho_{erreur}$  capital pour prendre en compte l'erreur de modèle.

Dans la suite, nous allons étudier différentes approches pour quantifier l'erreur de modèle  $\rho(M)$ . En particulier, nous détaillons les deux approches les plus développées dans la littérature ; à savoir l'approche dite "worst case" et l'approche dite Bayésienne.

## 3.2 L'approche *worst case*

### 3.2.1 Description

Cette approche consiste à mesurer l'erreur maximale commise en choisissant  $M^*$  comme modèle de référence par rapport à un ensemble de modèle  $M$  (contenant le modèle de référence et hypothétiquement le vrai modèle). Plusieurs mesures ont été proposées afin d'évaluer cette erreur. Dans cette section, les notations précédentes sont conservées.

Kerkhof et al [25] propose une mesure du risque de modèle "*worst case*". Le principe est de comparer la mesure de risque obtenue avec le modèle de référence avec celle obtenue avec le pire modèle, celui qui nécessite le plus de capital réglementaire.

**Définition 3.3.** La mesure "*worst case*" est définie comme suit :

$$\rho(M^*) = \max_{M_i \in M} \mathcal{R}(M_i) - \mathcal{R}(M^*)$$

Cette approche est très prudente dans la mesure où l'erreur est évalué par rapport au pire des cas. Cette méthode ne s'intéresse en aucun cas au vrai modèle qui pourrait être bien moins conservateur et dont la mesure du capital réglementaire est *a priori* moins importante que celle obtenue avec le "pire modèle".

*Remarque.*

- 1 : Nous pouvons remarquer que si le modèle de référence est celui qui requiert le plus de capital réglementaire au vu de la mesure sélectionnée (généralement la VaR) alors la mesure de l'erreur de modèle devient nulle. En effet, la réserve en capital requise par le modèle de référence est déjà maximale dans ce cas et l'erreur de modèle n'est donc pas sous-estimée.
- 2 : Une hypothèse importante dans cette approche est que le vrai modèle appartienne au set de modèles. Le choix de l'ensemble de modèle n'est pas anodin. Si l'ensemble est trop petit, la probabilité que le vrai modèle soit dans cette ensemble sera petite. A l'inverse un ensemble grand pourra contenir des modèles très éloignés du vrai modèle et la mesure "*worst case*" pourrait devenir très grande.

Cont [9] a proposé une autre mesure du risque de modèle basée sur le pire scénario.

**Définition 3.4.** La mesure de Cont est définie par :

$$\rho(M) = \max_{M_i \in M} \mathcal{R}(M_i) - \min_{M_i \in M} \mathcal{R}(M_i)$$

Cette mesure est indépendante du modèle de référence. C'est une mesure très conservatrice qui est très coûteuse en pratique car le capital réservé est maximal. Ce n'est donc pas une approche permettant d'optimiser l'allocation de capital.

Barrieu et al [3] propose deux mesures de risques construites à partir des précédentes.

**Définition 3.5.** La mesure absolue du risque de modèle :

$$AM = \rho(M^*) = \frac{\max_{M_i \in M} \mathcal{R}(M_i)}{\mathcal{R}(M^*)} - 1$$

Cette mesure quantifie à quel point le pire scénario est mauvais par rapport au modèle de référence. Si le pire scénario est donné par le modèle de référence alors la mesure absolue est nulle. Donc le risque de modèle est d'autant plus faible que le modèle de référence est proche du pire modèle.

**Définition 3.6.** La mesure relative du risque de modèle :

$$RM = \rho(M^*) = \frac{\max_{M_i \in M} \mathcal{R}(M_i) - \mathcal{R}(M^*)}{\max_{M_i \in M} \mathcal{R}(M_i) - \min_{M_i \in M} \mathcal{R}(M_i)}$$

Cette mesure donne la position relative du modèle de référence sur l'intervalle  $[\max_{M_i \in M} \mathcal{R}(M_i), \min_{M_i \in M} \mathcal{R}(M_i)]$ , en d'autres termes elle permet de définir si le modèle de référence est plus proche du meilleur scénario ou du pire. De même, cette mesure est nulle si le modèle de référence correspond au pire scénario et tend vers 1 lorsque ce dernier se rapproche du meilleur scénario.

Prolongement de la mesure :

Il est nécessaire avant d'évaluer le risque de modèle avec une de ces mesures de bien réfléchir au type de risque que l'on cherche à étudier. En particulier, nous utiliserons par la suite la value-at-risk (VaR) comme mesure de risque  $\mathcal{R}$ . Dans ce cas, il est indispensable de savoir si nous nous intéressons à la partie droite de la distribution (risque de hausse) ou à la partie gauche (risque de baisse). Nous traiterons notamment de modèles d'actifs où nous nous intéresserons au risque de baisse substantielle de l'actif à horizon 1 an. Cette baisse sera évaluée à l'aide du quantile à 0,5% de la distribution à un an et non le quantile à 99,5% (qui correspondrait à un risque de hausse). Ainsi, nous devons modifier les mesures ci-dessus qui ne sont viables que pour l'évaluation du risque par la VaR à 99,5%.

Les mesures à utiliser dans ce cas seront donc :

$$AM_- = 1 - \frac{\min_{M_i \in M} \mathcal{R}(M_i)}{\mathcal{R}(M^*)}$$

$$RM_- = \frac{\mathcal{R}(M^*) - \min_{M_i \in M} \mathcal{R}(M_i)}{\max_{M_i \in M} \mathcal{R}(M_i) - \min_{M_i \in M} \mathcal{R}(M_i)}$$

### 3.2.2 Critiques de l'approche

Cette approche présente un avantage majeur, sa simplicité. En effet, la mise en place est relativement aisée; dès lors que les modèles ont été choisis, il faut calculer le capital réglementaire requis pour chacun des modèles et la mesure du risque de modèle est alors immédiate par l'une des formules ci-dessus.

Pendant, cette approche comporte plusieurs inconvénients notamment un inconvénient économique. D'une part, cette approche requiert un ensemble de modèle qui soient en adéquation avec les hypothèses faites sur notre quantité d'intérêt et qui s'adaptent bien aux données. Comme nous l'avons expliqué auparavant, ce choix présente une première difficulté, le nombre de modèles sélectionnés. Par ailleurs, pour que la théorie ci-dessus soit valable, il est nécessaire que le vrai modèle appartienne à cet ensemble, or ne connaissant pas ce modèle, cette hypothèse est difficilement vérifiable. Le choix du modèle de référence est aussi capital puisque que ce dernier influence énormément le calcul de l'erreur de modèle.

D'autre part, économiquement, nous remarquons que ces différentes mesures de risque peuvent être, comme nous le verrons ultérieurement, très prudentes et donc augmentent significativement le capital réglementaire total  $\rho_{total}$  tel que définit précédemment. En effet, ces mesures s'intéressent uniquement aux cas extrêmes et non au cas le plus vraisemblable. Dans un soucis de gestion du risque et d'optimisation des stratégies d'allocation du capital, cette approche ne paraît pas efficace.

## 3.3 L'approche bayésienne

### 3.3.1 Principe

L'approche Bayésienne a été proposée en 1994 par Draper [14]. Comme son nom l'indique, cette approche est fondée sur la théorie de Bayes qui s'est développée à partir des années 60 dans différents domaines d'étude en particulier les statistiques, l'économétrie et l'intelligence artificielle. Draper réutilise cette théorie basée sur les probabilités conditionnelles pour introduire des poids afin de pondérer les différents modèles.

L'objectif de l'approche bayésienne est d'estimer la densité de probabilité  $p(y | x)$  en prenant en compte à la fois l'erreur de spécification du modèle et l'erreur liée aux paramètres. Ainsi, la prise en compte, dès l'estimation de la densité, des possibles erreurs de modélisation permet, contrairement à l'approche *worst case*, de diminuer l'erreur de modèle en adoptant une stratégie plus optimale.

Le fondement théorique de cette approche expliquée ci-dessus s'inspire de l'article de Cairns [7].

But : Trouver la distribution de  $p(y | x)$

$$p(y | x) = \sum_i p(y | x, M_i)p(M_i | x)$$

où

- $p(y | x)$  : la densité conditionnelle de  $y$  sachant les données historiques  $x$ .
- $p(y | x, M_i)$  : la densité conditionnelle de  $y$  sachant les données historiques  $x$  et le modèle  $M_i$
- $p(M_i | x)$  : la probabilité *a posteriori* du modèle  $M_i$ . C'est la probabilité que le modèle  $M_i$  soit le vrai modèle sachant l'historique.

Par ailleurs, nous pouvons réexprimer les quantités intervenant dans la somme :

$$p(y | x, M_i) = \int p(y | x, M_i, \theta_i)p(\theta_i | x, M_i)d\theta_i$$

où

- $p(y | x, M_i, \theta_i)$  : la densité conditionnelle de  $y$  sachant les données historiques  $x$ , le modèle  $M_i$  et le paramètre du modèle  $\theta_i$ .
- $p(\theta_i | x, M_i)$  : la distribution *a posteriori* du paramètre  $\theta_i$ .

Nous cherchons les distributions *a posteriori*  $p(M_i | x)$  et  $p(\theta_i | x, M_i)$ . En utilisant la loi de Bayes, nous obtenons :

$$p(M_i | x) = \frac{p(x | M_i)p(M_i)}{\sum_j p(x | M_j)p(M_j)} = c p(x | M_i)p(M_i)$$

et

$$p(\theta_i | x, M_i) = \frac{p(x | M_i, \theta_i)p(\theta_i | M_i)p(M_i)}{p(M_i, x)} = c_i p(x | M_i, \theta_i)p(\theta_i | M_i)$$

où

- $p(x | M_i)$  : la fonction de vraisemblance du modèle  $M_i$ .
- $p(M_i)$  : la probabilité *a priori* du modèle  $M_i$ . Cette probabilité doit être choisie consciencieusement.
- $p(x | M_i, \theta_i)$  : la fonction de vraisemblance du modèle  $(M_i, \theta_i)$
- $p(\theta_i | M_i)$  : la distribution *a priori* du paramètre  $\theta_i$  associé au modèle  $M_i$ . Cette probabilité est aussi choisie par nos soins de manière plus ou moins subjective.
- $c$  et  $c_i$  : des constantes de normalisation qui assurent que les probabilités se somment à 1.

Plusieurs auteurs ont proposé des formules permettant d'approximer la vraisemblance des modèles  $p(x | M_i)$  et la distribution  $p(\theta_i | x, M_i)$  qui peuvent s'avérer difficile à calculer en pratique.

*Remarque.*

- Draper a montré que  $p(\hat{\theta}_i | M_i)$  est négligeable en particulier lorsque la loi de probabilité de cet *a priori* est continue. En d'autres termes, les résultats de l'approche bayésienne ne sont pas très sensibles au choix de l'*a priori* pour les paramètres. Cependant, ce n'est pas le cas pour l'*a priori* relatif au modèle. Ce dernier point pourrait nous laisser penser que l'approche bayésienne n'est peut être pas la bonne approche pour mesurer l'incertitude du modèle à cause de sa sensibilité. En effet, 2 modèles peuvent avoir la même adéquation aux données mais donner des prédictions différentes.
- Il existe une autre approche appelée l'approche bayésienne partielle. Cette approche consiste à trouver le modèle  $M^*$  qui a la plus grande probabilité *a posteriori*. Une fois ce modèle trouvé (on considère que c'est le vrai modèle), on trouve la distribution de  $y$  sachant  $x$  de la même manière en ne conditionnant que par rapport à ce modèle  $M^*$ . Cette approche ne permet donc pas de prendre en compte l'erreur de spécification mais uniquement l'erreur de paramétrisation.

- Afin de simplifier la modélisation, l'approximation suivante peut être faite :  $p(y | x, M_i) \approx p(y | x, M_i, \hat{\theta}_i)$  où  $\hat{\theta}_i$  est l'EMV (Estimateur du Maximum de Vraisemblance). Cette approximation n'est valide que dans le cas où les données sont en nombre suffisant et en considérant que l'EMV est assez fiable. Dans ce cas, on néglige l'erreur liée à la paramétrisation en considérant qu'avec un grand nombre de données et avec les méthodes actuelles d'estimation des paramètres, cette source est négligeable.

### Etapes de l'approche bayésienne :

I.

Choix de la classe de modèles  $M$  en se basant sur des avis d'experts et sur la littérature académique. Ce choix est subjectif mais comme expliqué précédemment, cet ensemble de modèles doit contenir le vrai modèle pour que les résultats soient vérifiés. Cette décision n'est donc pas anodine dans notre étude et doit mériter une grande attention.

II.

Choix de l'*a priori* sur le modèle  $p(M_i)$  (i.e. la probabilité que  $M_i$  soit le vrai modèle) d'après des jugements subjectifs (intuitions, expérience, avis d'experts). Rappel : l'*a priori* sur les paramètres est négligé.

III.

Estimation des paramètres  $\theta_i$  pour chaque modèle  $M_i$  puis projection de la valeur de  $y$  à l'aide des  $(M_i, \theta_i)$ , c-à-d calcul de  $p(y | x, M)$ . Cette étape consiste à définir pour chaque modèle le capital réglementaire qui lui est associé en fonction de la mesure de risque choisie (généralement la VaR).

IV.

Détermination de l'*a posteriori* de chaque modèle  $p(M_i | x)$ . Cette probabilité va permettre d'introduire des poids à chaque modèle qui seront utilisés pour évaluer la projection de la valeur de  $y$ ,  $p(y | x)$ .

V.

Calcul de l'erreur de modèle. L'erreur de modèle est définie comme suit :

$$\rho(M) = \frac{\mathcal{R}_{Bootstrap}(M)}{\mathcal{R}(M^*)} - 1$$

où  $\mathcal{R}_{Bootstrap}(M)$  correspond à la mesure de risque obtenue en faisant un Bootstrap parmi nos différents modèles pondérés à l'aide des probabilités *a posteriori*.

#### Prolongement de la mesure :

De même que pour l'approche bayésienne, nous modifions légèrement la mesure ci-dessus afin de l'adapter à des situations inverses (ex : baisse de l'actif mesurée à l'aide d'une VaR). Ainsi la mesure correspondante sera :

$$\rho_-(M) = 1 - \frac{\mathcal{R}_{Bootstrap}(M)}{\mathcal{R}(M^*)}$$

### 3.3.2 Critiques de l'approche

L'approche bayésienne présente l'avantage de prendre en compte immédiatement et simultanément l'erreur de spécification et de paramétrisation. Cela permet d'avoir une mesure du risque plus précise. Ainsi les provisions en capital requises pour couvrir le risque de modèle sont moins importantes en comparaison avec l'approche *worst case*. Cette approche est donc conseillée pour une entreprise qui cherche à optimiser sa stratégie d'allocation du capital tout en ayant une bonne gestion de ses risques.

Cependant, elle présente tout de même plusieurs inconvénients. Le premier, similaire à l'approche *worst case*, réside dans le choix de l'ensemble de modèles qui doit potentiellement contenir le vrai modèle.

Le second concerne le choix des *a priori* des modèles de la classe qui peut s'avérer très complexe et loin d'être une évidence. Ce choix peut fortement impacter la probabilité *a posteriori* et ne peut donc pas être fait de manière aléatoire. Généralement, ce choix est subjectif et basé plus ou moins sur des connaissances théoriques ou liées à l'expérience. Il est nécessaire d'être très prudent lors de cette étape.

En outre, dans l'approche théorique plusieurs approximations sont faites sur les distributions afin de simplifier la modélisation de l'approche. Cependant ces approximations ne sont vraies que si le nombre de données est suffisant. De même, l'erreur liée aux paramètres ne peut être négligée que si nous possédons un grand nombre de données historiques permettant de supposer que l'EMV est fiable. La stabilité de l'EMV peut être démontrée à l'aide de mesures d'incertitude.

Pour finir, cette approche peut s'avérer difficile à mettre en application. En effet, la vraisemblance peut ne pas être calculable pour certains modèles. Cette dernière peut aussi être tellement grande (d'un ordre de grandeur bien supérieur à celui des autres modèles) que quasiment l'intégralité du poids (plus de 99%) est donné à ce modèle. Dans ce cas, l'approche par pondération n'apporte pas d'information sensiblement différente de l'approche *worst case*. Nous verrons donc d'autres pondérations lorsque celle-ci nous semblera non applicable.

## 4 L'erreur de modèle sur des structures de dépendance

En mathématiques de l'assurance, une hypothèse centrale est l'indépendance des variables aléatoires. Or, nous observons, en pratique, des dépendances entre les différents produits d'assurance qui s'expliquent notamment par leurs expositions à des facteurs de risque communs : inflation, cycle de marché, CAT extrêmes... Par conséquent, cette dépendance stochastique est présente dans beaucoup d'activités de l'assurance et doit être prise en compte lors de la modélisation des risques.

Par ailleurs, la réglementation actuelle oblige les assureurs et réassureurs à être prudents et en particulier à ne pas sous-estimer les dépendances comme cela a pu être le cas dans le passé ; notamment en finance avec l'utilisation excessive de la copule Gaussienne. Ainsi, la considération de la dépendance est aussi essentielle pour l'élaboration d'une gestion des risques solide et raisonnable.

Dans notre étude, nous nous concentrons uniquement sur la dépendance entre deux variables que nous choisissons de modéliser par une copule classique. Ce choix de modéliser la dépendance par des copules s'explique par deux éléments : la possibilité de prendre en compte des dépendances non linéaires et la possibilité de traiter séparément les erreurs de dépendance (liées au choix de la copule) et les erreurs liées au choix des modèles des marges.

### 4.1 Rappels et copules classiques

Une copule est une fonction qui relie les distributions marginales univariées à leur distribution jointe multivariée. Dans le cas de variables indépendantes, la loi jointe n'est autre que le produit des lois marginales.

Dans la suite, nous nous intéressons uniquement à la dépendance entre deux variables ; toutes les notions et copules définies seront donc en dimension 2.

**Définition 4.1.** Une copule de dimension 2 est une fonction  $C : [0, 1]^2 \rightarrow [0, 1]$  telle que :

- $C$  est "grounded"<sup>3</sup> et 2-croissante<sup>4</sup>
- $\forall u \in [0, 1], C(u, 1) = u$  et  $C(1, u) = u$

Le théorème fondamental dans la théorie des copules est le théorème de Sklar[41] de 1959 qui introduit cette notion. Ce théorème permet de décrire la dépendance de la distribution jointe en décomposant la fonction de répartition multivariée  $F$  en la copule  $C$  et les distributions marginales  $F_i$ .

**Théorème 4.1** (Théorème de Sklar). *Soit  $F$  une fonction de répartition bivariée ayant des marginales  $F_1$  et  $F_2$ , alors il existe une copule  $C$  telle que  $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$ . De plus, si les marges sont continues cette copule est unique.*

*Remarque* (Densité d'une copule). Sous des conditions usuelles de dérivabilité d'une fonction à deux variables, la densité de la copule est donnée par :

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}$$

Ainsi nous pouvons calculer la vraisemblance d'une copule.

D'autres propriétés sur les copules pourront être mentionnées dans la suite de notre exposé en fonction des besoins. Nous allons désormais présenter les copules classiques que nous étudierons ultérieurement. Ces copules ont été choisies puisqu'elles permettent, en dimension 2, de modéliser la plupart des formes de dépendance possible. Par exemple, la copule gaussienne correspond à la dépendance linéaire tandis que les copules de Clayton et Gumbel modélisent les dépendances de queues inférieures et supérieures respectivement.

3.  $C$  est "grounded" si  $C(u, 0) = C(0, v) = 0, \forall (u, v) \in [0, 1]^2$

4.  $C$  est 2-croissante si  $V_C(B) = C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0, \forall B = [u_1, u_2] \times [v_1, v_2] \in [0, 1]^2$

## Copules elliptiques

Les copules elliptiques sont les copules dont les contours de la densité forment des ellipsoïdes. En d'autres termes, ce sont les copules associés aux vecteurs elliptiques définis par :

$$X \sim \psi(\mu, \Sigma, \Phi) \text{ (loi elliptique de paramètres } (\mu, \Sigma, \Phi)) \Leftrightarrow \mathbb{E}[\exp(i(X - \mu)t)] = \Phi(t'\Sigma t)$$

En particulier, deux copules usuelles appartiennent à la famille des copules elliptiques ; la copule Gaussienne et la copule de Student.

### La copule Gaussienne

**Définition 4.2.** Soit  $\Sigma$  une matrice carrée d'ordre 2, symétrique, définie positive avec tous les éléments de la diagonale valant 1. On note  $\Phi_\Sigma$  la fonction de répartition bivariée de la loi normale de variance-covariance  $\Sigma$  et  $\Phi$  la fonction de répartition univariée de la loi normale centrée réduite. La copule gaussienne bi-dimensionnelle est définie par :

$$C_{Gaussienne, \Sigma}(u, v) = \Phi_\Sigma(\Phi^{-1}(u), \Phi^{-1}(v))$$

### La copule de Student

**Définition 4.3.** Soit  $\Sigma$  une matrice carrée d'ordre 2, symétrique, définie positive avec tous les éléments de la diagonale valant 1. On note  $T_{\Sigma, \nu}$  la fonction de répartition bivariée de la loi Student à  $\nu$  degrés de liberté ( $\nu > 0$ ) et de variance-covariance  $\Sigma$  et  $t_\nu$  la fonction de répartition univariée de la loi Student à  $\nu$  degrés de liberté. La copule de Student bi-dimensionnelle est définie par :

$$C_{Student, \Sigma, \nu}(u, v) = T_{\Sigma, \nu}(t_\nu^{-1}(u), t_\nu^{-1}(v))$$

*Remarque.* Lorsque le nombre de degrés de liberté  $\nu$  tend vers l'infini, la copule de Student de paramètres  $(\Sigma, \nu)$  converge vers la copule Gaussienne de paramètres  $\Sigma$ . Un avantage de la copule de Student est qu'elle permet de modéliser des variables ayant des queues de distribution plus épaisses que la copule gaussienne (lorsque  $\nu$  est petit).

## Copules archimédiennes

Une copule archimédienne est définie par Genest et al. [21] comme suit :

**Définition 4.4.** Soit  $h$  un générateur strict, c-à-d. une fonction  $\mathcal{C}^2$  avec  $h(1) = 0$ ,  $h'(u) < 0$  et  $h''(u) > 0$ ,  $\forall u \in [0, 1]$ . La copule archimédienne est :

$$C(u, v) = \begin{cases} h^{-1}(h(u) + h(v)) & \text{si } h(u) + h(v) \leq h(0) \\ 0 & \text{sinon} \end{cases}$$

*Remarque.* Les copules archimédiennes sont très utilisées en pratique car elles permettent de modéliser la dépendance pour de grandes dimensions avec seulement un paramètre anisi que la dépendance des événements extrêmes.

Voici des exemples de copules archimédiennes dont nous nous servons pour notre étude :

### La copule de Clayton

**Définition 4.5.**

$$C(u, v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-\frac{1}{\alpha}}, \quad \alpha > 0$$

### La copule de Gumbel

**Définition 4.6.**

$$C(u, v) = \exp\left(-\left((-\ln(u))^\alpha + (-\ln(v))^\alpha\right)^{\frac{1}{\alpha}}\right), \quad \alpha \geq 1$$

### La copule de Frank

**Définition 4.7.**

$$C(u, v) = -\frac{1}{\alpha} \ln\left(1 - \frac{1}{e^{-\alpha} - 1}(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)\right), \quad \alpha > 0$$

## 4.2 Estimation du paramètre de copule

Afin d'estimer le paramètre de la copule à partir de nos données, plusieurs méthodologies ont été développées. Cependant, selon le contexte (les données, ce que l'on cherche à modéliser...), il faut choisir la méthode la plus pertinente. Nous présentons ici deux méthodes, la méthode du maximum de la pseudo-vraisemblance et la méthode dite du CQEP qui est celle utilisée en pratique par la SCOR pour les risques non-vie.

### Maximisation de la pseudo-vraisemblance

Cette méthode semi-paramétrique a été proposée par Genest et al [20]. Cette méthode est semi-paramétrique car nous supposons connue la copule mais pas les marges associées. Il faut donc estimer à la fois les marges et le paramètre de la copule. Les marges sont estimées par les fonctions de répartition empiriques normalisées et le paramètre de la copule est obtenu en maximisant la pseudo log-vraisemblance qui est donnée par :

$$L(\theta) = \sum_{k=1}^n \log (C_{\theta}(F_{1,n}(X_{1,k}), F_{2,n}(X_{2,k})))$$

avec

- $\theta$  : le paramètre de la copule
- $n$  : la taille de l'échantillon de données
- $F_{i,n}(X_{i,k}) = \frac{n-1}{n} \sum_{j=1}^n \mathbb{1}_{X_{i,j} \leq X_{i,k}}$
- $C_{\theta}$  : la copule paramétrique supposée

Les auteurs ont prouvé que cet estimateur est convergent et asymptotiquement normal.

### CQEP

La CQEP (conditionnal quantile exceedance probability) est une mesure de dépendance entre deux facteurs de risques qui se concentre sur les queues de distribution. Elle permet de déterminer la probabilité que les deux facteurs de risque s'éloignent beaucoup de la moyenne au même moment. Cette mesure est symétrique au même titre que la corrélation. On distingue la CQEP lower tail qui évalue la dépendance dans la queue gauche de la distribution et la CQEP upper tail qui s'intéresse à la corrélation entre les deux facteurs dans la queue droite de la distribution.

La CQEP upper tail est définie par :

$$CQEP_{up,\alpha}(X, Y) = \mathbb{P}(X > F_X^{-1}(\alpha) \mid Y > F_Y^{-1}(\alpha))$$

avec :

- $X$  et  $Y$  les facteurs de risque
- $\alpha$  le niveau du quantile, souvent 99% ou 99.5% pour prendre en compte les événements dans les queues de distribution et être en accord avec la réglementation.
- $F_X^{-1}$  l'inverse généralisée de la fonction de répartition associé au facteur de risque  $X$ .

*Remarque.* Si  $X$  et  $Y$  sont indépendants alors  $CQEP_{up,\alpha}(X, Y) = 1 - \alpha$

La CQEP lower tail est définie par :

$$CQEP_{low,\alpha}(X, Y) = \mathbb{P}(X \leq F_X^{-1}(\alpha) \mid Y \leq F_Y^{-1}(\alpha))$$

avec  $\alpha$  le niveau du quantile généralement égal à 1%.

Cette dernière fonction est moins utilisée puisque dans beaucoup de cas, nous nous intéressons aux événements extrêmes présents dans la queue droite. Cependant, nous verrons par la suite que dans certains cas, l'objet d'étude est la queue gauche de la distribution et il est alors nécessaire d'utiliser cette fonction.

*Remarque.* Si  $X$  et  $Y$  sont indépendants alors  $CQEP_{low,\alpha}(X, Y) = \mathbb{P}(X \leq F_X^{-1}(\alpha)) = \alpha$

Notons  $C(u, v, \theta)$  la copule sélectionnée avec  $\theta$  le paramètre de la copule que nous cherchons à estimer. Soit  $\alpha$  un niveau fixé. D'après la formule de Bayes, la CQEP est donnée par :

$$\begin{aligned} CQEP_{up,\alpha}(X, Y) &= \frac{\mathbb{P}(X > F_X^{-1}(\alpha), Y > F_Y^{-1}(\alpha))}{\mathbb{P}(Y > F_Y^{-1}(\alpha))} \\ &= \frac{1 - \mathbb{P}(Y \leq F_Y^{-1}(\alpha)) - \mathbb{P}(X \leq F_X^{-1}(\alpha)) - \mathbb{P}(X \leq F_X^{-1}(\alpha), Y \leq F_Y^{-1}(\alpha))}{1 - \mathbb{P}(Y \leq F_Y^{-1}(\alpha))} \\ &= \frac{1 - 2\alpha + C_\theta(\alpha, \alpha)}{1 - \alpha} = q_{up}(\theta) \end{aligned}$$

Par le même raisonnement,

$$CQEP_{low,\alpha}(X, Y) = \frac{C_\theta(\alpha, \alpha)}{\alpha} = q_{low}(\theta)$$

En pratique, les CQEP s'obtiennent de manière empirique pour un niveau  $\alpha$  fixé, notées  $cqep_{n,up}$  et  $cqep_{n,low}$ . Le paramètre de la copule est ensuite calculé en inversant numériquement la copule. En d'autres termes, nous recherchons le paramètre qui minimise l'erreur :

$$\hat{\theta} = \arg \min_{\theta} (q(\theta) - cqep_{n,\cdot})^2$$

*Remarque.* Si pour différents niveaux  $\alpha$  le paramètre estimé est très variable, cela signifie que la copule n'est pas adaptée pour représenter l'ensemble de la distribution, elle ne convient que pour modéliser la queue de la distribution. Cette remarque peut nous permettre d'évaluer la pertinence de la copule choisie surtout dans les cas où la distribution n'est pas particulièrement lourde et où les données sont suffisantes pour avoir une bonne idée de la distribution jointe.

## Conclusion

Les deux méthodes ne peuvent pas être utilisées dans les mêmes contextes. Dans le cas où nous supposons une forte dépendance entre les facteurs de risque dans les extrêmes, nous allons davantage nous intéresser aux queues de distribution et le QEP sera un bon indicateur de la corrélation recherchée. C'est notamment particulièrement vrai lorsque nous étudions les risques non-vie dont les données sont peu nombreuses et dont les experts conjecturent des queues lourdes de distribution jointe.

A l'inverse, lorsque la distribution jointe n'est pas très lourde et que nous disposons de suffisamment de données, la méthode qui consiste à maximiser la pseudo log-vraisemblance semble plus adaptée car elle prend en compte toute la distribution et l'adéquation s'améliore avec le nombre de données.

## 4.3 Proposition de méthodologie de quantification de l'erreur de modèle

Dans cette sous-partie, nous présentons la méthodologie mise en place afin de quantifier l'erreur de spécification du modèle. Tout d'abord, nous nous concentrerons uniquement sur l'erreur liée au choix de la copule puis nous tenterons d'évaluer conjointement l'erreur de dépendance et l'erreur de modélisation des marginales.

### L'erreur de dépendance pure

Pour évaluer l'erreur de dépendance pure, nous fixons la loi des marginales que nous supposons être la "vraie loi" et nous notons  $F_1$  et  $F_2$  les fonctions de répartition du premier facteur de risque et du deuxième respectivement. Afin d'être en adéquation avec le cadre réglementaire, ces distributions correspondent aux distributions projetées à horizon 1 an. Dans notre étude, nous modélisons la dépendance uniquement entre deux variables (copule de dimension 2) à l'aide des copules classiques mentionnées ci-dessus ; à savoir la copule Gaussienne, la copule de Student, la copule de Clayton, la copule de Gumbel et la copule de Frank. Pour chacune des copules  $C$ , nous procédons comme suit

1. Définir le paramètre  $\hat{\theta}$  de la copule à l'aide d'une des deux méthodes présentées

2. Générer des vecteurs (de dimension 2) de nombres aléatoires à partir de la copule paramétrée par le paramètre estimé, ie simuler  $(u, v) \sim C_{\hat{\theta}}$
3. Obtenir la distribution jointe  $(x, y) = (F_1^{-1}(u), F_2^{-1}(v))$

Pour évaluer l'erreur de modèle, il faut ensuite se rapporter à l'exposition du portefeuille. Soient  $E_1$  et  $E_2$  les expositions respectives des facteurs de risque 1 et 2 dans le portefeuille. Après mise à l'échelle des grandeurs étudiées si nécessaire, la distribution à 1 an est donnée par :

$$xE_1 + yE_2$$

Pour chaque copule, on peut ainsi évaluer la VaR pour différente période de retour.

La mise en place de l'approche *worst case* est alors immédiate et similaire au cas où nous n'étudions qu'une seule variable. Il suffit de fixer les marginales et d'évaluer la VaR pour chaque copule. Le calcul de l'erreur se fait en comparant la mesure de risque minimale avec la mesure de risque de la copule de référence. L'approche bayésienne est aussi mise en œuvre de la même manière même si le choix des poids sera plus complexe et pourra nécessiter des réajustements voire davantage d'avis d'expert afin de les déterminer. Les aspects plus spécifiques seront discutés dans la partie applicative.

### L'erreur liée à la modélisation des marginales et de la copule

Le calcul de l'erreur de modèle dans le cas de l'approche *worst case* se fait de la même manière que pour la dépendance pure sauf que la VaR n'est pas évaluée pour une marginale fixée mais pour chaque marginale. Ensuite, le risque de modèle (la mesure relative) est évalué en choisissant une copule et les marginales associées.

Dans le cas de l'approche par pondération, l'erreur de modèle est obtenue en faisant simultanément deux simulations par bootstrap ; une sur les marginales pour définir la distribution projetée à 1 an et une sur les copules. Nous utilisons pour ce faire les mêmes pondérations que dans le cas d'une variable aléatoire seule pour la pondération des marginales et de la dépendance pure pour la pondération de la copule. Ainsi, l'erreur de modèle prend en compte simultanément le risque lié à la dépendance et celui du à la modélisation des marginales.

## Partie III

# Etude de modèles d'actifs

## 5 Application à un risque isolé

Nous allons commencer par mettre en œuvre les méthodologies décrites dans les parties précédentes sur un modèle d'actifs. La modélisation d'actifs a fait l'objet de nombreuses recherches. En effet, les financiers ont cherché des modèles décrivant, de façon pertinente, les données passées du prix des actifs afin de prédire leurs rendements futurs ; aussi bien les hausses que les baisses. Le modèle le plus connu est le modèle de Black-Scholes bien que ce dernier soit trop simple pour décrire de manière réaliste les événements de marché.

Plusieurs sources d'erreur de modèle interviennent lors de la conception d'un modèle de valorisation. Tout d'abord, il y a l'erreur de données issue à la fois du type de données utilisées, de la propriété de la base de données ainsi que de la période étudiée. Nous avons évoqué dans la description de l'erreur de modèle que cette source avait un impact dès la création du modèle. Par ailleurs, l'autre source principale d'erreur de modèle, qui sera l'objet de notre étude, est l'erreur de spécification du modèle et en particulier le choix du modèle.

Nous allons donc modéliser un actif à l'aide de plusieurs modèles de valorisation utilisés dans la pratique. Nous appliquerons ensuite les deux approches de quantification de l'erreur de modèle à cet ensemble de modèle. Pour cela, nous projeterons la valeur de l'actif à un an et nous nous intéresserons aux situations extrêmes de baisse exceptionnelle de l'actif et en particulier aux situations ne se produisant qu'une fois tous les 200 ans. Ces situations correspondent au quantile à 0,5% de la distribution de l'actif à un an.

Nous commencerons par présenter l'actif choisi pour notre étude ainsi que les différents modèles sélectionnés afin de valoriser cet actif.

Par la suite, nous développerons la mise en place de l'approche bayésienne. Cette approche repose sur le choix d'un *a priori*. Nous chercherons donc à faire ce choix grâce à différents outils statistiques ou de l'ordre du jugement. Le choix d'un *a priori* étant toujours subjectif, nous analyserons la sensibilité des résultats à des changements d'*a priori*.

Pour finir, nous regarderons les résultats obtenus avec l'approche worst case afin de comparer les deux approches et l'impact du choix d'une approche sur le capital nécessaire à la couverture du risque de modèle. D'ores et déjà, nous avons l'intuition que l'approche bayésienne nécessitera la réserve de moins de capital puisqu'elle permet d'intégrer directement l'erreur de modèle.

### 5.1 Présentation des modèles et des données

#### 5.1.1 Les données

Nous étudions l'indice boursier journalier britannique GDDLUK de 2001 à 2014. Cet indice, calculé dans la monnaie locale (ici la Livre Sterling £), est publié par la société MSCI et mesure le prix moyen brut d'un ensemble d'actifs en utilisant une technique de moyenne pondérée couplée avec une méthode de chaînage des indices. L'indice est calculé 5 jours par semaine du lundi au vendredi avec des exceptions pour certains jours fériés et dimanche.

Nous cherchons à trouver un modèle qui nous permette de prédire, à l'aide de l'historique de 2001 à 2013, le cours de l'indice pour l'année 2014 et en particulier au dernier jour ouvré de l'année 2014. Etant donné que nous ne chercherons pas à modéliser l'erreur liée aux données, nous avons effectué un certain nombre de retraitements de nos données (explicités ci-dessous) préalablement à l'étude afin de limiter au maximum cette source d'erreur.

Nous travaillerons essentiellement sur les données mensualisés pour des raisons de simplicité et de rapidité de paramétrisation et projection des modèles. Nous présenterons cependant les résultats obtenus sur données journalières dans la dernière partie afin d'évaluer les différences notables et l'impact du choix de la fréquence étudiée.

Dans les sources potentielles d'erreur de données, nous avons évoqué le choix des données. Ici, les données sont celles utilisées en interne par la SCOR pour la modélisation et la projection des cours boursiers. Ces données font l'objet de validation de leur adéquation au profil de risque. L'historique de données commence à partir de fin 1969 mais nous ne travaillerons que sur les données à compter de 2001 puisqu'auparavant les données ne sont que mensuelles et non journalières. Bien que l'historique retenu ne soit pas très long, nous possédons suffisamment de données et l'étendue de temps considérée présente des périodes de croissance économique ainsi que des périodes de récession notables comme la crise des *subprimes*. L'erreur relative au choix des données nous paraît donc limitée au vu de tout ce que nous venons de mentionner.

Par ailleurs, nous avons remarqué qu'il était important que les données soient bien traitées avant utilisation afin d'enlever les données erronées et de prendre en compte les données manquantes ou tronquées. Ici, nous avons ôté de la base les données en doublon. En effet, presque sûrement, les cours de l'action en clôture de deux jours ouvrés successifs sont différents, même au centième près. Nous avons pu constater que, lorsque les cours sont identiques à deux dates successives dans notre base, il s'agit bien souvent d'un jour férié auquel le cours de la veille a été attribué par défaut (c'est le cas notamment pour les fêtes de fin d'année).

Nous présentons ci-dessous le graphe 2 correspondant au cours boursier britannique entre 2001 et 2014 ainsi que le graphe 3 des rendements logarithmiques associés (après retraitement des données).

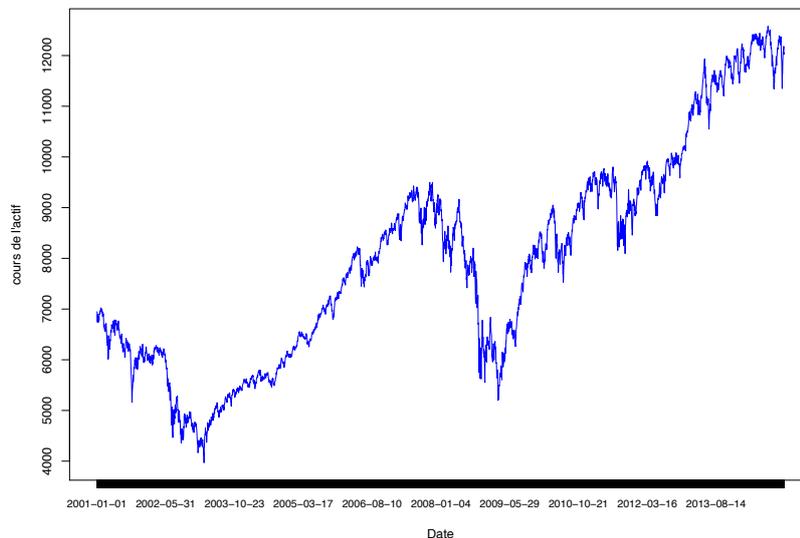


FIGURE 2 – Cours boursier britannique entre 2001 et 2014

### 5.1.2 Les modèles

Nous allons chercher à inférer la volatilité à partir de l'historique de rendements boursiers. En effet, la prédiction de la volatilité future permet de prendre en compte de futures variations des cours et de prédire le risque. Plusieurs propriétés observables des rendements sont à noter avant de créer un modèle de valorisation bien que nous verrons par la suite que plusieurs modèles populaires n'en tiennent pas compte :

- La moyenne des rendements est nulle à court terme.
- La moyenne est très nettement inférieure à la volatilité (dans notre application, elle est environ 70 fois plus petite), c'est pourquoi la seule connaissance de la moyenne n'est pas suffisante pour prédire les rendements.
- Les rendements ont une faible autocorrélation, il est par conséquent difficile de prédire un rendement à partir de son historique.

- Il y a phénomène de volatilité clustering, c.-à-d. que les grands rendements et les petits ont tendance à se produire successivement, en grappe. Cette constatation est obtenue à partir de l'autocorrélation des carrés des rendements.

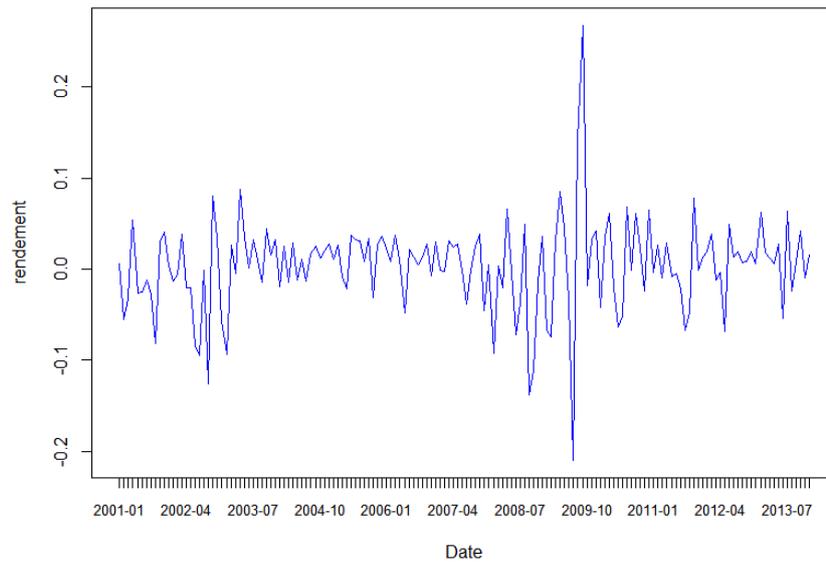


FIGURE 3 – Rendements logarithmiques mensualisés du cours britannique entre 2001 et 2014

Nous avons retenu quatre modèles d'actifs plus ou moins utilisés en pratique : le modèle de Black-Scholes, le modèle de Merton, le modèle de Garch et le modèle de Black-Scholes à changement de régimes markovien ou regime-switching. Le modèle de Garch étant le plus répandu et le plus pertinent au dire d'expert, nous le choisissons comme modèle de référence, en d'autres termes, nous cherchons à modéliser l'erreur liée à l'utilisation de ce modèle plutôt qu'un des trois autres modèles sélectionnés.

Dans la suite de notre étude, nous concentrerons notre attention uniquement sur le risque lié au choix du modèle. En conséquence, nous ne traiterons pas de l'erreur d'estimation qui peut être considérée comme négligeable avec les méthodes actuelles d'estimation. Cependant, nous veillerons au même titre que les données à obtenir des estimateurs de bonne qualité afin de limiter cette source d'erreur.

### Le modèle de Black-Scholes

Le modèle de Black-Scholes[4] est le modèle de base dans les théories de valorisation d'actifs. Les hypothèses centrales de ce modèle sont que le logarithme des rendements du prix d'un actif suit un mouvement Brownien géométrique et que la tendance et la volatilité des prix sont constants au cours du temps. Le cours de l'action  $S_t$  est donc modélisé par le processus stochastique suivant :

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

où :

- $S_t$  : le cours de l'action à la date  $t$
- $W_t$  : un mouvement Brownien géométrique,  $W_t \sim \mathcal{N}(0, t)$
- $\mu$  : la tendance du cours de l'actif
- $\sigma$  : la volatilité ou l'écart-type du rendement de l'actif
- $t = 1 \dots T$  où  $T$  la période étudiée

La solution de cette équation différentielle stochastique peut être obtenue à l'aide du lemme d'Itô [24].

Le cours de l'actif est donc défini par :

$$S_t = S_0 \exp \left( \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma dW_t \right)$$

Par conséquent, le rendement logarithmique défini par  $r_t = \log \left( \frac{S_t}{S_{t-1}} \right)$  suit une loi normale ;  
 $r_t \sim \mathcal{N} \left( \mu - \frac{\sigma^2}{2}, \sigma^2 \right), \forall t.$

Estimation des paramètres : Nous choisissons d'estimer les paramètres par la méthode des moments à partir de l'historique des rendements. Ainsi les paramètres estimés de  $\mu$  et  $\sigma$  sont

$$\hat{\sigma} = \sqrt{\frac{1}{T} \sum_{t=1}^T (r_t - \bar{r})^2} \quad \text{avec} \quad \bar{r} = \frac{1}{T} \sum_{t=1}^T r_t$$

et

$$\hat{\mu} = \bar{r} + \frac{\hat{\sigma}^2}{2}$$

La vraisemblance : Bien que nous ayons choisi d'estimer les paramètres par la méthode des moments, nous maximiserons la vraisemblance dans un second temps pour affiner l'estimation de nos paramètres et mettre en œuvre l'approche bayésienne. La vraisemblance du modèle de Black-Scholes s'écrit :

$$L(r_1, \dots, r_T | \theta) = \prod_{t=1}^T \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{(r_t - (\mu - \frac{\sigma^2}{2}))^2}{2\sigma^2} \right) \quad \text{où} \quad \theta = (\mu, \sigma)$$

Avantages et inconvénients : Comme explicité précédemment, ce modèle présente un avantage, sa simplicité qui est aussi son principal défaut. En effet, beaucoup d'hypothèses sont trop simplistes et par conséquent pas réalistes. Les principaux inconvénients de ce modèle sont :

- l'hypothèse de gaussianité des rendements qui ne permet pas de prendre en compte les événements extrêmes pour lesquels une distribution à queue épaisse serait plus adaptée.
- l'hypothèse de constance des paramètres qui n'est réaliste qu'en période de stabilité économique.
- l'hypothèse de continuité de l'actif qui empêche la prise en compte des sauts boursiers que l'on peut constater en période de troubles sur les marchés.

## Le modèle de Merton

Le modèle de Merton [32] s'inscrit dans la continuité du modèle de Black-Scholes et vise à corriger certains défauts cités ci-dessus. En particulier, ce modèle introduit un processus de Poisson Composé dans la formule du cours de l'action afin de modéliser les sauts boursiers. Les sauts se produisent donc de manière aléatoire selon un processus de Poisson.

Ainsi le cours de l'actif est donné par :

$$S_t = S_0 \exp \left( \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma dW_t + \sum_{i=1}^{N_t} Y_i \right)$$

où

- $N = (N_t)_t$  : Un processus de Poisson d'intensité  $\lambda$  qui modélise la fréquence des sauts
- $Y = (Y_i)_i$  : Une suite de variables i.i.d. de loi normale  $\mathcal{N}(0, \sigma_Y^2)$  qui modélise l'amplitude des sauts. Nous choisissons la loi normale pour des questions de simplicité. Cette loi est symétrique ce qui sous-tend que les sauts sont symétriques à la hausse et à la baisse avec une amplitude moyenne nulle. En pratique, les sauts ont, en général, une amplitude plus forte à la baisse.
- $N$ ,  $Y$  et  $W$  sont supposés mutuellement indépendants.

Estimation des paramètres : Nous devons estimer quatre paramètres  $(\mu, \sigma, \lambda, \sigma_X)$ . Pour ce faire, nous utilisons la méthode des moments qui est décrite dans l'article de Planchet et al. [37] avec d'autres méthodes d'estimations. Nous n'utilisons pas le maximum de vraisemblance pour des questions pratiques mais nous présentons la vraisemblance du modèle ci-après qui nous servira pour mettre en place l'approche bayésienne.

Pour cela, nous avons besoin de calculer 4 moments afin d'obtenir un système non-linéaire de quatre équations et quatre inconnues. Nous utilisons, d'une part, le premier moment, c.-à-d. l'espérance. Par ailleurs, les moments centrés d'ordre impair sont nuls par symétrie de la distribution. Nous calculons donc les trois premiers moments centrés d'ordre pair qui s'écrivent :

$$m_{2p} = \mathbb{E}[(r - m)^{2p}] = e^{-\lambda} \frac{(2p)!}{2^p p!} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} (\sigma^2 + n\sigma_X^2)^p, \quad p = 1, 2, \dots$$

Ainsi nous obtenons le système d'équation suivant :

$$\mathbb{E}[r] = \mu - \frac{\sigma^2}{2} \quad (1)$$

$$m_2 = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} (\sigma^2 + n\sigma_X^2) = \sigma^2 + \lambda\sigma_X^2 \quad (2)$$

$$m_4 = 3e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} (\sigma^2 + n\sigma_X^2)^2 = 3[(\sigma^2)^2 + 2\sigma^2\sigma_X^2\lambda + \lambda(1 + \lambda)(\sigma_X^2)^2] \quad (3)$$

$$m_6 = 15e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} (\sigma^2 + n\sigma_X^2)^3 \quad (4)$$

$$= 15[(\sigma^2)^3 + 3(\sigma^2)^2\sigma_X^2\lambda + 3\sigma^2(\sigma_X^2)^2(\lambda + \lambda^2) + 6(\lambda + 3\lambda^2 + \lambda^3)(\sigma_X^2)^3] \quad (5)$$

La résolution du système peut s'effectuer simplement à l'aide de différents algorithmes de résolution comme celui de Newton-Raphson. Nous définissons la fonction  $F_1$  par :

$$F_1 : \quad \mathbb{R}^4 \quad \rightarrow \quad \mathbb{R} \\ (\mu, \sigma, \lambda, \sigma_X) \mapsto \left( \frac{1}{T} \sum_{t=1}^T r_t - \mathbb{E}[r] \right)^2$$

et les fonctions  $F_i$ ,  $i = 2, 4, 6$  par :

$$F_i : \quad \mathbb{R}^4 \quad \rightarrow \quad \mathbb{R} \\ (\mu, \sigma, \lambda, \sigma_X) \mapsto \left( \frac{1}{T} \sum_{t=1}^T (r_t - \bar{r})^i - m_{2i} \right)^2$$

Nous proposons ici de minimiser la fonction G telle que

$$G : \mathbb{R}^4 \rightarrow \mathbb{R} \\ x \mapsto F_1(x) + F_2(x) + F_4(x) + F_6(x)$$

Le minimum de la fonction G devrait converger vers zéro dans la mesure où chaque fonction  $F_i$  doit être asymptotiquement nulle si les estimateurs sont convergents.

La vraisemblance : La vraisemblance du modèle de Merton s'écrit :

$$L(r_1, \dots, r_T | \theta) = \prod_{t=1}^T \left( \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \frac{1}{\sqrt{\sigma^2 + k\sigma_X^2} \sqrt{2\pi}} \exp \left( -\frac{(r_t - (\mu - \frac{\sigma^2}{2}))^2}{2(\sigma^2 + k\sigma_X^2)} \right) \right) \quad \text{où } \theta = (\mu, \sigma, \lambda, \sigma_X)$$

Avantages et inconvénients : Ce modèle présente, comme le modèle de Black-Scholes, l'avantage d'être relativement simple à utiliser tout en améliorant l'adéquation aux données. En effet, la prise en compte des sauts permet notamment de modéliser l'actif à l'aide d'une distribution à queue plus épaisse que la loi normale ce qui est important dans le calcul d'un capital réglementaire puisque nous nous intéressons, dans ce cas, aux événements extrêmes.

Le modèle ne corrige pas toutes les limites du modèle de Black-Scholes notamment concernant la volatilité supposée constante. Par ailleurs, la modélisation des sauts pourrait faire l'objet d'une étude plus approfondie sur le choix de la distribution de leur amplitude. L'hypothèse de gaussianité des sauts peut être remise en cause comme nous l'avons mentionné et une distribution non-symétrique avec une queue plus épaisse à gauche pourrait être plus adaptée.

## Le modèle de Garch

Le modèle de Garch (Generalized Autoregressive Conditional Heteroscedasticity), développé par Bollerslev[5] en 1986, permet de modéliser la volatilité non plus par une constante mais en supposant que la volatilité conditionnelle varie. En effet, on constate en pratique que la volatilité varie localement et que la volatilité de la veille influe sur celle d'aujourd'hui notamment en période d'emballement de marchés.

Pour sélectionner les paramètres du modèle de Garch qui s'adaptent le mieux à nos données, nous avons utilisé le critère d'information d'Akaike [1](AIC). Ce critère est défini par :

$$AIC = 2k - 2\ln(L)$$

où :

- $L$  : la vraisemblance du modèle
- $k$  : le nombre de paramètres du modèle

Ainsi, nous avons calculé l'AIC pour des modèles Garch(p,q),  $p = 1, 2, 3$ ,  $q = 1, 2, 3$  et nous avons finalement retenu le modèle Garch(1,1) qui minimise le critère AIC. Nous ne développerons dans la suite que le modèle théorique de Garch(1,1) qui sera celui dont nous nous servirons dans l'application pratique.

Les rendements logarithmiques vérifient l'équation suivante :

$$r_t = \lambda + \epsilon_t \sqrt{h_t(\theta)}$$

avec :

- $\lambda$  : la moyenne empirique des rendements ou encore la prime de risque
- $(\epsilon_t)_t$  : un bruit blanc gaussien
- $\theta = (\omega, \alpha, \beta)$  : les paramètres du modèle
- $h_t(\theta) = \mathbb{V}(r_t | r_{t-1}, \dots, r_0, \theta)$  : la variance conditionnelle des rendements

L'équation de la variance conditionnelle pour un modèle Garch(1,1) est donnée par :

$$h_t(\theta) = \omega + \alpha(r_{t-1} - \lambda)^2 + \beta h_{t-1}(\theta)$$

et la contrainte de stationnarité par :

$$\alpha + \beta < 1$$

Le modèle possède donc quatre paramètres qui sont estimés par maximum de vraisemblance.

La vraisemblance : La vraisemblance du modèle s'écrit :

$$L(r_1, \dots, r_T | \theta_G) = \prod_{t=1}^T \frac{1}{\sqrt{h_t(\theta)} \sqrt{2\pi}} \exp\left(-\frac{(r_t - \lambda)^2}{2h_t(\theta)}\right) \quad \text{où } \theta_G = (\omega, \alpha, \beta, \lambda)$$

Avantages et inconvénients : Ce modèle présente l'avantage d'être plus réaliste que les modèles présentés ci-dessus puisque les hypothèses de départ correspondent mieux aux observations faites en pratique. En particulier, la variance conditionnelle est non constante ce qui permet d'avoir des prévisions de volatilité plus fiables et réactives. De plus, la distribution des rendements est leptokurtique et non-gaussienne puisqu'elle est fortement concentrée autour de la moyenne mais que ses queues sont plus épaisses que celles d'une loi normale.

Le choix des résidus  $(\epsilon_t)_t$  gaussiens n'est pas toujours adapté notamment du fait de l'hypothèse de symétrie que cela implique. Un des reproches fait au modèle est son invariance résultant de trop fortes contraintes.

## Le modèle de Black-Scholes à changement de régimes markovien

Le modèle de Black-Scholes à changements de régime permet d'incorporer une volatilité stochastique au modèle de Black-Scholes classique. Le modèle suppose que la volatilité prend un nombre fini de valeurs discrètes qui seront appelées des régimes. Chaque régime possède ses propres paramètres (volatilité et dérive). Le changement de régime est aléatoire et est décrit par un processus markovien. Par conséquent, la probabilité de changer de régime ne dépend que du régime actuel et non des régimes passés.

Nous supposons ici qu'il existe deux régimes, un régime correspondant aux périodes de marché dites "calmes" et un régime associé aux périodes d'emballement des marchés pour lequel la volatilité sera plus élevé.

A chaque instant  $t = 1, \dots, T$ , le rendement est supposé gaussien conditionnellement au régime s'appliquant sur  $[t, t+1[$ , i.e.

$$r_t \mid e_t \sim \mathcal{N}(\mu_{e_t}, \sigma_{e_t})$$

où

- $e_t$  : le régime sur  $[t, t+1[$
- $\mu_{e_t}$  : la moyenne des rendements dans le régime  $e_t$
- $\sigma_{e_t}$  : l'écart-type des rendements dans le régime  $e_t$

La probabilité de passer du régime  $i = \{1, 2\}$  à  $j = \{1, 2\}$  est constante dans le temps et ne dépend pas du passé puisque le processus est markovien ; elle sera notée  $p_{i,j} = \mathbb{P}(e_{t+1} = j \mid e_t = i)$ .

### La vraisemblance :

Nous avons estimé les paramètres par le maximum de vraisemblance à l'aide de la démarche proposée par Hardy [23].

Tout d'abord, nous calculons la probabilité invariante définie par :

$$\pi P = \pi$$

avec

- $\pi = (\pi_1, \pi_2)^T$
- $P = (p_{i,j})_{i=1,2; j=1,2}$

Par ailleurs, comme  $p_{1,1} + p_{1,2} = 1$  et  $p_{2,1} + p_{2,2} = 1$ , on a :

$$\pi_1 = \frac{p_{2,1}}{p_{1,2} + p_{2,1}} \quad \text{et} \quad \pi_2 = \frac{p_{1,2}}{p_{1,2} + p_{2,1}}$$

Nous allons utiliser une méthode récursive. Quelques notations :

- $\Theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, p_{1,2}, p_{2,1})$
- $\phi$  désigne la fonction de densité d'une loi normale  $\mathcal{N}(0, 1)$
- $f$  désigne une fonction de densité

Déroulement des itérations :

- $t = 1$  :
- $$f(e_1 = 1, r_1 \mid \Theta) = \pi_1 \phi\left(\frac{r_1 - \mu_1}{\sigma_1}\right)$$
- $$f(e_1 = 2, r_1 \mid \Theta) = \pi_2 \phi\left(\frac{r_1 - \mu_2}{\sigma_2}\right)$$

Par conséquent, la densité conditionnelle du premier rendement s'écrit :

$$f(r_1 \mid \Theta) = f(e_1 = 1, r_1 \mid \Theta) + f(e_1 = 2, r_1 \mid \Theta)$$

et la distribution de probabilité du premier régime conditionnellement au rendement et aux paramètres, qui sera utilisée pour les itérations suivantes, est donnée par :

$$p(e_1 | r_1, \Theta) = \frac{f(e_1, r_1 | \Theta)}{f(r_1 | \Theta)}$$

—  $t = 2, \dots, T$  :

$$f(e_t = i, e_{t-1} = j, r_t | \Theta, r_1, \dots, r_{t-1}) = p(e_{t-1} = j | \Theta, r_1, \dots, r_{t-1}) p_{i,j} f(r_t | \Theta, e_t)$$

$$\text{où } p(e_{t-1} = j | \Theta, r_1, \dots, r_{t-1}) = \frac{f(e_{t-1}=j, e_{t-2}=1, r_{t-1} | \Theta, r_1, \dots, r_{t-2}) + f(e_{t-1}=j, e_{t-2}=2, r_{t-1} | \Theta, r_1, \dots, r_{t-2})}{f(r_{t-1} | \Theta, r_1, \dots, r_{t-2})}$$

La densité conditionnelle du t-ième rendement s'écrit alors :

$$f(r_t | \Theta, r_1, \dots, r_{t-1}) = \sum_{\{i, e_t=i\}\{j, e_{t-1}=j\}} f(e_t = i, e_{t-1} = j, r_t | \Theta, r_1, \dots, r_{t-1})$$

Ainsi, la vraisemblance du modèle est simplement :

$$L(r_1, \dots, r_T | \Theta) = \prod_{t=1}^T f(r_t | \Theta, r_1, \dots, r_{t-1})$$

Avantages et inconvénients : Ce modèle présente l'avantage d'autoriser la volatilité à être stochastique. Il est donc plus réaliste que le modèle de Black-Scholes.

Cependant, ce modèle est moins simple à mettre en œuvre notamment pour estimer les paramètres. Le choix des paramètres initiaux de volatilité et de probabilité de passage doivent faire l'objet d'une étude préalable. D'autre part, le choix du nombre de régimes peut considérablement changer les résultats. Si ce nombre n'est pas assez élevé, alors toutes les variations de volatilité ne seront pas captées par le modèle. Néanmoins, nous fixons ce nombre à deux car un modèle présentant plus de régimes est très difficile à mettre en application. Il est à noter que cette hypothèse est forte et peut impacter conséquemment les résultats.

## 5.2 Validation des Modèles

Avant de calculer l'erreur de modèle pour l'ensemble des modèles sélectionnés, nous avons procédé à la vérification des hypothèses des différents modèles. Afin d'évaluer si la distribution des prévisions obtenue avec chacun des modèles coïncide avec la vraie distribution, nous avons effectué un backtesting. Beaucoup de tests peuvent être utilisés à cet égard mais nous avons choisi de regarder le PIT (Probability Integral Transform) et de réaliser quelques tests d'adéquation usuels (Kolmogorov-Smirnov et Anderson-Darling).

### 5.2.1 Vérification des hypothèses

Nous avons procédé à plusieurs vérifications des hypothèses à l'aide d'outils graphiques essentiellement.

#### Le modèle de Black-Scholes

Le modèle de Black-Scholes suppose d'une part que les paramètres du modèle (dérive et volatilité) sont constants. Un moyen simple d'évaluer cette hypothèse est d'estimer les paramètres du modèle en commençant par le rendement de l'année 2001 et en augmentant progressivement la période considérée jusqu'aux rendements de 2001 à 2013. Les valeurs de ces paramètres sont mises en annexe [A.1.1](#). Nous constatons que les paramètres estimés ne sont pas constants. Le paramètre de dérive semble relativement stable à partir d'un seuil. Le manque de stabilité pour les périodes antérieures est expliqué par la faible quantité de données puisque nous augmentons progressivement la taille des rendements

sur lesquels sont effectués nos estimations. Au contraire, le paramètre de volatilité semble être plutôt stochastique. Ce deuxième constat était d'ailleurs un des principaux arguments lors de l'introduction des autres modèles dans la section précédente. Nous remarquons, par ailleurs, que les paramètres évoluent de manière opposée. Un rendement moyen faible est associé à une forte volatilité. Cela nous semble pertinent puisque, lorsque les marchés sont perturbés, la volatilité va augmenter et les rendements journaliers ont tendance à baisser (cf crise des subprimes).

Les graphiques ci-dessous présentent la moyenne et l'écart-type empiriques des rendements logarithmiques évalués pour chaque année de 2001 à 2013.

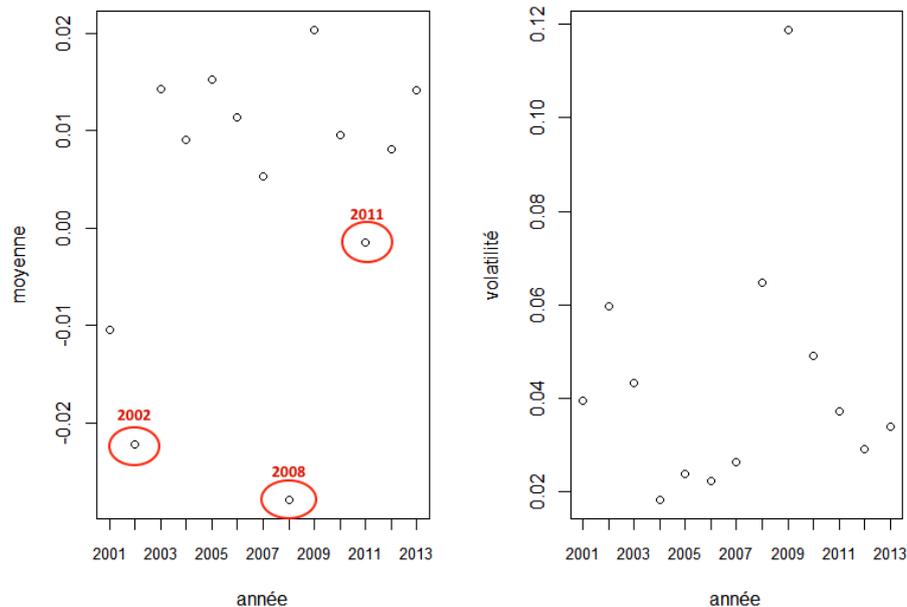


FIGURE 4 – Moyenne et volatilité des rendements logarithmiques mensuels par année de 2001 à 2013

Sur ces graphes, nous avons mis en avant 3 points correspondant aux années 2002, 2008 et 2011. Historiquement, ces années correspondent à trois périodes de crises financières notables. Suite à la faillite de Lehman Brothers le 15 septembre 2008, la confiance chute et les cours sur les différentes places boursières se déprécient dramatiquement. Ainsi les rendements sont très volatiles au lendemain de cette faillite et baissent de manière conséquente en moyenne. En 2001-2002, le krach boursier qui découle de la bulle internet de la fin des années 2000, est marqué par la faillite de Enron. Après quelques années de performances exceptionnelles, le Nasdaq atteint son niveau le plus bas. En 2011, ce sont les prémices de la crise souveraine européenne qui a affecté assez sévèrement les marchés financiers (en particulier les actions et la dette obligataire d'Etat) et qui continue actuellement à engendrer des périodes de troubles périodiquement.

Au niveau de la volatilité, il est plus difficile de voir clairement une hausse de volatilité pour ces trois années car cette hausse est lissée par la mensualisation des rendements. Pour bien observer l'impact des crises sur la volatilité, les graphes obtenus similairement sur données journalières (cf annexe A.1.1) sont plus explicites.

Par ailleurs, nous réalisons un graphique Quantile-Quantile (QQ-Plot) des rendements centrés et réduits à l'aide de la moyenne et de l'écart type estimés afin de vérifier l'hypothèse de gaussianité des rendements.

Sur le graphe 5, nous remarquons que les quantiles empiriques ne correspondent pas aux quantiles de la loi normale centrée réduite. En particulier, les quantiles présentent des différences notables sur les queues de distribution. Cela vient du fait que les rendements ont des queues plus épaisses que la loi normale à cause d'évènements extrêmes. En effet, les points les plus extrêmes (cf graphe ci-dessus) correspondent aux fortes variations (à la hausse et à la baisse) qui ont suivi la faillite de Lehman Brothers. Les autres points des extrémités (quantiles empiriques autour de -2 et 2) sont associés à la

période de crise de 2008 mais aussi au krach boursier de 2001-2002 et à la crise souveraine européenne de 2011. Par conséquent, le choix de la loi normale pour modéliser les rendements logarithmiques ne permet pas de prendre en compte les valeurs extrêmes et ne semble pas adéquat. Le test de Shapiro-Wilk qui permet de tester la normalité d'un échantillon, confirme cette constatation puisque l'hypothèse nulle est rejetée ( $p\text{-value}=2.7e-08$ ).

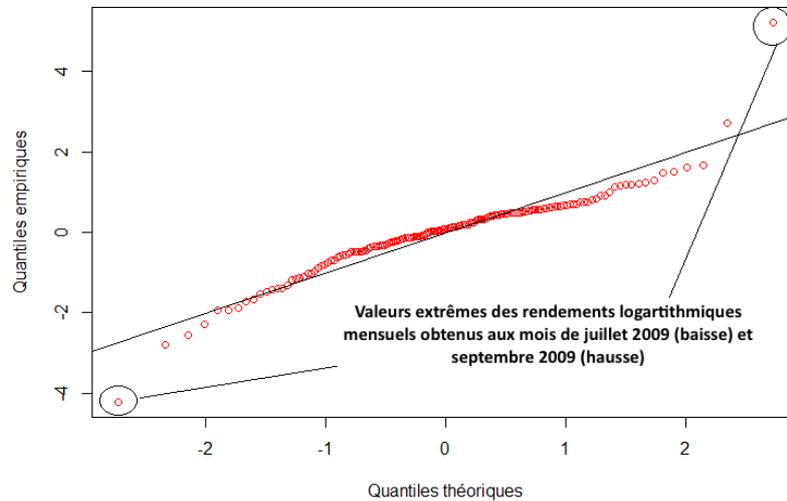


FIGURE 5 – Graphique Quantile-Quantile des rendements logarithmiques mensuels 2001-2013

Au vu de ces résultats, le modèle de Black-Scholes ne semble pas très adapté, nous constatons exactement sur nos données les défauts souvent reprochés à ce modèle. Nous garderons cependant ce modèle qui est encore très utilisé en pratique.

### Le modèle de Merton

Le modèle de Merton possède les mêmes hypothèses que le modèle de Black Scholes. Les paramètres estimés sont présentés en annexe [A.1.2](#).

Nous constatons, dans un premier temps, que la méthode d'estimation semble appropriée et ne donne pas de valeurs aberrantes pour les paramètres. En particulier,  $\sigma^2$ ,  $\lambda$ , et  $\sigma_X^2$  sont positifs et si nous supposons qu'il n'existe pas de sauts en fixant les paramètres du processus de Poisson composé à zéro, nous obtenons bien les mêmes paramètres  $\mu$  et  $\sigma$  que pour le modèle de Black-Scholes.

Par ailleurs, les volatilités, comme pour le modèle de Black-Scholes, ne semblent pas constantes au cours du temps ce qui remet en cause l'hypothèse de stabilité des paramètres.

Pour finir, la fréquence des sauts n'est pas très élevée ce qui correspond bien à la constatation que nous pouvons faire sur le graphique représentant le cours de l'actif.

### Le modèle de Garch

Le modèle de Garch est basé sur l'hypothèse que la volatilité d'un jour influe sur la volatilité du lendemain. En d'autres termes, cela implique que les carrés des rendements boursiers sont corrélés et en particulier, qu'ils possèdent une forte corrélation entre deux dates successives. Le graphe [6](#) suivant permet d'affirmer cette hypothèse sur nos données.

Cette constatation est moins évidente lorsque l'on s'intéresse aux rendements mensuels (cf graphe [7](#)) puisqu'il ne s'agit pas de dates successives ; il y a cependant bien une dépendance.

Par ailleurs, comme pour les autres modèles, les paramètres estimés en augmentant progressivement l'historique de données sont présentés en annexe [A.1.3](#). Nous constatons essentiellement que la contrainte de stationnarité est vérifiée à chaque période.

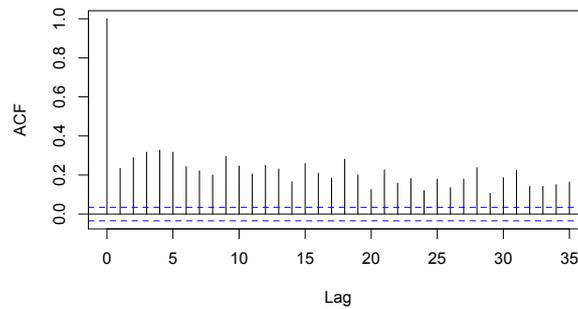


FIGURE 6 – Autocorrélation des carrées des rendements logarithmiques journaliers sur 2001-2013

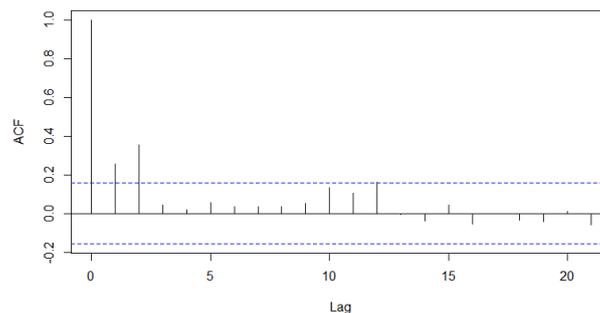


FIGURE 7 – Autocorrélation des carrées des rendements logarithmiques mensuels sur 2001-2013

Pour finir, l'intérêt majeur du modèle de Garch est de tenir compte des variations temporelles de la volatilité. Nous avons donc regardé si la volatilité prédite était proche de celle observée. Pour cela, à l'aide des paramètres estimés, nous avons effectué des prévisions sur les valeurs que prendraient les rendements s'ils étaient issus d'un modèle de Garch ayant pour paramètres les paramètres estimés. Ainsi, nous obtenons un intervalle de confiance à 95% des rendements prédits par le modèle que nous confrontons avec les valeurs historiques de rendements. Nous obtenons le graphe 8 ci-dessous sur les données journalières.

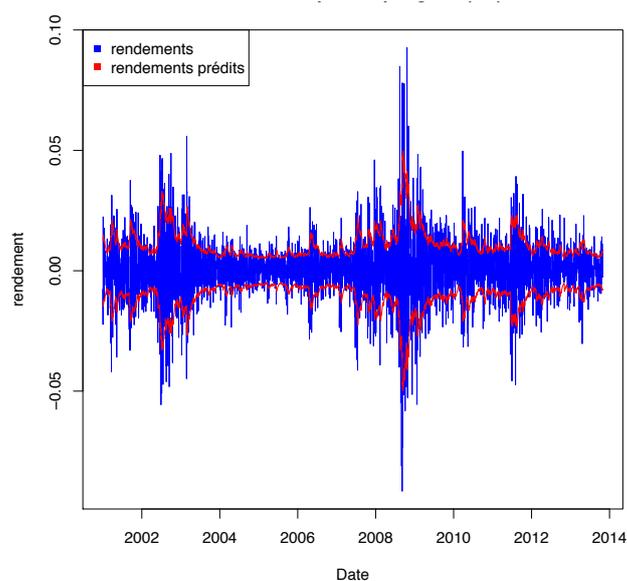


FIGURE 8 – Prévisions des rendements logarithmiques par le modèle de Garch en journalier

Nous remarquons que le modèle capte les fluctuations de la volatilité même s'il ne réussit pas à modéliser les grands écarts de volatilité en particulier ceux ayant eu lieu au cœur de la crise, en 2008-2009.

Sur les données mensuelles, les fluctuations sont moins bien prises en compte. Le modèle de Garch semble être réactif aux variations à court terme et moins sur des fréquences plus grandes (cf graphe 9 ci-dessous).

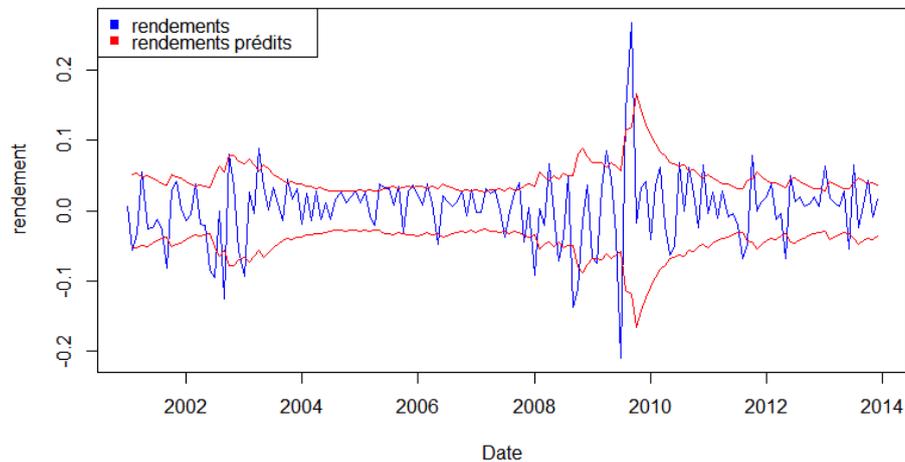


FIGURE 9 – Prévisions des rendements logarithmiques par le modèle de Garch en mensuel

### Le modèle de Black-Scholes à changements de régime

Les paramètres initiaux ont été choisis à l'aide des graphiques de la moyenne et de la volatilité des rendements logarithmiques en départageant deux types de périodes, périodes dites "calmes" et périodes de trouble. Le choix de la probabilité de passage a été effectué à l'aide de nos connaissances sur les cycles financiers.

Tout d'abord, nous présentons le graphique 10 de la prévision faite par le modèle de Black-Scholes à changements de régime pour l'année 2014 que nous confrontons avec les valeurs des rendements effectivement observées sur l'année 2014.

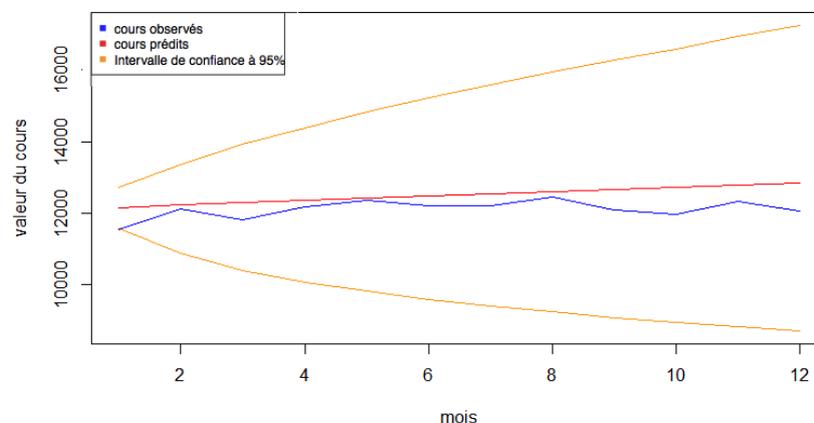


FIGURE 10 – Prévisions du cours boursier pour l'année 2014 par le modèle de BS-RS

Nous remarquons que la prévision effectuée reste, en moyenne, assez éloignée des rendements observés. Cependant, les rendements sont bien dans l'intervalle de confiance à 95%. En moyenne, les prévisions effectuées à l'aide de ce modèle ont tendance à surestimer la valeur du cours.

Une hypothèse de ce modèle est qu'il existe deux états de la nature, un état avec faible volatilité et un état avec plus grande volatilité correspondant aux périodes de troubles. Cependant, le graphique des rendements nous laisse penser qu'il y aurait plus de deux états puisque toutes les périodes à plus forte volatilité ne sont pas identiques. En effet, entre 2002 et 2003, nous observons une plus forte volatilité mais qui est bien moins importante que le pic observé vers la fin de l'année 2008 au moment de la crise des *subprimes*. Dans notre modèle, les deux périodes sont modélisées avec la même volatilité "haute" alors qu'il faudrait différencier les volatilités de ces deux périodes. Cependant, la modélisation avec plus d'états s'avère difficile à mettre en place en pratique.

### 5.2.2 Probability Integral Transform (PIT)

Le PIT est une notion introduite en 1932 par Fisher [19] qui utilise un théorème que nous rappelons ci-dessous.

*Rappel.* Si une variable aléatoire  $X$  a une fonction de répartition continue  $F(\cdot)$ , alors la variable aléatoire  $U = F(X)$  présente une répartition uniforme sur l'intervalle  $[0, 1]$ , c-à-d  $U \sim \mathcal{U}[0, 1]$ .

Le PIT est particulièrement utilisé pour tester si la distribution estimée correspond bien à la vraie densité. Ce test a été introduit par Diebold[13].

Soit  $(r_t)_{t=1}^d$  les rendements journaliers sur la période de taille  $d$ . Ces rendements ont pour "vraie" fonction de répartition  $(F_t(\cdot | \Omega_t))_{t=1}^d$  conditionnellement à l'historique  $\Omega_t = (r_{t-1}, \dots, r_1)$ . On note  $(\hat{H}_t(\cdot | \Omega_t))_{t=1}^d$  la fonction de répartition associée au modèle choisi.

Le PIT est défini par :

$$z_t = \hat{H}_t(r_t | \Omega_t)$$

Par le théorème précédent, cette variable suit une loi uniforme ;  $z_t \sim \mathcal{U}[0, 1]$ ,  $\forall t \in \{1, \dots, d\}$ . Attention, cette condition n'est pas suffisante pour dire que  $(z_1, \dots, z_d)$  est un échantillon car nous n'avons pas la propriété d'indépendance.

De plus, par la formule de Bayes, nous avons :

$$\hat{H}_t(r_t, \dots, r_1 | \Omega_1) = \hat{H}_t(r_t | \Omega_t) \hat{H}_{t-1}(r_{t-1} | \Omega_{t-1}) \dots \hat{H}_1(r_1 | \Omega_1)$$

et par conséquent

$$\hat{H}_t(r_t, \dots, r_1 | \Omega_1) = z_t z_{t-1} \dots z_1$$

La distribution jointe étant égale au produit des distributions marginales, il y a indépendance entre les variables  $z_t$ . Ainsi :

$$(z_t)_{t=1}^d \stackrel{iid}{\sim} \mathcal{U}[0, 1]$$

En pratique : Nous cherchons à tester l'hypothèse d'indépendance et d'uniforme distribution des variables issues du PIT, i.e. :

$$H_0 : (z_t)_{t=1}^d \stackrel{iid}{\sim} \mathcal{U}[0, 1], \quad \forall d \in \{1, \dots, T\}$$

$$\text{vs : } H_1 : \exists d \in \{1, \dots, T\}, \quad (z_t)_{t=1}^d \not\stackrel{iid}{\sim} \mathcal{U}[0, 1]$$

Pour ce faire, nous allons construire les vecteurs de variables  $(z_t)_{t=1}^d$  pour plusieurs valeurs de  $d$  et ce pour chaque modèle. Nous allons ensuite, à l'aide de plusieurs méthodes, tester notre hypothèse.

### Procédure :

Il existe, tout d'abord, des tests d'adéquation connus tels que le test de Kolmogorov-Smirnov ou le test d'Anderson-Darling. Nous constatons que ces tests manquent de clarté quant à la raison du rejet de l'hypothèse nulle d'adéquation de la distribution si tel est le cas. En effet, il est difficile de savoir si l'hypothèse violée dans ce cas est l'hypothèse d'uniformité ou le caractère i.i.d. du vecteur. Le test d'Anderson-Darling nous semble plus pertinent dans la mesure où il s'intéresse à l'ensemble de la distribution tout en donnant plus de poids aux observations sur les queues de distribution. Ainsi ce test prend bien en compte les observations extrêmes pour vérifier l'adéquation de la loi. Nous avons pu observer auparavant que les rendements présentaient des valeurs extrêmes et il est important que la distribution estimée par le modèle sélectionné soit adaptée sur les queues. Le test de Kolmogorov-Smirnov a tendance à ne considérer que l'écart maximal entre la fonction de répartition empirique et la fonction de répartition simulée et ne prend donc pas en compte l'ensemble de la distribution.

Pour pallier cette difficulté pratique, nous utilisons la méthode du test PIT proposée par Diebold qui est essentiellement graphique. Dans un premier temps, nous réalisons des statistiques descriptives sur le vecteur pour vérifier l'uniformité des données. Pour ce faire, nous réalisons un histogramme des variables estimées que nous comparons avec l'histogramme d'une variable uniforme. Nous comparons de même la distribution empirique de nos variables avec celle de la loi uniforme. Ainsi, nous pouvons, au vu de ces résultats, rejeter ou non l'hypothèse d'uniformité des variables.

Par la suite, nous nous intéressons à l'indépendance des variables. Pour tester l'indépendance, nous regardons la corrélation entre les variables à l'aide des fonctions d'autocorrélation. Nous commençons par la corrélation linéaire, c-à-d nous étudions la corrélation des variables  $((z_t - \bar{z}))_{t=1}^d$  où  $\bar{z} = \frac{1}{d} \sum_{t=1}^d z_t$ . Cependant, nous voulons aussi évaluer la corrélation à des ordres supérieurs pour assurer qu'il n'y ait pas de corrélation plus sophistiquée sur notre série. Diebold a montré que l'étude des corrélations aux quatre premiers ordres  $((z_t - \bar{z}))_{t=1}^d$ ,  $((z_t - \bar{z})^2)_{t=1}^d$ ,  $((z_t - \bar{z})^3)_{t=1}^d$  et  $((z_t - \bar{z})^4)_{t=1}^d$  permettait de prendre en compte les dépendances impliquées par des erreurs de moyenne, de variance, d'asymétrie et d'aplatissement conditionnels de la distribution prédite.

### Résultats :

Nous avons, d'une part, créé 13 vecteurs de rendements correspondant respectivement aux rendements de l'année 2001 à l'année 2001+k,  $k=\{0, \dots, 12\}$ . Pour chaque vecteur et chaque modèle  $M_i$ ,  $i = 1, \dots, 4$ , nous avons estimé les paramètres  $\hat{\theta}_{M_i}$  puis évalué la fonction de répartition associée aux valeurs prises par les rendements de l'année 2001+k+1.

$$z_{k+1}^{M_i} = \hat{H}_{(k+1, \hat{\theta}_{M_i})}(r_{2001+k+1} | \Omega_{k+1}) \quad \text{où} \quad \Omega_t = (r_{2001}, \dots, r_{2001+k}).$$

Pour chaque vecteur  $(z_k^{M_i})_{k=1}^{13}$ , nous dressons l'histogramme et la fonction de répartition empirique. Nous effectuons aussi les quatre autocorrélogrammes correspondant aux quatre premiers moments centrés.

Cependant, nous ne présenterons pas ces résultats qui sont trop nombreux et que nous avons choisi de regrouper. En effet, le vecteur  $(z_1^{M_i}, \dots, z_{13}^{M_i})$  doit aussi être un vecteur uniforme et i.i.d. Ainsi pour chaque modèle nous présentons les résultats pour ce vecteur et les statistiques de test de Kolmogorov-Smirnov et d'Anderson-Darling seront effectuées sur ce vecteur.

Les histogrammes, fonctions de répartition empiriques et autocorrélogrammes obtenus pour les modèles de Black-Scholes, Merton et Garch sont présentés en annexe A.2, nous montrons ici ceux réalisés pour le modèle de Black-Scholes à changements de régime (cf graphes 11 et 12).

### Analyse des résultats :

Au vu des histogrammes, nous remarquons qu'aucun des modèles ne semble être très adapté puisque sinon les histogrammes seraient uniformes.

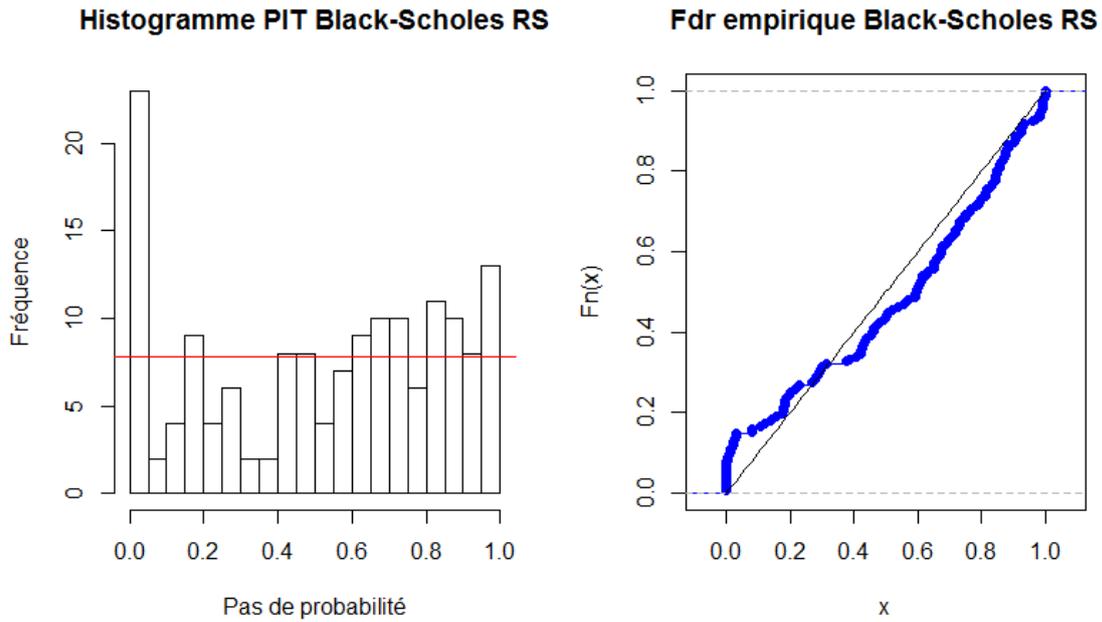


FIGURE 11 – PIT modèle de Black-Scholes à changements de régime

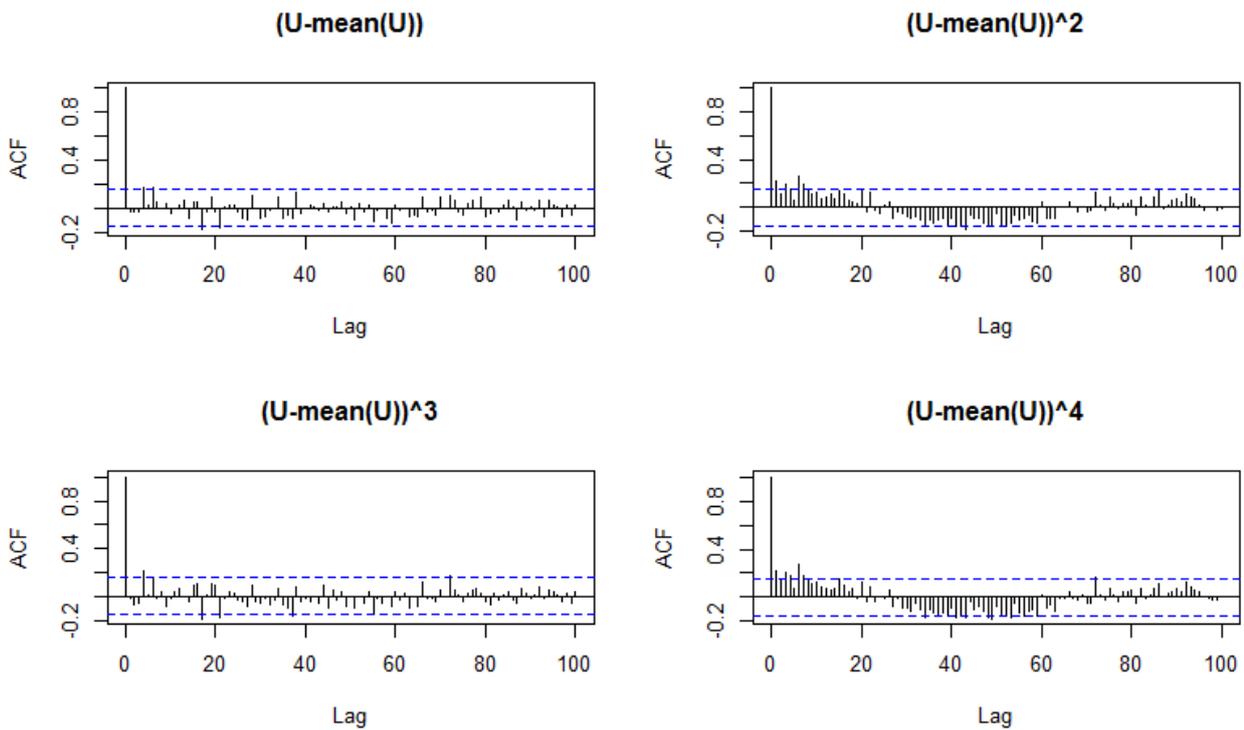


FIGURE 12 – Autocorrélogrammes modèle de Black-Scholes à changements de régime

Ces constatations se retrouvent avec la fonction de répartition empirique. En effet, la fonction de répartition empirique obtenue avec le modèle de Garch est proche de la fonction de répartition de la loi uniforme qui est la fonction linéaire de pente 1. Pour les autres modèles, la fonction empirique n'est pas uniforme notamment aux extrémités pour le modèle de Black-Scholes à changements de régime et autour de 0.5 pour les deux autres modèles.

Pour finir, nous analysons les autocorrélogrammes des quatre modèles pour les quatre premiers moments centrés. Ainsi, nous pouvons remarquer que le modèle de Garch ne présente pas ou peu de corrélation aux différents ordres. Cela signifie que le modèle arrive bien à capter la moyenne, la variance, l'asymétrie et l'aplatissement de la distribution des rendements conditionnellement à l'historique des données. Ce n'est pas le cas pour les autres modèles qui présentent notamment une forte corrélation à l'ordre 2. Par conséquent, ces modèles n'arrivent pas à modéliser la variance conditionnelle. Ce résultat est cohérent avec la description des modèles puisque pour les modèles de Black-Scholes et de Merton la volatilité est supposée constante. Or, en pratique, ce n'est pas ce qui est observé et donc la volatilité estimée pour ces modèles n'est pas en accord avec la "vraie" volatilité des rendements. Pour le modèle de Black-Scholes à changements de régime, il y a aussi de la corrélation à l'ordre 2 et 4 mais légèrement plus faible. Ce modèle aurait peut-être mieux capté la volatilité si nous avions supposé la présence de plus de deux régimes.

Au vu du test de PIT, nous retenons clairement le modèle de Garch. Il est plus difficile de différencier les autres modèles même si le modèle de Merton semble mieux que celui de Black-Scholes et celui de Black-Scholes à changements de régime semble mieux que les deux précédents.

#### Les tests statistiques :

Afin que le résultat du test statistique ne soit pas que purement graphique, nous avons réalisé quelques tests statistiques classiques. Pour tester notre hypothèse, nous procédons à deux types de test au même titre que les analyses graphique, un test d'adéquation de nos données à la loi uniforme et un test d'indépendance.

Pour vérifier l'uniformité de nos variables PIT, nous avons effectué un test de Kolmogorov-Smirnov et un test de Anderson-Darling sur ces variables. Ces tests ne sont pas directement comparables. Les résultats sont présentés dans le tableau 1 suivant :

TABLE 1 – Tests d'adéquation

		KS	AD
BS	Statistique	0.14768	4.355
	p-value	0.002217	0.005878
Merton	Statistique	0.1049	2.5843
	p-value	0.08409	0.04483
Garch	Statistique	0.085195	2.2939
	p-value	0.302	0.06377
BS à RS	Statistique	0.11474	9.676
	p-value	0.0329	1.585e-05

Les p-values des deux tests sont inférieures à 0.05 pour les modèles de Black-Scholes et de Black-Scholes à changements de régime. Par conséquent, l'hypothèse nulle  $H_0$ , à savoir l'hypothèse d'uniformité des variables du PIT, est donc rejetée pour ces deux modèles. Cependant, la p-value du test de Kolmogorov-Smirnov pour le modèle de Black-Scholes à changements de régime n'est pas loin de 5% ce qui pourrait nous laisser penser que ce dernier est plus adapté que le modèle de Black-Scholes. Concernant les deux autres modèles, le test de Kolmogorov-Smirnov ne permet pas de rejeter l'hypothèse nulle et le test d'Anderson-Darling ne permet pas de rejeter cette hypothèse nulle pour un seuil légèrement inférieur à 5%. Ainsi, nous considérons que l'uniformité des variables aléatoires issues du PIT est vérifiée pour le modèle de Garch et de Merton.

Pour tester l'indépendance des variables PIT, nous utilisons le test de Ljung-Box aussi connu sous le nom de test du portemanteau (cf tableau 2). Ce test évalue la corrélation entre les résidus et nous avons choisi de réaliser ce test pour un lag (retard) égal à 1 mais pour différents moments centrés, les mêmes que ceux proposés par Diebold.

TABLE 2 – Test d'indépendance

	Ljung-Box	1er ordre	2ième ordre	3ième ordre	4ième ordre
BS	Statistique	0.0052708	12.438	2.1143	12.318
	p-value	0.9421	0.0004208	0.1459	0.0004486
Merton	Statistique	0.43356	10.045	0.40548	10.973
	p-value	0.5102	0.001528	0.5243	0.0009245
Garch	Statistique	0.0019793	1.2572	0.0048778	0.63134
	p-value	0.9645	0.2622	0.9443	0.4269
BS-RS	Statistique	0.091889	7.941	0.024837	7.5274
	p-value	0.7618	0.004833	0.8748	0.006077

Le test d'indépendance confirme les constatations graphiques faites sur les ACF. An effet, l'hypothèse d'indépendance n'est pas rejetée au 1er et 3ième ordre pour les 4 modèles. Cependant, elle est rejetée pour tous les modèles sauf le modèle de Garch au 2ième et 4ième ordre. Ainsi, seul le modèle de Garch arrive à prendre en compte la variance conditionnelle de nos données (par définition du modèle) et permet donc un meilleur ajustement.

Au vu des statistiques de test, le modèle de Garch semble clairement le meilleur des quatre. Le choix de ce modèle comme référence n'est donc pas aberrant. Le modèle de Merton arrive à s'adapter plutôt bien aux données mais capte difficilement les dépendances des données. Les deux autres modèles ne semblent pas adaptés à nos données.

### 5.2.3 Procédure

Nous rappelons tout d'abord les différentes étapes de la mise en œuvre de l'approche bayésienne en les adaptant au cadre d'application.

#### Etapes de l'approche bayésienne :

- I. Choix de la classe de modèles :  $M = (M_1, M_2, M_3, M_4)$ . La classe de modèle choisie consiste donc en quatre modèles : Black-Scholes ( $M_1$ ), Merton ( $M_2$ ), Garch ( $M_3$ ) et Black-Scholes à changements de régime ( $M_4$ ). Le modèle de référence qui sera utilisé pour le calcul de l'erreur de modèle est le modèle de Garch.
- II. Choix de l'a priori sur le modèle  $p(M_i)$  L'a priori pour nos modèles sera  $p(M) = (p(M_1), p(M_2), p(M_3), p(M_4)) = (24\%, 24\%, 28\%, 24\%)$ . Le choix de cet a priori est issu à la fois des tests effectués précédemment et des connaissances théoriques que nous avons des modèles.
- III. Estimation des paramètres  $\theta_i$  pour chaque modèle  $M_i$ .
- IV. Calcul de la VaR à 0.5% : Le risque actif est un risque de baisse significative du prix de l'actif. Par conséquent, on s'intéresse au quantile à 0.5% de la distribution du cours de l'actif à un an, soit fin  $T+1$ , notée  $X_{T+1}$ . Ainsi, nous calculons pour chaque  $i$ ,

$$VaR_{0.5\%(M_i, \theta_i)}(X_{T+1})$$

- V. Détermination de l'a posteriori de chaque modèle  $p(M_i | r_1, \dots, r_T)$  :

$$p(M_i | r_1, \dots, r_T) = \frac{L(r_1, \dots, r_T | \theta_i)p(M_i)}{\sum_j L(r_1, \dots, r_T | \theta_j)p(M_j)}$$

avec  $(r_1, \dots, r_T)$  l'historique de rendements et  $L(r_1, \dots, r_T | \theta_i)$  la vraisemblance du modèle  $i$ .

**Simulation par Bootstrap :** Nous effectuons  $N$  tirages par Bootstrap parmi nos quatre modèles. Chaque modèle  $i$  est tiré avec probabilité  $p(M_i | r_1, \dots, r_T)$ . Une fois le modèle tiré, nous estimons ses paramètres, projetons une valeur à horizon 1 an. Ainsi nous obtenons  $N$  valeurs projetées à 1 an provenant de différents modèles. Sur cette distribution, nous évaluons la  $VaR_{0.5\%}$ . En répétant cette opération  $I$  fois afin de veiller à la stabilité des résultats, nous obtenons finalement  $I$  Value-at-Risk ( $VaR_{0.5\%(M_n, \theta_n)}(X_{T+1})_{n=1, \dots, I}$  où  $M_n \in (M_1, M_2, M_3, M_4)$ ). Nous quantifions le risque associé à cet ensemble de modèle par

$$\mathcal{R}_{Bootstrap}(M) = \frac{1}{I} \sum_{n=1}^I VaR_{0.5\%(M_n, \theta_n)}(X_{T+1}).$$

L'erreur de modèle pour l'approche bayésienne est définie par :

$$\rho_-(M) = 1 - \frac{\mathcal{R}_{Bootstrap}(M)}{\mathcal{R}(M^*)}$$

où  $\mathcal{R}(M^*) = \frac{1}{I} \sum_{n=1}^I VaR_{0.5\%(M^*, \theta^*)}(X_{T+1})$ .

En effet, nous nous intéressons, dans le cas d'un choc d'actif, à l'étude d'un choc à la baisse. Dans le cas de l'évaluation d'une baisse de l'actif, l'erreur de modèle provient du fait de choisir un modèle de référence pour lequel le quantile à 0.5% est plus élevé que le "vrai" quantile à 0.5%. Il faut donc être attentif aux événements extrêmes étudiés, en particulier il faut savoir si l'on s'intéresse aux événements présents dans la queue gauche ( $q_{0.5\%}$ ) ou ceux présents dans la queue droite ( $q_{99.5\%}$ ).

La méthode du Bootstrap permet de prendre en compte directement l'erreur liée au choix du modèle dans la quantification de la mesure de risque VaR. En effet, les facteurs de risques sont projetés à l'aide de modèles sélectionnés selon leur probabilité à être le vrai modèle.

#### 5.2.4 Résultats

Pour chaque modèle, nous avons évalué la vraisemblance dont les formules ont été présentées précédemment. Plus précisément, nous avons maximisé la log-vraisemblance.

TABLE 3 – Vraisemblance de chaque modèle

	BS	Merton	Garch	BSRS
Log-vraisemblance	244.4669	259.4574	271.0486	267.729
Vraisemblance	1.48e+106	4.80e+112	5.19e+117	1.88e+116

Ainsi, nous obtenons les probabilités *a posteriori* résumées dans le tableau 4 ci-dessous.

TABLE 4 – Probabilités *a posteriori*

	B&S	Merton	Garch	B&S RS
Nbre de paramètres	2	4	4	6
$\mathbb{P}(M_i)$	24%	24%	28%	24%
$p(x   M_i)$	244.4669	259.4574	271.0486	267.729
$\mathbb{P}(M_i   r_1, \dots, r_T)$	2.37e-12	7.69e-06	96.99%	3.01%

Nous constatons que le modèle de Garch est très nettement favorisé par rapport aux autres modèles. En effet, c'est le modèle qui possède à la fois la plus grande vraisemblance et la plus grande probabilité *a priori*. Ainsi le modèle sera nécessairement davantage sélectionné lors du bootstrap et la mesure de risque VaR obtenue avec le bootstrap devrait être assez proche de celle obtenue en ne sélectionnant que le modèle Garch, c-à-d la VaR de référence.

Le tableau ci-dessous présente la valeur du quantile à 0.5% de la distribution prédite à 1 an par la méthode bayésienne. Nous présentons ici ce quantile pour évaluer les événements extrêmes présents dans la queue gauche de la distribution, soit les événements correspondant à une importante baisse de l'actif ne se produisant qu'une fois tous les 200 ans. Nous aurions également pu étudier les événements extrêmes présents dans la queue droite de la distribution (hausse substantielle de l'actif) grâce au quantile à 99.5%. Cependant, une très forte hausse de l'actif ne présente pas un risque en soi-même si cela peut indiquer la formation d'une bulle boursière ayant un risque d'implosion et consécutivement de fortes baisses de l'actif.

*Remarque (CoV).* Nous évaluons le coefficient de variation (CoV) défini par

$$CoV = \frac{\sigma}{|\mu|}$$

avec

- $|\mu|$  : la valeur absolue de la moyenne
- $\sigma$  : l'écart-type

Ce coefficient permet de mesurer la dispersion relative de la distribution de probabilité.

TABLE 5 – VaR à 0.5% sur 100 itérations de distributions de taille 20 000

<i>Bootstrap</i>	$q_{0.5\%}$
<b>moyenne VaR</b>	6910.148
<b>écart-type VaR</b>	64.01
<b>CoV VaR</b>	0.93%

Pour finir, nous n'avons plus qu'à calculer l'erreur de modèle grâce à la formule ci-dessus. Nous mesurons aussi le choc à la baisse à 0.5% sur les actifs défini par :

$$Choc_{\alpha\%} = \frac{VaR_{\alpha\%}}{S_0} - 1$$

où  $S_0$  : la dernière valeur du cours de l'année T.

La mesure de risque de baisse de l'actif obtenue avec le modèle de référence Garch est  $\mathcal{R}(M^*)$ .

TABLE 6 – Erreur de Modèle sur des distributions de taille 20000 avec estimation sur 100 itérations

<i>Bootstrap</i>	$q_{0.5\%}$
<b>Choc 0.5%</b>	-42.37%
$\mathcal{R}(M^*)$	6766.885
<b>Erreur=<math>\rho(M)</math></b>	-2.12%

Nous remarquons que l'erreur de modèle est négative. En effet, cela provient du fait que le modèle de Garch possède une queue gauche de distribution plus épaisse que celle des autres modèles et que celle de la distribution obtenue par bootstrap. En d'autres termes, le quantile à 0.5% de la distribution à 1 an issue du modèle de Garch est plus petit que celui de la distribution du bootstrap. Par conséquent, le modèle de Garch est déjà très prudent et il n'y a donc pas d'erreur de modèle en choisissant ce dernier comme référence et donc pas de nécessité de rajouter du capital pour prendre en compte ce risque.

### 5.2.5 Tests de sensibilité

Afin de nous assurer que les résultats obtenus ne sont pas trop influencés par le choix de l'*a priori*, nous avons réitéré les mêmes calculs pour quatre *a priori*. Ces *a priori* restent relativement proches de ceux choisis initialement car nous les choisissons toujours au regard des résultats des tests d'adéquation effectués au préalable. Les *a priori* retenus sont donc :

- $AP_1$  :  $\mathbb{P}(M_i) = (25\%, 25\%, 25\%, 25\%)$ . On suppose dans ce cas que les quatre modèles sont équidistribués.
- $AP_2$  :  $\mathbb{P}(M_i) = (20\%, 20\%, 40\%, 20\%)$ . Dans ce cas, le modèle de Garch est clairement privilégié et les autres modèles ont tous la même probabilité d'être sélectionnés.
- $AP_3$  :  $\mathbb{P}(M_i) = (0\%, 20\%, 35\%, 35\%)$ . Nous choisissons d'enlever le modèle de Black-Scholes qui, bien que très utilisé, est loin d'être réaliste.
- $AP_4$  :  $\mathbb{P}(M_i) = (20\%, 20\%, 20\%, 40\%)$ . Le modèle de Black-Scholes à changements de régime présentant une vraisemblance élevée et des résultats au PIT acceptables, nous décidons de le privilégier puisque les probabilités *a posteriori* pourraient être modifiées et donc la distribution issue du bootstrap aussi.

Les distributions *a posteriori* sont les suivantes :

TABLE 7 – Distributions *a posteriori*

		B&S	Merton	Garch	B&S RS
$AP_1$	$\mathbb{P}(M_i)$	25%	25%	25%	25%
	$\mathbb{P}(M_i   r_1, \dots, r_T)$	2.76e-12	8.92e-06	96.51%	3.49%
$AP_2$	$\mathbb{P}(M_i)$	20%	20%	40%	20%
	$\mathbb{P}(M_i   r_1, \dots, r_T)$	1.40e-12	4.54e-06	98.22%	1.78%
$AP_3$	$\mathbb{P}(M_i)$	0%	30%	35%	35%
	$\mathbb{P}(M_i   r_1, \dots, r_T)$	0%	7.65e-06	96.51%	3.49%
$AP_4$	$\mathbb{P}(M_i)$	20%	20%	20%	40%
	$\mathbb{P}(M_i   r_1, \dots, r_T)$	2.66e-12	8.62e-06	93.33%	6.75%

Le tableau 8 suivant illustre les quantiles à 0.5% obtenus pour chaque cas sur des distributions de taille 20000 avec estimation sur 100 itérations :

TABLE 8 – Quantiles à 0.5%, Chocs à 0.5% et Erreur de modèle associée

	$AP_1$	$AP_2$	$AP_3$	$AP_4$
moyenne VaR(0,5%)	6901.669	6902.089	6912.744	6933.152
écart-type VaR	69.23	72.44	68.33021	67.16552
CoV VaR	1.003%	1.050	0.988	0.969
Choc 0.5%	-42.44%	-42.35%	-42.18%	-42.176%
Erreur= $\rho(M)$	-1.99%	-1.20%	-2.16%	-2.46%

Nous observons que quel que soit l'*a priori* l'erreur de modèle est toujours négative ce qui est évident dans la mesure où le modèle de Garch est le plus prudent. En effet, quel que soit le modèle sélectionné parmi les quatre, la distribution à 1 an associée sera moins ou aussi prudente que celle obtenue à partir du modèle de Garch. Nous remarquons que plus nous donnons de poids *a priori* aux autres modèles que celui de Garch, plus la Value-at-Risk augmente ce qui s'explique par le fait que les autres modèles sont moins prudents comme nous l'avons observé dans l'approche *worst case*.

L'intérêt de cette étude de sensibilité est principalement d'évaluer si la méthode est stable et donc si les résultats obtenus varient beaucoup ou non en fonction de l'*a priori* choisi. Les variations sont faibles ici puisqu'elles ne dépassent pas 2.46% (en valeur absolue) par rapport au modèle de Garch. La log-vraisemblance du modèle de Garch étant sensiblement plus grande que les autres, la pondération du modèle de Garch *a posteriori* sera toujours prépondérante par rapport aux autres modèles, l'*a priori* n'ayant quasiment aucun impact sur l'*a posteriori*.

### 5.3 Comparaison des approches Bayésienne et *worst case*

#### 5.3.1 Approche *worst case*

Pour l'approche *worst case*, il est nécessaire d'évaluer, pour chaque modèle, la distribution du cours de l'actif à 1 an afin d'obtenir la Value-at-Risk propre à chaque modèle. Pour cela, la méthode est similaire à celle du Bootstrap excepté qu'au lieu de tirer un modèle parmi quatre à chaque itération, nous tirons toujours le même modèle. Nous avons déjà calculé celle du modèle de Garch pour la partie précédente et nous faisons de même avec les trois autres modèles. Les résultats sont présentés ci-dessous :

TABLE 9 – VaR à 0,5% et erreur de modèle sur 200 itérations des distributions de taille 50000

	Moyenne	Ecart-type	CoV	Choc
<b>BS</b>	7956,54	30,44718	0,383%	-33,640%
<b>Merton</b>	7737,512	35,73471	0,462%	-35,467%
<b>Garch</b>	6766.885	34.66013	0.512%	-43.563%
<b>BSRS</b>	7737,700	37,77645	0,488%	-35,466%

De même que pour la mesure de l'erreur de modèle dans l'approche bayésienne, les différentes mesures de l'erreur de modèle dans le cas de l'approche *worst case* utilisées pour le risque de baisse de l'actif sont :

$$AM_- = 1 - \frac{\min_{M_i \in M} \mathcal{R}(M_i)}{\mathcal{R}(M^*)} = 1 - \frac{\min_{M_i \in M} VaR_{0.5\%}(M_i)}{VaR_{0.5\%}(M^*)}$$

$$RM_- = \frac{\mathcal{R}(M^*) - \min_{M_i \in M} \mathcal{R}(M_i)}{\max_{M_i \in M} \mathcal{R}(M_i) - \min_{M_i \in M} \mathcal{R}(M_i)} = \frac{VaR_{0.5\%}(M^*) - \min_{M_i \in M} VaR_{0.5\%}(M_i)}{\max_{M_i \in M} VaR_{0.5\%}(M_i) - \min_{M_i \in M} VaR_{0.5\%}(M_i)}$$

L'erreur de modèle pour l'approche *worst case* est donc nulle quelque soit la mesure utilisée puisque le modèle de Garch (faisant référence) est le modèle déjà le plus prudent.

Le tableau suivant illustre l'erreur de modèle dans le cas où chacun des trois autres modèles seraient choisis comme référence.

TABLE 10 – Erreur de modèle pour l'approche *worst case*

Modèle de référence	B&S	Merton	Garch	B&S RS
AM	13,55%	11,11%	0%	11,11%
RM	100%	79.688%	0%	79.706%

Le tableau 10 précédent illustre l'erreur de modèle dans le cas où chacun des trois autres modèles seraient choisis comme référence. Ainsi, le choix du modèle de Garch n'implique pas d'erreur de modèle puisque c'est le modèle le plus prudent. Par contre, le choix d'un des autres modèles comme modèle de référence entraîne une potentielle erreur dans le modèle qui devrait être prise en compte dans le calcul du capital réglementaire. Cette erreur, en absolu, est maximale si le modèle de Black-Scholes est référent et s'élève à 13.55%.

L'interprétation de cette erreur dans l'approche *worst case* doit être faite avec prudence. En effet, le résultat obtenu ne doit pas mener à la conclusion que le modèle de Black-Scholes est le pire modèle. Le résultat nous informe juste que la queue gauche de la distribution des rendements prédits obtenue à partir de ce modèle est moins épaisse que pour les autres modèles. Par conséquent, si ce modèle n'est pas le "vrai" modèle, alors la queue gauche sera sous-estimée. Cette approche étant une approche prudente, la mesure de l'erreur est calculée par rapport à la distribution qui aura la queue gauche la plus épaisse sans considération de la pertinence de chaque modèle.

### 5.3.2 Analyse comparative

Tout d'abord, les deux approches reposent sur la même hypothèse selon laquelle le "vrai" modèle appartient à l'ensemble des modèles choisis.

Nous avons montré que pour les deux approches, l'erreur de modèle est nulle si le modèle de Garch est choisi comme référence. Cela provient du fait que le modèle de référence est alors celui qui a la queue de distribution gauche la plus épaisse. Par conséquent, la baisse potentielle de l'actif, arrivant une fois tous les 200 ans, est la plus importante en choisissant ce modèle et la provision pour faire face à ce risque sera plus élevée en utilisant ce modèle.

Cependant, si le "vrai" modèle n'appartient pas à l'ensemble et que ce dernier présente une queue de distribution plus épaisse que celle du modèle de Garch alors l'erreur de modèle commise en sélectionnant le modèle de Garch n'est pas prise en compte. Ainsi, la véracité de l'hypothèse de base est primordiale bien que difficile à vérifier en pratique.

Par ailleurs, nous pouvons remarquer qu'une erreur de modèle apparaît si nous choisissons un des trois autres modèles comme référence. Dans ce cas, l'erreur de modèle obtenu avec l'approche bayésienne sera moins élevée que celle obtenue avec l'approche *worst case*. Nous présentons, ci-dessous, l'erreur de modèle pour l'approche Bootstrap dans le cas où chacun des trois autres modèles est choisi comme référence et nous rappelons les erreurs obtenues par l'approche *worst case*.

TABLE 11 – Synthèse des erreurs par les deux approches

Modèle de référence	B&S	Merton	Garch	B&S RS
Erreur Bayésienne	13.15 %	10.69%	-2.12%	10.69%
AM	13,55%	11,11%	0%	11,11%
RM	100%	79.688%	0%	79.706%

Ce résultat s'explique par le fait que l'approche bayésienne inclut l'erreur de modèle dans le calcul de la Value-at-Risk. Ainsi, nous n'évaluons pas l'erreur dans le pire des cas mais plutôt l'erreur commise par rapport au cas le plus vraisemblable. La distribution obtenue par cette approche possède une queue de distribution à gauche moins épaisse que celle du modèle de Garch et donc nécessite la provision d'un capital moins important.

L'approche bayésienne nous semble plus adéquate dans un environnement porté sur la gestion des risques et l'allocation optimale des ressources.

La modélisation d'un actif n'engendre pas une grande erreur liée au choix du modèle. Ce résultat peut s'expliquer par plusieurs points : le grand nombre de données disponibles permet de bien adapter les modèles ; les modèles d'actifs ont fait l'objet de nombreuses recherches et sont plutôt fiables.

## 5.4 Etude comparative sur rendements journaliers

### Erreur de modèle évaluée par les deux approches

Nous présentons rapidement les résultats obtenus sur les rendements journaliers ; l'objectif principal étant de comparer les résultats en fonction de la fréquence de temps considérée. Le tableau suivant

récapitule l'erreur de modèle pour l'approche *worst case* et l'approche bayésienne avec des poids similaires à l'étude précédente.

Les poids obtenus précédemment ont été évalués par l'approche bayésienne sur les rendements mensuels et non journaliers. Cependant, la vraisemblance obtenue sur les données journalières tend vers l'infini du fait du nombre important de données. Par conséquent, l'approche bayésienne n'est alors pas applicable et nous choisissons d'utiliser les poids *a posteriori* obtenus pour les rendements mensuels afin de pouvoir établir une vraie comparaison des valeurs des bootstrap pour les deux fréquences de rendements.

TABLE 12 – VaR à 0,5% et erreur de modèle sur 100 itérations de distributions de taille 20 000

	Moyenne	Ecart-type	CoV	Choc	Erreur
<b>BS</b>	7466.003	48.27	0.647%	-37.73164%	7.206 %
<b>Merton</b>	7433.871	49.15	0.661%	-37.99963%	6.805%
<b>Garch</b>	6928.014	87.53	1.263%	-42.21860%	0%
<b>BSRS</b>	7561.381	79.86	1.056%	-36.93616%	8.376%
<b>Bayésien</b>	6957.73	134.92	1.94%	-41.97 %	-0.429%

Pour l'approche *worst case* (les 4 premières lignes du tableau), l'erreur de modèle correspond à l'erreur commise en choisissant ce modèle comme référence et elle est calculée par la mesure relative. Notre modèle de référence étant le modèle de Garch, l'erreur est nulle. Concernant l'approche bayésienne, le calcul est similaire à l'étude journalière et nous obtenons une erreur de modèle négative étant donné que le modèle de Garch est le plus prudent.

Comme dans le cas mensuel, l'erreur de modèle est nulle et négative pour l'approche *worst case* et bayésienne respectivement si le modèle de Garch est choisi comme référence. Nous constatons par ailleurs que les écarts entre les VaR du modèle de Garch et des autres modèles sont plus grands en mensuel qu'en journalier, ainsi la sélection d'un autre modèle impliquera une plus grande erreur de modèle en approche *worst case*.

En outre, en analysant les rendements mensuels, nous diminuons le CoV de 20 à 100 points de base. Ainsi, les résultats sont plus stables en mensuel puisque l'indice fluctue davantage sur un intervalle journalier que sur un intervalle de temps plus grand qui lisse la volatilité à court terme.

Nous constatons aussi que dans le cas des rendements mensuels, c'est le modèle de Black-Scholes et non celui de Black-Scholes à changements de régime qui est le moins prudent. Cela peut s'expliquer par le fait que le modèle de Black-Scholes à changements de régime avec seulement deux régimes arrive mieux à s'adapter aux rendements mensuels qui sont plus stables et dont les deux régimes sont plus faciles à identifier. Ainsi, l'hypothèse forte de l'existence de deux régimes aura moins tendance à être violée dans ce cas. Par ailleurs, nous constatons grâce au graphique quantile-quantile (cf annexe A.3) que les rendements journaliers ne sont pas gaussiens du fait d'évènements extrêmes correspondant aux périodes de crise. Le modèle de Black-Scholes n'est donc évidemment pas plus adapté dans le cas de rendements journaliers qui sont d'autant plus volatils (cf annexe A.3).

## 5.5 Conclusion

L'erreur de modèle est nulle voire négative si le modèle de Garch est choisi comme référence. Ce n'est pas le cas si un des trois autres modèles est référent mais dans ce cas l'erreur ne dépasse pas 13.55% (si le modèle de Black-Scholes est référence). Ainsi, le choix du modèle de référence n'est pas anodin et peut engendrer des erreurs de modèle notables. Cependant, dans la plupart des cas, le modèle de Garch semble être le plus adéquat pour modéliser les actifs. Nous pouvons donc imaginer que l'erreur de modèle d'actifs est faible, nulle voire négative (modélisation prudente).

Il est important de rappeler que le calcul de cette erreur est effectué dans les deux approches par rapport à un modèle de référence. Ainsi l'erreur de modèle nous permet juste de conclure si ce modèle

est plus ou moins prudent que les autres modèles mais cette mesure ne nous permet en aucun cas de déterminer l'adéquation du modèle. Pour cela, il faut faire une étude préalable de backtesting des modèles comme dans la section 5.2.

Pour finir, une hypothèse fondamentale et forte de notre étude est la détermination de l'ensemble de modèles étudiés. En effet, nous ne pouvons pas nous assurer que le 'vrai' modèle appartient à notre ensemble. Or, la théorie n'est vraie que si ce dernier appartient bien à l'ensemble et que nous pouvons être certain que le capital réservé pour prendre en compte l'erreur potentielle de modèle est suffisant puisqu'il serait alors calculé par rapport au 'vrai' modèle. Cette hypothèse ne peut être vérifiée et le mieux que l'on puisse faire est de choisir l'ensemble de modèles les plus vraisemblables et adéquates au vu de nos données. Dans notre étude, nous avons remarqué dans la partie validation (cf section 5.2) que plusieurs modèles, et en particulier le modèle de Black-Scholes, n'étaient pas adaptés.

## 6 Application à des structures de dépendance

Dans cette section, nous allons évaluer l'erreur de spécification lors de la modélisation jointe de deux actifs. Nous allons ici nous intéresser au risque provenant d'une mauvaise prise en compte de la dépendance entre des actifs d'un portefeuille. En l'occurrence, pour des questions de simplification de l'étude, nous nous concentrons sur la modélisation jointe de deux actifs. Nous reprenons l'indice boursier journalier britannique GDDLUK de 2001 à 2014 que nous avons utilisé précédemment (cf section 5.1.1) pour étudier l'erreur liée au choix d'un modèle d'actif. Nous allons l'analyser conjointement avec son homologue américain, l'indice boursier journalier GDDLUS de 2001 à 2014. Les données américaines ont subi les mêmes retraitements que l'indice britannique ; ainsi toutes les dates coïncident parfaitement.

Dans cette section, la dépendance sera ici modélisée grâce à des copules. Nous commencerons donc par choisir les copules étudiés et définir leur paramètre "optimal", celui ayant la meilleure adéquation à nos données. Nous comparerons les résultats obtenus par différentes méthodologies (CQEP et maximisation de la pseudo-vraisemblance).

Par la suite, nous calculerons l'erreur de dépendance pure, i.e. l'erreur de spécification uniquement due au choix de la copule et non au choix des marginales (des modèles d'actifs). Nous appliquerons l'approche *worst case* d'une part puis nous pondérerons chaque copule par l'approche bayésienne afin d'intégrer l'erreur de modèle directement dans la mesure de risque.

Pour finir, nous évaluerons l'erreur de modèle prenant en compte à la fois l'erreur de spécification du modèle d'actifs (les marginales) et l'erreur de spécification de la copule modélisant la dépendance entre les actifs.

### 6.1 Choix de la copule et ses paramètres

#### 6.1.1 Choix de la copule

La dépendance sera ici modélisée essentiellement grâce à des copules classiques (copule Gaussienne, de Student, de Clayton, de Gumbel et de Frank) dont les définitions et caractéristiques ont été présentées précédemment (cf section 4.1). Ces copules ont été choisies, d'une part, parce qu'elles permettent de modéliser la plupart des types de dépendance en dimension 2. D'autre part, elles sont relativement faciles à implémenter. Enfin, le graphe 13 représentant le nuage de points obtenu sur les rangs normalisés de nos observations justifie également ce choix.

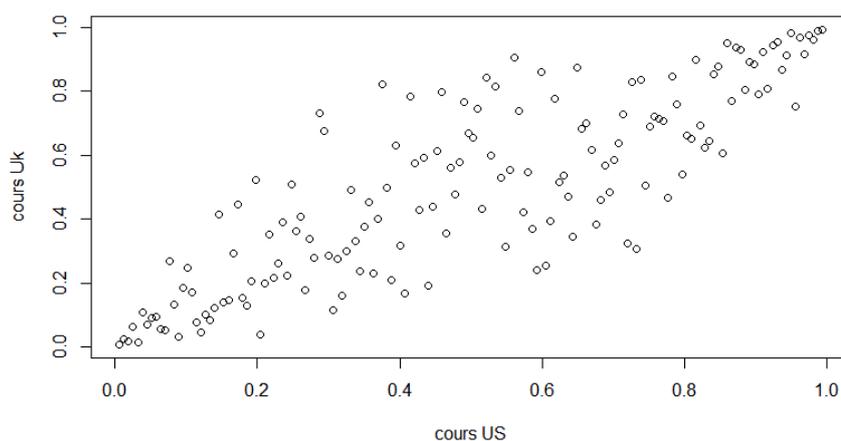


FIGURE 13 – Nuage de points des rangs sur les indices mensualisés US et UK

Ce graphique nous permet de faire une première conjecture quant au type de copule étant potentiellement le plus adéquat. Nous pouvons notamment observer la symétrie de la copule ce qui réduit les possibilités aux copules elliptiques. Par ailleurs, les queues de distributions ne paraissent pas particulièrement lourdes et il semblerait qu'il y ait une certaine dépendance entre les deux variables.

Les mesures de dépendance classiquement utilisées viennent confirmer cette constatation. En effet, la corrélation de Pearson (ou corrélation linéaire) vaut ici 90.6% et la corrélation de Spearman (ou corrélation des rangs) vaut 83.8%.

Dans la suite, des tests d'adéquation de copules de type Cramér-von Mises sont implémentés afin d'avoir un outil quantitatif d'évaluation de la pertinence de chaque copule.

### 6.1.2 Choix des paramètres de copule

Dans cette section, nous allons chercher le paramètre de copule qui permette de modéliser au mieux la dépendance entre nos actifs et ce pour différentes copules. Pour cela, nous allons mettre en pratique les deux méthodes présentées dans la section présentant l'erreur de modèle (cf section 4.2) sur des structures de dépendance à partir de nos données historiques.

#### la CQEP lower tail

Dans notre étude, nous avons choisi de calculer la CQEP (cf section 4.2) pour la queue gauche de la distribution puisque comme mentionné précédemment, le risque encouru sur les modèles d'actifs est un risque de baisse conséquente des actifs. Dans le cas d'un portefeuille d'actifs, le risque est lié à la baisse simultanée et prépondérante des actifs du portefeuille.

Nous présentons ci-dessous les différentes valeurs du CQEP obtenues pour différents niveaux :

TABLE 13 – Valeurs du CQEP low tail pour différents niveaux

<i>alpha</i>	<i>CQEP low tail</i>
0.5%	1
1%	0.5
5%	0.5
10%	0.6875
20%	0.7742

Nous remarquons que la corrélation dans l'extrême partie de la queue gauche de la distribution jointe est de 100%. Ceci s'explique par le faible nombre de données et le fait que les événements extrêmes des deux distributions correspondent à la période de crise de 2008 ainsi il y a une forte corrélation à ces dates. Lorsque nous prenons un quantile d'ordre plus élevé, la dépendance est moins élevée ; en d'autres termes, la valeur des actifs ne chute pas particulièrement aux mêmes dates lorsque nous nous intéressons aussi à des baisses moins importantes que celle de 2008.

Excepté le quantile à 0.5% qui correspond à un cas extrême (1 point de la distribution jointe), la dépendance augmente avec le niveau étudié ce qui laisse penser que la dépendance entre ces deux actifs n'est pas particulièrement au niveau de la queue de distribution. Cela peut aussi s'expliquer par le manque de données extrêmes présentes dans nos observations. Par conséquent, la dépendance n'est ici pas concentrée dans les queues de distributions mais est plutôt globale ce qui est en adéquation avec la constatation graphique ci-dessus.

TABLE 14 – Paramètres estimés pour différentes copules et différents niveaux

$\alpha$	<i>Normal</i>	<i>Student</i>				<i>Clayton</i>	<i>Gumbel</i>	<i>Frank</i>
		<b>4 degrés</b>	<b>5 degrés</b>	<b>11 degrés</b>	<b>12 degrés</b>			
<b>0.5%</b>	Diverge	Diverge	Diverge	Diverge	Diverge	Diverge	Diverge	Diverge
<b>1%</b>	0.8789	0.7788	0.8009	0.8463	0.8492	0.9925	4.9437	diverge
<b>5%</b>	0.8041	0.7287	0.7450	0.7786	0.7809	0.9586	3.3303	19.2485
<b>10%</b>	0.8992	0.8762	0.8812	0.8914	0.8921	1.8301	4.5974	19.9199

Le tableau 14 précédent présente les paramètres obtenus par la méthode du CQEP. Il peut être assez difficile de juger de la pertinence de chaque résultat.

Tout d'abord, l'inversion ne peut se faire que lorsque le quantile conditionnel est inférieur strictement à 1 sinon il n'y a pas convergence de la méthode d'optimisation (cf section 4.2) et nous notons Diverge dans le tableau 14 de résultats.

Nous pouvons déjà mettre de côté la copule de Frank qui semble avoir une queue droite de distribution bien plus lourde que celle observée.

Par ailleurs, la grande variabilité des paramètres obtenus pour les copules de Clayton et de Gumbel s'explique par le fait que les paramètres ne sont viables que sur une partie de la distribution à chaque fois. En augmentant le niveau du quantile, nous nous intéressons à la dépendance sur davantage de données et la valeur du paramètre de copule observée sur l'extrémité de la distribution ne coïncide plus avec le nouveau paramètre. Ces copules peuvent être adaptées sur une partie extrême de la distribution mais non sur l'ensemble. En effet, ce résultat était envisageable au vu du graphe des rangs puisque les copules de Clayton et Gumbel sont des copules utilisées pour modéliser les valeurs extrêmes et ont une dépendance de queue unilatérale.

La méthode du maximum de la pseudo-vraisemblance semble plus adaptée à nos données qui ne disposent pas de queues lourdes. Ainsi, nous utiliserons cette seconde méthode pour définir nos paramètres.

### Maximisation de la pseudo-vraisemblance

Pour maximiser la pseudo-vraisemblance, nous utilisons la fonction *fitCopula* sous R en spécifiant la méthode 'mpl'. Puis nous appliquons la fonction *gofCopula* sous R pour effectuer un test d'adéquation. Ce test étudié par Genest et al. [22] est une version basée sur les rangs du test de Cramér-von Mises. Ce test s'écrit de la manière suivante :

$$S_n = \int_{[0,1]^2} \sqrt{n}(C_n(u, v) - C_{\hat{\theta}}(u, v))^2 dC_n(u, v)$$

avec

- $C_n$  : la copule empirique, si  $(U, V) \sim C$ ,  $C_n = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(U_i \leq u, V_i \leq v)}$
- $C_{\hat{\theta}}$  : la copule supposée "vraie" paramétrée par le paramètre estimé.

Les résultats sont présentés ci-dessous pour les mêmes copules que précédemment sachant que la copule de Student avec un degré de liberté de 12 est la copule de Student qui maximise la pseudo-vraisemblance.

TABLE 15 – Paramètres estimés par pseudo maximum de vraisemblance

degrés	Normal	Student				Clayton	Gumbel	Frank
		4	5	11	12			
paramètre	0.8776	0.8610	0.8661	0.8746	0.8750	2.8709	2.8833	9.2502
statistique	0.0198	0.0178	0.0177	0.0186	0.0187	0.0731	0.0333	0.0453
p-value	0.1214	0.2153	0.2133	0.1683	0.1673	0.0015	0.0155	0.0025
log-vrais	111	110.67	110.92	111.204	111.205	94.81	105.10	92.06

Concernant les 5 copules elliptiques étudiées, les p-values du tests d'adéquation sont toutes supérieures à 5% ce qui nous permet pas de rejeter l'hypothèse nulle d'adéquation des données à la copule paramétrée par le paramètre estimé. Au contraire, aucune des copules archimédiennes ne passe ce test ce qui coïncide avec toutes les remarques et observations que nous avons faites précédemment. Nous

pourrons ainsi voir si l'utilisation (à tort) d'une de ces copules pourrait entraîner de potentielles pertes subséquentes.

La copule de Student semble meilleure que la copule gaussienne. Cette copule permet de modéliser des queues de distribution plus lourde que la gaussienne ce qui peut expliquer ce résultat, bien que la distribution jointe ne semble pas avoir des queues particulièrement lourdes. Il est difficile au vu des ces résultats de choisir la copule de Student qui possède la meilleure adéquation car la méthode de maximisation de la vraisemblance sélectionne la copule à 12 degrés de liberté (qui sera la copule de référence) et le test d'adéquation basé sur la statistique de Cramér-von Mises a tendance à préférer celle à 4 degrés de liberté.

## 6.2 L'erreur de dépendance pure

A présent que nous avons estimé le paramètre  $\hat{\theta}$  de chaque copule, nous pouvons évaluer la distribution jointe à 1 an.

**Etape 1 :** Simulation d'un vecteur  $(u, v)$  de taille 10000 de loi  $C_{\hat{\theta}}$

**Etape 2 :** Projection à un an (fin 2014) des actifs américains et britanniques séparément pour chacun des 4 modèles d'actifs étudiés et calcul des rendements espérés :

$$r_i = \frac{S_{1,i} - S_0}{S_0}, i \in \{\text{indice américain, indice britannique}\} \text{ avec } S_0 \text{ le dernier indice de 2013}$$

et  $S_{1,i}$  l'indice projeté à fin 2014 à l'aide d'un des quatre modèles d'actifs.

**Etape 3 :** Calcul de  $(x, y) = (F_1^{-1}(u), F_2^{-1}(v))$  où  $F_1$  et  $F_2$  sont les fonctions de répartition respectives des variables  $r_1$  et  $r_2$ . Pour cela, nous utilisons la fonction inverse généralisée et la propriété du rang :

Soit  $X$  un vecteur aléatoire de taille  $n$  et de fonction de répartition  $F$ , alors

$$F^{-1}(u) = \inf\{x, F_i(x) \geq u\} = X_{\lfloor nu \rfloor + 1} \text{ avec } X_{[i]} \text{ la } i\text{-ème valeur du vecteur trié.}$$

Par conséquent,  $(x, y) = (r_{1 \lfloor nu \rfloor + 1}, r_{2 \lfloor nv \rfloor + 1})$  a pour distribution la loi jointe des deux actifs.

**Etape 4 :** A partir de l'exposition de chacun des actifs dans le portefeuille, nous déterminons la distribution de la valeur du portefeuille à fin 2014. Pour des raisons de confidentialité, les expositions ne correspondent en aucun cas à celles de la SCOR sur ces deux indices. Nous supposons donc que la valeur initiale (à fin 2013) de notre portefeuille est de 100M (nous ne précisons pas d'unités) dont 70M en actions américaines et 30M en actions britanniques. La valeur projetée à fin 2014 du portefeuille est donc :

$$70000000 * x + 30000000 * y$$

**Etape 5 :** Calcul de la VaR à 0,5% de la distribution finale pour chaque modèle et chaque copule.

### 6.2.1 Résultats sur données mensuelles

Dans cette section, nous ne voulons nous intéresser qu'au risque lié à la spécification de la copule. Ainsi, nous choisissons le modèle de Garch (comme référence) pour projeter nos actifs, les marges suivent donc un modèle de Garch paramétré à l'aide des paramètres estimés sur nos échantillons. Nous comparons ensuite les résultats obtenus pour les différentes copules répertoriés dans le tableau 16 ci-dessous.

Nous observons que la variation maximale entre deux VaR est de presque 4.5M ce qui ne représente que 4.5% de la valeur initiale du portefeuille. Ainsi, ce fort écart nous permet, d'une part, de constater l'importance de traiter d'une potentielle erreur de modèle liée à la dépendance et, d'autre part, l'importance du choix des copules en fonction de leur pertinence et du choix de la pondération pour l'approche bayésienne.

Nous allons maintenant évaluer l'erreur de modèle pure par les deux approches proposées. La copule de référence sera la copule de Student avec 12 degrés de liberté.

TABLE 16 – VaR à 0,5% sur 100 itérations et le modèle de Garch pour les marges

Degrés	<i>Garch</i>				
	<i>Normal</i>	<i>Student</i>			
		4	5	11	12
<b>Moyenne</b>	-36.806M	-37.149M	-37.020M	-37.120M	-36.971M
<b>Ecart-type</b>	1 288 206	1 252 835	1 272 383	1 277 726	1 296 645
<b>CoV</b>	3.500%	3.372%	3.437%	3.442%	3.507%
	<i>Clayton</i>	<i>Gumbel</i>	<i>Frank</i>		
<b>%</b>	-38.116M	-35.026M	-32.766M		
<b>Ecart-type</b>	1 472 730	1 249 083	952 547		
<b>CoV</b>	3.864%	3.566%	2.907%		

### Approche worst case

Nous calculons ici les mesures absolues de risques par le même calcul que précédemment. Il est à noter que les valeurs de pertes étant négatives, il faut multiplier par (-1) la formule utilisée pour garder un pourcentage :

$$AM_- = - \left( 1 - \frac{\min_{C_i \in \mathbb{C}} VaR_{0,5\%}(C_i)}{VaR_{0,5\%}(C^*)} \right)$$

avec

- $\mathbb{C}$  : l'ensemble des copules sélectionnées pour l'étude
- $C_i$  : la  $i$ -ème copule
- $C^*$  : la copule de référence, ici la copule de Student à 12 degrés de liberté

Les résultats sont présentés dans le tableau 17 ci-dessous en retenant successivement chaque copule comme copule de référence. Nous constatons que l'erreur de modèle peut être très élevée selon le modèle de référence choisi. L'erreur maximale qui peut être faite est de 16.331% si nous choisissons la copule de Frank comme copule de référence ; cette erreur n'est pas négligeable. Cependant, étant donné que les copules qui semblent les plus adaptées pour modéliser la dépendance de nos actifs sont les copules elliptiques, le choix de la copule de Frank comme référence paraît peu vraisemblable. Cependant, l'erreur de modèle correspondant à notre copule de référence (copule de Clayton) est de 3.096%, ce qui n'est pas négligeable mais reste inférieur à 5% (la marge d'erreur communément tolérée).

TABLE 17 – AM en supposant successivement chaque copule comme copule de référence

Modèle de référence		AM	AM bis
<i>Normal</i>		3.560%	3.560%
<i>Student</i>	4 degrés	2.604%	-
	5 degrés	2.962%	-
	11 degrés	2.684%	-
	12 degrés	3.096%	3.096%
<i>Clayton</i>		0%	0%
<i>Gumbel</i>		8.822%	8.822%
<i>Frank</i>		16.331%	16.331%

Etant donné que la copule la plus prudente est la copule de Clayton, l'erreur de modèle ne change pas si nous décidons de considérer plus ou moins de copule de Student avec des degrés de liberté différents.

Nous présentons, sur le graphe 26 ci-dessous, la moyenne des VaR à 0.5% obtenue pour 100 simulations sur des distributions de taille 50 000 et ce pour chaque copule et chaque modèle d'actifs.

Nous observons que le choix de la copule a un impact sur le quantile à 0.5% ; en particulier, le choix de la copule de Frank sera peu prudent au contraire du choix de la copule de Clayton. Par

ailleurs, nous remarquons que le choix des marges aura aussi une incidence sur le capital réglementaire mais ce de façon différente de l'étude d'un actif individuellement. En effet, le choix du modèle de Black-Scholes à changements de régime pour modéliser les marges sera le plus prudent. Par ailleurs, bien que l'écart entre les quantiles pour différents choix de marge soit assez important, cet écart reste quasiment constant entre les différentes copules sauf quelques exceptions pour les marges issues des modèles de Garch et Merton. La copule la plus prudente, la copule de Clayton, n'étant pas une copule elliptiques, nous supposons que sont adéquations aux données est faible. L'application de l'approche par pondération va nous permettre de diminuer considérablement cette erreur de 3.096% en donnant moins de crédibilité à cette copule.

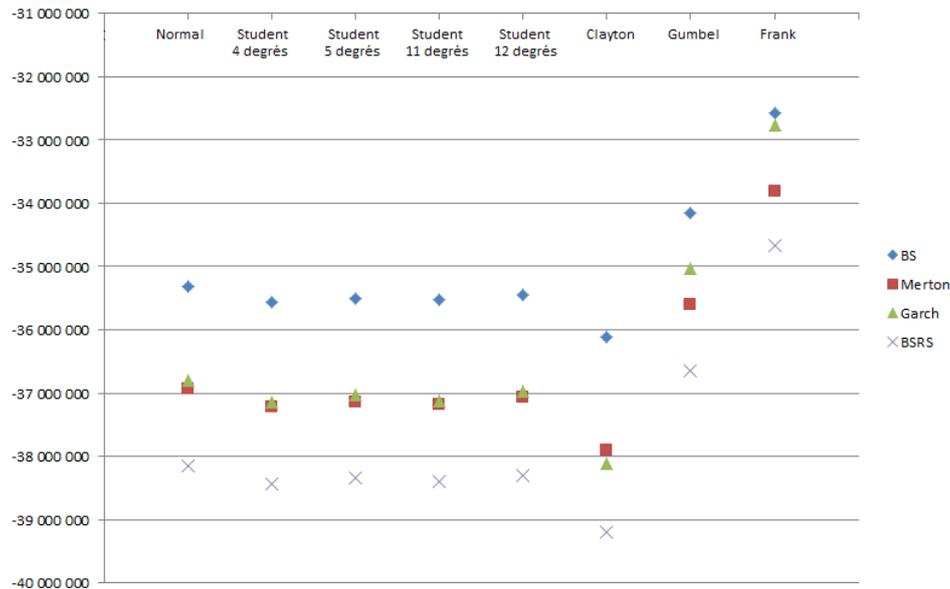


FIGURE 14 – Comparaison des VaR à 0,5% pour différentes marges et copules

### Approche par pondération bayésienne

Nous cherchons ici à pondérer nos copules afin d'intégrer l'erreur de dépendance pure dans le calcul de la VaR à 0.5%. Pour ce faire, nous utiliserons l'approche bayésienne avec les vraisemblances évaluées pour chaque copule que nous avons présentées ci-dessus. Nous choisissons l'*a priori* à l'aide des p-values obtenues pour le test d'adéquation (cf section 6.1.2) hormis pour les copules de Clayton et de Frank qui avaient de très petites p-values et dont nous avons fixé l'*a priori* à 0.5% pour ne pas les enlever de l'étude. Le calcul de l'*a posteriori* se fait par la méthode vue dans la partie 3.3.1.

Nous présentons ci-dessous (cf table 18) les pondérations ainsi que les valeurs des mesures de risque obtenues en prenant en considération plus ou moins de copule de Student et l'erreur de modèle associée. L'erreur de modèle est calculée en prenant la copule de Student à 12 degrés de liberté comme copule de référence. Elle est définie comme suit :

$$Erreur_{Boot} = - \left( 1 - \frac{VaR_{0,5\%}(C_{Boot})}{VaR_{0,5\%}(C^*)} \right)$$

avec  $C_{Boot}$  : la copule obtenue en faisant des simulations par bootstrap parmi les copules à l'aide des pondérations *a posteriori*.

Comme nous ne traitons que de l'erreur liée au choix de la copule dans cette partie, nous avons évalué l'erreur en fixant les marges. Pour un modèle d'actifs fixé modélisant les marges, le modèle de référence est le modèle de dépendance composé de la copule de Student à 12 degrés de liberté et les marges issues de ce modèle d'actifs.

Nous observons que l'erreur de modèle est désormais négative quelles que soit les marges choisies. Cela s'explique par le fait que la copule de Student à 12 degrés de liberté est globalement plus prudente que les autres copules et plus prudente que la copule issue du bootstrap. Si nous ne considérons que 5 copules, le poids donné *a posteriori* se concentre essentiellement sur les copules gaussienne et de Student. La copule gaussienne étant moins prudente que la copule de Student, la copule issue de la pondération des deux est aussi moins prudente que la copule de Student d'où une erreur de modèle négative pour cette approche. Cependant, l'erreur reste très faible en valeur absolue et n'est pas considérée comme significative. Par ailleurs, cette erreur augmente avec le nombre de copules de Student considérées. En effet, en ajoutant les 3 copules de Student retenues, le poids est réparti de façon assez homogène entre les copules gaussienne et les copules de Student, autres que celle de référence, sont plus prudentes que la référence. Ainsi le quantile extrême obtenu à l'issue des pondération des 8 copules reste moins prudent que la référence mais s'en rapproche beaucoup plus que celui obtenu après pondération des 5 copules; d'où une erreur de dépendance toujours négative mais quasiment nulle.

TABLE 18 – Log-vraisemblances, pondérations et erreur de modèle sur 100 itérations de distributions de taille 10 000

Degrés	Normal	Student				Clayton	Gumbel	Frank
		4	5	11	12			
log-vrais	111	110,6694	110,9236	111,2037	111,2048	94,81	105,1	92,06
a priori	13,300%	23,700%	23,400%	18,500%	18,400%	0,500%	1,700%	0,500%
a posteriori	13,673%	17,506%	22,286%	23,315%	23,215%	0,000%	0,005%	0,000%

VAR à 0.5%	Moyenne	%	Ecart-type	CoV	Erreur
BS	-35 426 370	-35.426%	595 035	1.680%	-0.068%
Merton	-37 061 833	-37.062%	759 816	2.050%	-0.024%
Garch	-36 901 370	-36.901%	1 259 418	3.413%	-0.190%
BSRS	-38 253 394	-38.253%	834 567	2.182%	-0.107%

Degrés	Normal	Student	Clayton	Gumbel	Frank
		12			
log-vrais	111	111,2048	94,81	105,1	92,06
a priori	34,42%	60,73%	0,50%	3,85%	0,50%
a posteriori	31,589%	68,402%	0,000%	0,010%	0,000%

VAR à 0.5%	Moyenne	%	Ecart-type	CoV	Erreur
BS	-35 256 360	-35.256%	513 999	1.458%	-0.547%
Merton	-36 845 559	-36.846%	663 767	1.801%	-0.607%
Garch	-36 557 496	-36.557%	1 081 892	2.959%	-1.120%
BSRS	-38 034 364	-38.034%	725 007	1.904%	-0.679%

En conclusion, nous pouvons retenir qu'il existe des écarts significatifs entre les quantiles à 0.5% obtenus par différentes copules, l'erreur de modèle peut dépasser 15% dans le pire des cas avec la copule de Frank comme référence. Cependant, nous remarquons que le modèle que nous avons choisi est plutôt prudent et l'erreur qui lui est associée est faible dans le pire des cas et négative dans le cas de la pondération qui prend en compte la pertinence des différentes copules sélectionnées. L'erreur de dépendance pure sera donc considérée nulle entre ces deux actifs pour le traitement des données mensuelles.

## 6.2.2 Comparaison avec les résultats sur rendements journaliers

Nous effectuons la même étude sur les rendements journaliers. Nous voulons ainsi vérifier qu'il n'y a pas de différences significatives dans l'erreur du modèle dues au fait d'utiliser les rendements journaliers ou mensuels. Les résultats sont présentés dans l'annexe A.4. Nous retenons ici aussi quatre copules de Student mais pas avec les mêmes degrés de liberté puisque les copules utilisées précédemment n'ont plus la meilleure adéquation. La copule de référence sera la copule de Student à 2.56 degrés de liberté qui a la plus grande vraisemblance (cf annexe A.4.1).

Le nuage de points obtenu sur les rangs normalisés de nos observations est le suivant :

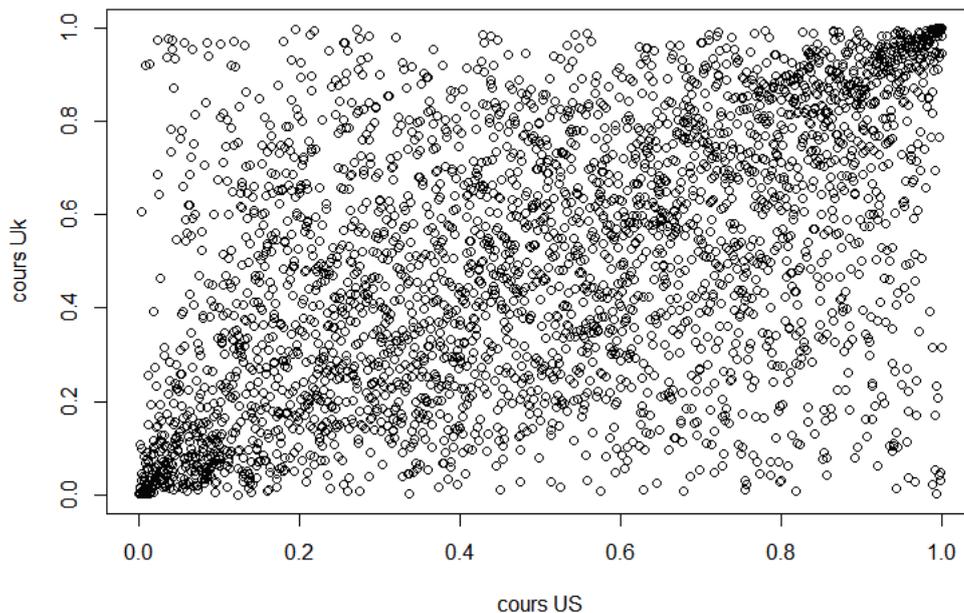


FIGURE 15 – Nuage de points des rangs sur les indices journaliers US et UK

L'erreur est calculée similairement excepté pour les poids qui ne sont pas évalués à partir de l'approche bayésienne. En effet, nous avons un grand échantillon de données, les vraisemblances associées à chaque copule tendent donc vers l'infini. Ainsi, nous choisissons d'utiliser les pondérations obtenues sur les rendements mensuels. Comme les degrés de liberté des copules de Student retenues sont différents, nous décidons de nous référer à la log-vraisemblance obtenue sur les rendements journaliers pour attribuer les poids de la copule de la plus vraisemblable à la moins (cf annexe A.4.2).

Nous constatons que les pertes potentielles bicentennaires sont assez proches en mensuel et en journalier pour le modèle de Garch. Cependant, les écarts entre les différentes copules sont plus marqués sauf entre les différentes copules de Student (cf annexe A.4.5). Par ailleurs, nous remarquons que le modèle avec marges issues du modèle Black-Scholes à changements de régime devient le moins prudent alors qu'il était le plus prudent en mensuel. Ce résultat s'explique par la difficulté rencontrée lors de la mise en application de ce modèle. En effet, nous avons choisi de nous limiter à deux régimes, cela limite nécessairement l'adéquation aux données. Sur des données journalières, les sauts sont très fréquents et le modèle ne parvient pas à s'adapter aux données en ne considérant que deux régimes. De plus, la probabilité de passage d'un régime à l'autre est aussi un paramètre sensible du modèle dont la calibration n'est pas évidente. Ainsi le modèle aura tendance à être plus fiable sur les données mensuelles même s'il est nécessaire de prendre des précautions puisque les choix d'une paramétrisation simplifiée du modèle engendre nécessairement une moins bonne adéquation même sur les données mensuelles.

Ici, nous nous intéressons surtout à l'erreur liée au choix de copule avec la copule de Student à 2.56 degrés de liberté comme référence puisque c'est elle qui a la plus grande vraisemblance. Nous observons que l'erreur de modèle est très légèrement supérieure dans le cas de l'approche *worst case* (AM=3.732%) mais est plus faible pour l'approche bayésienne puisque la copule de référence est l'une des plus prudentes (voire la plus prudente) parmi les copules ayant une pondération significative (cf annexe A.4.4). Le nombre de copules de Student prises en compte ne change pas l'erreur dans le pire des cas. Pour l'approche bayésienne, nous constatons que l'erreur augmente avec le nombre de copules de Student incluses. Ces résultats sont identiques à ceux obtenus sur les données mensuelles et s'expliquent similairement.

La conclusion est donc similaire à celle pour les données mensuelles. L'erreur de dépendance pure sur ces actifs est non significative pour les copules sélectionnées.

### 6.3 L'erreur de modèle d'actifs

Dans cette section, nous cherchons à évaluer l'erreur de modèle globale, c-à-d. à évaluer conjointement l'erreur de dépendance pure et l'erreur de modèles des marginales. Nous sélectionnons pour ce faire une copule de référence et un modèle d'actif de référence. Pour les données journalières ce sera la copule de Student à 2.56 degrés de liberté qui sera référente tandis que ce sera la copule de Student à 12 degrés de liberté pour les données mensuelles. Par ailleurs, le modèle de Garch sera le modèle de référence pour modéliser les marges.

#### Données mensuelles

Pour évaluer l'erreur totale par l'approche *worst case*, il est nécessaire d'évaluer la VaR à 0.5% pour chaque modèle et chaque copule. Par la suite, nous nous référons à la formule de la mesure absolue 6.2.1 pour évaluer l'erreur totale de modèle en fonction de la référence sélectionnée ci-dessus.

Modèle de référence	Garch-Student à 12 degrés de liberté
AM	6.017%

Quel que soit le nombre de copules retenues, l'erreur de modèle pour l'approche par le pire des cas est identique. Cette erreur dépasse 5%, seuil à partir duquel l'erreur est considérée communément comme significative. Par ailleurs nous observons que l'erreur augmente du fait de la dépendance.

Pour évaluer l'erreur par l'approche bayésienne, nous effectuons deux simulations par bootstrap : une pour simuler la distribution à 1 an des marges et une pour simuler la copule. Les marges sont supposées être obtenues à partir du même modèle pour être cohérent puisque des actifs, ayant des propriétés similaires, seront très probablement modélisés de façon équivalente (les paramètres variant). Par ailleurs, cette hypothèse permet de simplifier les simulations. Les poids utilisés sont les mêmes que précédemment, obtenus par la théorie bayésienne.

<i>Modèle de référence</i>	VaR*	-36 971 480
<i>Pondération de 8 copules</i>	$VaR_{Boot}$	-36 364 060
	Erreur	-1.643%
<i>Pondération de 5 copules</i>	$VaR_{Boot}$	-36 117 853
	Erreur	-2.309%

La mesure de risque après pondération est moins prudente que la mesure de risque du modèle de référence ; c'est pourquoi, l'erreur est négative. Ce n'est pas le cas dans l'approche *worst case* où il existe plusieurs modèles plus prudents que le modèle de référence notamment les modèles avec les marges issues du modèle de Black-Scholes à changement de régime. Cependant, au vu des tests effectués sur ce modèle d'actifs et des intuitions que nous avons, nous privilégions le résultat de l'approche bayésienne qui prend en compte la vraisemblance de ce modèle qui est en l'occurrence bien plus faible que celle du modèle de Garch.

Au vu des résultats par l'approche bayésienne, nous pouvons supposer que le modèle de référence est très prudent. Cependant, si ce dernier s'avérait être le 'vrai' modèle, cette réserve serait nécessaire. Comme nous avons pu le mentionner au préalable, une erreur de modèle négative n'implique pas, par prudence, la reprise de capital.

## Données journalières

Nous suivons la même démarche sur les données journalières. L'erreur est bien moins élevée dans l'approche du pire des cas que sur les données mensuelles, cependant elle est positive dans l'approche par pondération. L'erreur est toutefois très faible (inférieure à 1%) et n'est donc pas significative. Nous retiendrons que l'erreur dans le pire des cas est plus faible du fait que, sur les données journalières, le modèle de référence est l'un des plus prudents (cf annexe A.4.5).

<i>Modèle de référence</i>	VaR*	-37 545 597
<i>Pondération de 8 copules</i>	AM	3.732%
	$VaR_{Boot}$	-38 870 522
	Erreur	0.865%
<i>Pondération de 5 copules</i>	AM	3.732%
	$VaR_{Boot}$	-37 692 336
	Erreur	0.391%

## 6.4 Conclusion

Pour conclure, l'ajout d'une dépendance entre deux actifs n'entraîne pas d'augmentation significative de l'erreur de modèle. En effet, lorsque nous avons étudié l'erreur de dépendance pure, nous avons observé que le choix de la copule modélisant cette dépendance pouvait avoir un impact important. Cependant, l'erreur évaluée avec la copule de référence n'est pas significative. Il en est de même pour la dépendance globale même si l'erreur est plus importante pour l'approche *worst case*.

L'erreur de modèle totale n'est tout de même pas très élevée dans le cas de l'approche *worst case* mais est supérieure à l'erreur de modèle dans le cas d'un actif seul (l'erreur *worst case* était nulle). Cette augmentation s'explique par l'erreur liée à la dépendance qui n'est pas nulle. Il n'y a cependant toujours pas d'erreur de modèle dans le cas de l'approche bayésienne, elle est, en tout cas, non significative.

La comparaison entre les rendements journaliers et les rendements mensuels a montré que l'implémentation des modèles sur les données mensuelles est plus stable et plus aisée. Cependant, les résultats ne sont pas fondamentalement différents. Nous noterons quand même que nous avons comparé les résultats en terme d'erreur et non en terme de capital requis. Nous observons ainsi que le capital requis pour couvrir les pertes bicentennaires sur les données mensuelles pour le modèle de référence est de 37M tandis qu'il est de 37.5M sur les données journalières. Cette différence ne représente que 0.5% du capital initial pour ce portefeuille et s'explique par l'imprécision et l'instabilité des modèles en particulier sur les données journalières. Nous préférons donc travailler sur données mensuelles.

## 7 Analyse et interprétations des résultats sur les modèles d'actifs

L'agrégation de deux risques n'a pas montré d'impact significatif sur le capital réglementaire pour notre application. Nous avons pu observer que l'erreur restait nulle dans le cas des modèles d'actif pour l'approche par pondération mais augmentait significativement en approche *worst case*. Comme nous l'avons mentionné, le choix de l'ensemble de modèle est essentiel dans l'étude d'un risque isolé. Dans l'étude des modèles de dépendance, c'est d'autant plus vrai que deux ensembles sont sélectionnés, celui des modèles pour les marges et celui des copules. Ainsi, la probabilité de réaliser un mauvais choix d'ensemble augmente nécessairement. En effet, nous avons gardé certaines copules dans l'étude qui semblent assez éloignées de la 'vraie' copule, par exemple, la copule de Clayton pour les actifs qui augmente considérablement l'erreur dans le pire des cas. Dans l'approche par pondération, le poids donné à cette copule est quasiment nul puisque qu'elle n'est pas bonne en terme d'adéquation. L'erreur de modèle avec cette approche est donc bien inférieure à celle dans le pire des cas. L'approche par pondération semble donc préférable.

Dans toute la partie traitant des actifs, nous avons pu montrer qu'il n'existait pas ou peu d'erreur de modèle avec les références sélectionnées. L'erreur reste relativement faible si nous changeons le modèle de référence excepté pour quelques modèles de dépendance qui ne semblent, quoi qu'il en soit, pas adéquats. Cette constatation souligne cependant l'importance du choix du modèle de référence. Il est nécessaire d'insister sur le fait qu'il ne s'agit que de deux actifs particuliers et qu'il est difficile d'émettre une conclusion générale concernant tous les actifs modélisés. Cependant, les modèles d'actifs ont fait l'objet d'une abondante littérature et sont plutôt bien maîtrisés par les praticiens désormais. L'historique de données étant relativement long, l'adéquation des modèles est plus aisée à vérifier.

Au vu de cette étude sur un segment particulier, il ne semble pas nécessaire d'allouer du capital pour l'erreur de modèle. Il faut néanmoins être attentif à plusieurs éléments :

- Le type d'actifs modélisé -> les résultats obtenus sont considérés comme valables pour tout actif ayant des propriétés similaires (volatilité, tendance...)
- Le choix du modèle de référence -> il faut effectuer un backtesting du modèle afin de vérifier l'adéquation du modèle.
- Le choix de la copule de référence
- Vérification de la prudence du modèle (queue de distribution gauche pas trop fine)

## Partie IV

# Etude des modèles de provisionnement non-vie

## 8 Application à un risque isolé

### 8.1 Données, problématiques et modèles

Dans cette partie, nous présentons les données que nous avons utilisées pour mettre en place nos approches de calcul de l'erreur de modèles. Nous développons ensuite les différents modèles de provisionnement que nous avons sélectionnés afin de constituer notre ensemble de modèle.

#### 8.1.1 Les données

Les données sélectionnées ici sont des données provenant d'un segment de transport maritime. Les sinistres sous-jacents proviennent principalement des assurances des navires marchands et des biens transportés (dommages matériels), des interruptions d'activité, des dommages physiques (notamment sur les plateformes offshore) et des assurances de responsabilité. Nous ne détaillerons pas les montants de ces sinistres pour des questions de confidentialité. Nous choisissons de travailler sur le triangle des sinistres payés, i.e. le triangle contenant tous les montants déjà payés par année d'origine et par période de développement. Ce triangle présente l'avantage d'avoir un nombre d'incrément négatifs faible (ici nul) puisqu'il prend en compte uniquement les montants payés, contrairement au triangle des charges de sinistres qui se compose de la somme des montants payés et des montants restant à payer estimés par la cédante.

Le graphique ci-dessous présente les sinistres cumulés par période de développement pour chaque année d'origine de 1999 à 2013.

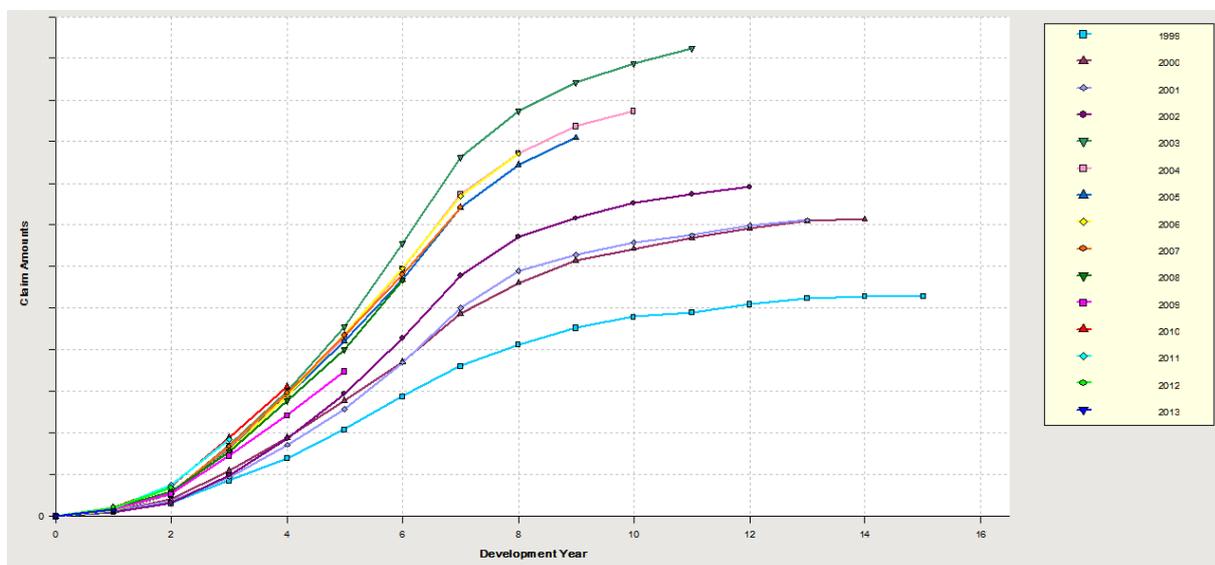


FIGURE 16 – Sinistres cumulés par période de développement pour chaque année d'origine

Nous remarquons que les données sont plutôt lisses avec un développement assez proche entre les différentes années d'origine. Nous n'effectuons donc pas de retraitements sur ce triangle (suppression de données, lissage des données...). Les années d'origine les plus récentes ont des montants de sinistres plus élevés pouvant s'expliquer en grande partie par l'inflation. Les paiements cumulés semblent se stabiliser vers la 12ème période de développement, ce qui signifie que la liquidation des sinistres d'une année d'origine est d'environ 12 ans. Cependant, en 2013, aucun montant incrémental payé n'est négatif ou nul ce qui signifie que les sinistres d'aucune année d'origine ne sont totalement liquidés. Ce triangle est donc plutôt à développement long.

### 8.1.2 Notation et introduction à la problématique du provisionnement

Nous commençons par introduire un certain nombre de notations utilisées dans la description des modèles de provisionnement ci-dessous.

- $i = 1, \dots, n$  l'année d'origine ou de survenance du sinistre
- $j = 1, \dots, n$  la période de développement
- $D_{i,j}$  les paiements cumulés
- $C_{i,j}$  les paiements incrémentaux pour l'année de développement  $j$  survenus l'année  $i$ .

Pour provisionner des provisions en assurance, nous cherchons à prédire au plus juste et de manière prudente les paiements futurs sachant une information historique (fréquence, sévérité, durée de liquidation ...). L'information peut être notée :

$$\mathcal{F}_n = \{C_{i,j}, i + j \leq n + 1\} = \{D_{i,j}, i + j \leq n + 1\}$$

Ainsi, nous cherchons à prédire, pour chaque année de survenance  $i$ , le montant cumulé  $D_{i,n}$  conditionnellement à l'information passée  $\mathcal{F}_n$ .

Pour évaluer la qualité du modèle à prédire les provisions futures, nous ne pouvons pas nous contenter d'estimation des provisions en moyenne (*best estimate*), il nous faut davantage de statistiques (variance, quantiles...). Pour ce faire nous utilisons des méthodes numériques telles que le bootstrap pour obtenir des distributions des provisions.

Nous distinguerons l'erreur d'estimation de l'erreur de processus qui est l'erreur liée au choix de la distribution prédictive et qui prend en compte une partie de l'erreur de modèle.

$$\text{Erreur de prédiction} = \text{erreur d'estimation} + \text{erreur de processus}$$

### 8.1.3 Le modèle de Chain-Ladder

La méthode de Chain-Ladder repose sur les cadences de paiements. Les méthodes basées sur les cadences de paiements sont utilisées depuis les années 30 pour le provisionnement en assurance et sont très appréciées de par leur simplicité. Cependant, nous verrons que ce modèle pose, malgré la théorie mathématique sous-jacente, des problèmes lors de la mise en pratique, notamment pour le respect des hypothèses.

#### Les hypothèses

Le modèle repose sur deux hypothèses fortes :

H1 : Les paiements cumulés  $D_{i,j}$  des années de survenance  $i$  sont indépendants.

H2 :  $\forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, n - 1\}$ , il existe  $\lambda_{j+1}$  tels que

$$D_{i,j+1} = \lambda_{j+1} D_{i,j}$$

En d'autres termes, les paiements cumulés d'une année de développement à la suivante sont proportionnels et le coefficient de proportionnalité est constant pour chaque année de survenance. L'hypothèse sous-jacente est l'indépendance entre  $\frac{D_{i,k+1}}{D_{i,k}}$  et  $\frac{D_{i,k}}{D_{i,k-1}}$ .

*Remarque.*

Avantages : La méthode de Chain-Ladder présente l'avantage d'être simple et de pouvoir être utilisée sur des triangles non cumulés présentant des incréments négatifs.

Inconvénients : Les hypothèses ne spécifient que le moment d'ordre 1 et modélisent la tendance de manière constante sur la période de développement ; ce modèle ne permet pas d'obtenir des intervalles de confiance. Nous verrons d'autres modèles qui précisent la forme de la variance (ex : modèle de Mack) afin de ne pas évaluer uniquement les provisions *best estimate* mais aussi leur volatilité en ultime. Par ailleurs, les hypothèses sont fortes et doivent être vérifiées

préalablement à l'utilisation du modèle notamment sur les segments à long développement puisque certains changements (ex : modification du portefeuille, des politiques de l'entreprises, des clauses contractuelles, forte inflation...) peuvent remettre en cause ces hypothèses. En effet, les cadences sont souvent irrégulières.

## Le modèle

Sous les hypothèses H1 et H2, nous pouvons déterminer les provisions estimées  $\hat{R}_i$  en *best estimate*,  $i \in 1, \dots, n$  par :

$$\hat{R}_i = \mathbb{E}[D_{i,n} | \mathcal{F}_n] - D_{i,n-i+1} = D_{i,n-i+1}(\lambda_{n-i+1} \dots \lambda_{n-1} - 1)$$

Les coefficients de Chain-Ladder  $\lambda_j$  sont estimés par :

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} D_{i,j}}{\sum_{i=1}^{n-j+1} D_{i,j-1}}$$

Les provisions totales estimées l'année  $n$  correspondent à la somme des provisions pour chaque année d'origine  $i$  :

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

## Le Bootstrap

Nous réalisons un bootstrap sur les résidus des estimations afin d'obtenir non plus une estimation des provisions mais une distribution empirique de ces provisions. Cette distribution nous permet d'avoir un certain nombre de statistiques sur les estimations (quantile...) et de pouvoir différencier l'erreur d'estimation de la variance du processus (erreur de prédiction). Le bootstrap est établi à partir des résultats de la méthode de Chain-Ladder (cf Etape 1).

Nous reprenons l'algorithme de bootstrap développé par England et al [16] pour Poisson sur-dispersé qui est équivalent au modèle de Chain-Ladder lorsque tous les paramètres de ligne et de colonne sont conservés (cf section suivante sur GLM). Cette méthode ne s'applique que si la somme des paiements incrémentaux par période de développement est positive, sinon la variance du processus de Poisson serait négative.

La procédure Bootstrap se fait en deux étapes :

**Etape 1 :** L'étape 1 consiste à calculer les résidus sur notre triangle de payés.

1. Calcul des facteurs de développements de Chain-ladder  $\hat{\lambda}_j$
2. Par récursion inverse, calcul des montants cumulés ajustés :

$$\hat{D}_{i,n-i+1} = D_{i,n-i+1} \text{ pour la diagonale}$$

$$\hat{D}_{i,k-1} = \frac{1}{\hat{\lambda}_k} \hat{D}_{i,k}, \quad i \in \{1, \dots, n-1\} \text{ et } k \in \{n-i+1, \dots, 2\}$$

3. Déduction des montants incrémentaux ajustés sur le nouveau triangle :

$$\hat{m}_{i,j} = \hat{C}_{i,j} = \hat{D}_{i,j} - \hat{D}_{i,j-1}$$

4. Calcul des résidus de Pearson : Le choix s'est porté sur ces résidus car ils sont simples à implémenter et sont normalisés, c-à-d. asymptotiquement, ces résidus suivent une loi normale (cf annexe B.2.1).

$$r_{i,j}^p = \frac{C_{i,j} - \hat{m}_{i,j}}{\sqrt{\hat{m}_{i,j}}}$$

Les résidus sont asymptotiquement gaussiens.

5. Calcul du coefficient de sur-dispersion, le paramètre d'échelle, qui est soit constant :

$$\Phi^{const} = \frac{\sum_{\substack{i,j=1 \\ i+j \leq n+1}}^{n-i+1} (r_{i,j}^p)^2}{\frac{1}{2}n(n+1) - 2n + 1}$$

où le dénominateur correspond au degré de liberté, c-à-d. au nombre d'observations  $Nb = \frac{1}{2}n(n+1)$  moins le nombre de paramètres  $p = 2n - 1$ .

soit variable par période de développement :

$$\Phi_j = \frac{\sum_{\substack{i,j=1 \\ i+j \leq n+1}}^{n-i+1} (r_{i,j}^p)^2}{n_j} \frac{\frac{1}{2}n(n+1)}{\frac{1}{2}n(n+1) - 2n + 1}, j \in \{1, \dots, n\}$$

avec  $n_j$  le nombre d'observations par période de développement.

Le paramètre d'échelle sera noté  $\Phi$  dans la suite et pourrait faire référence à l'un des deux paramètres ci-dessus en fonction de la paramétrisation choisie. Dans notre partie applicative, le paramètre d'échelle est choisi constant.

6. Les résidus finaux utilisés pour le bootstrap sont ajustés pour corriger le biais :

$$r_{i,j}^{fin} = \frac{r_{i,j}^p}{\sqrt{\Phi}} \sqrt{\frac{Nb}{Nb-p}}$$

**Etape 2 :** L'étape 2 consiste à ré-échantillonner les résidus  $N$  fois afin d'obtenir une distribution des provisions estimées. A chaque itération, nous réalisons les étapes suivantes :

1. Ré-échantillonnage des résidus ajustés avec remise pour créer un nouveau triangle supérieur de résidus. Les nouveaux résidus sont notés  $r'_{i,j}$ .
2. Calcul des pseudo montants incrémentaux associés en inversant la formule des résidus dans le cas d'une estimation non-paramétrique

$$C'_{i,j} = r'_{i,j} \sqrt{\Phi \hat{m}_{i,j} + \hat{m}_{i,j}}$$

ou en spécifiant une distribution, par exemple la loi normale, ainsi :

$$C'_{i,j} \sim \mathcal{N}(\hat{m}_{i,j}, \Phi \hat{m}_{i,j})$$

3. Calcul des pseudo montants cumulés

$$\hat{D}'_{i,j} = \sum_{k=1}^j C'_{i,k}$$

4. Calcul des nouveaux facteurs de développement associés à  $\hat{D}'_{i,j}$ , notés  $\hat{\lambda}'_j, j \in \{1, \dots, n\}$
5. Projection du futur triangle (inférieur) des paiements cumulés espérés :

$$\hat{D}'_{i,n-i+2} = \hat{\lambda}'_{n-i+2} \hat{D}'_{i,n-i+1}$$

et

$$\hat{D}'_{i,j} = \hat{\lambda}'_j \hat{D}'_{i,j-1}, j \in \{n-i+3, \dots, n\}$$

6. Différenciation pour avoir les paiements futurs incrémentaux notés  $\tilde{m}_{i,j}$  qui correspondent à la moyenne de la distribution du processus. La variance du processus est donnée par  $\Phi \tilde{m}_{i,j}$
7. Simulation d'un paiement futur à partir de la distribution supposée du processus et des paramètres ci-dessus. Par exemple, si la distribution du processus est normale, le triangle des paiements incrémentaux futurs sera :

$$T_{i,j} \sim \mathcal{N}(\tilde{m}_{i,j}, \Phi \tilde{m}_{i,j}), i \in \{2, \dots, n\}, j \in \{n-i+2, \dots, n\}$$

8. Sommation des paiements incrémentaux par année d'origine pour obtenir une valeur de la distribution des provisions pour chaque année d'origine.

Les distributions des provisions ainsi obtenues pour chaque année d'origine permettent d'obtenir des statistiques détaillées ainsi que la distribution des provisions totales en sommant les distributions des différentes années d'origine.

Nous remarquons que plusieurs choix peuvent être faits à différentes étapes. Ces derniers ont un impact sur la distribution finale des provisions et sur les erreurs d'estimation et de prédiction. D'une part la distribution choisie pour les pseudo-données peut engendrer une première source d'erreur ; ce choix intervient à l'étape 2.2. En général, une distribution non-paramétrique est ici privilégiée bien qu'il soit également possible de spécifier une loi connue et d'avoir ainsi une approche paramétrique. La variabilité des résultats due à ce choix est appelée variance d'estimation ou incertitude des paramètres. La distribution du processus peut être simulée par une loi Gamma, une loi Log-Normale, une loi Normale ou par rééchantillonnage. L'autre source de variabilité est liée à cette distribution généralement spécifiée (loi paramétrique mentionnée ci-dessus) mais qui peut aussi être non-paramétrique (cf étape 2.7). La variance du processus est donc la variance de la distribution prédite.

#### 8.1.4 Le modèle de Mack

Ce modèle basé sur la méthode des cadences (*link ratio*) a été introduit en 1993 par Thomas Mack [29].

##### Les hypothèses

Le modèle repose sur trois hypothèses, les hypothèses H1 et H2 étant les mêmes que pour le modèle de Chain-Ladder :

H1 : Les paiements cumulés  $D_{i,j}$  des années de survenance  $i$  sont indépendants.

H2 :  $\forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, n-1\}$ , il existe  $\lambda_{j+1}$  tels que

$$\mathbb{E}[D_{i,j+1} \mid D_{i,1}, \dots, D_{i,j}] = \mathbb{E}[D_{i,j+1} \mid D_{i,j}] = \lambda_{j+1} D_{i,j}$$

H3 :  $\forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, n-1\}$ , il existe  $\sigma_{j+1}^2$  tels que  $\mathbb{V}[D_{i,j+1} \mid D_{i,j}] = \sigma_{j+1}^2 D_{i,j}$ .

*Remarque.*

Avantages : Ce modèle nous permet de déterminer l'erreur de prédiction ainsi que des intervalles de confiance pour nos provisions.

Inconvénients : Ayant des hypothèses similaires au modèle de Chain-Ladder, ce modèle présente les mêmes limites. Une limite non-mentionnée ci-dessus mais présente dans les deux modèles concerne l'estimation des provisions pour les dernières années de survenance puisqu'il y a très peu de données pour vérifier les hypothèses de même que pour l'estimation du dernier facteur de développement.

##### Le modèle

Le facteur de développement  $\lambda_j$  est estimée par  $\hat{\lambda}_j$ , similairement au modèle de Chain-Ladder. Le paramètre de variance est estimé par

$$\begin{aligned} \hat{\sigma}_j^2 &= \frac{1}{n-j} \sum_{i=1}^{n-j+1} D_{i,j} (\lambda_{i,j} - \hat{\lambda}_j)^2 & \text{si } j < n-1 \text{ avec } \lambda_{i,j} = \frac{D_{i,j}}{D_{i,j-1}} \\ \hat{\sigma}_{n-1}^2 &= \min\left\{\frac{\hat{\sigma}_{n-2}^4}{\hat{\sigma}_{n-3}^2}, \min\{\hat{\sigma}_{n-3}^2, \hat{\sigma}_{n-2}^2\}\right\} & \text{sinon} \end{aligned}$$

##### Le bootstrap

Nous reprenons le déroulement du bootstrap pour la méthode Chain-Ladder dont nous changeons certains estimateurs à plusieurs étapes comme décrit ci-dessous.

**Étape 1** : La première étape est similaire à celle décrite pour le modèle de Chain-Ladder bien que les résidus changent (cf étape 1.4). En effet, les résidus de Mack sont donnés par :

$$r_{i,j} = \frac{D_{i,j+1} - \hat{\lambda}_j D_{i,j}}{\sqrt{D_{i,j} \hat{\sigma}_j^2}} \sqrt{\frac{Nb}{Nb-p}}$$

où

- $\hat{\sigma}_j$  : le paramètre de variance.
- $\hat{\lambda}_j$  : le facteur de développement de Chain-Ladder.

**Etape 2 :** A l'étape 2.2, nous simulons des pseudo ratios de passage plutôt que des pseudo montants incrémentaux. Si nous choisissons une approche non-paramétrique, les pseudo ratios sont :

$$\lambda_{i,j} = r'_{i,j} \sqrt{\frac{\hat{\sigma}_j^2}{\tilde{D}_{i,j-1}}} + \hat{\lambda}_j$$

avec  $r'_{i,j}$  les résidus après ré-échantillonnage.

Sinon

$$\lambda_{i,j} \sim \mathcal{L}(\hat{\lambda}_j, \frac{\hat{\sigma}_j^2}{\tilde{D}_{i,j-1}})$$

où

- $\mathcal{L}$  une loi paramétrique connue
- $\hat{\lambda}_j$  la moyenne
- $\frac{\hat{\sigma}_j^2}{\tilde{D}_{i,j}}$  la variance

Les facteurs de développement associés sont donc

$$\hat{\lambda}'_j = \frac{\sum_{i=1}^{n-j+1} D_{i,j-1} \lambda_{i,j}}{\sum_{i=1}^{n-j+1} D_{i,j-1}}$$

Ainsi, nous pouvons projeter le triangle des paiements cumulés initial afin d'obtenir la partie inférieure du triangle telle que :

$$D'_{i,n-i+2} = \hat{\lambda}'_{n-i+2} D_{i,n-i+1}$$

et

$$D'_{i,j} = \hat{\lambda}'_j D'_{i,j-1}, j \in \{n-i+3, \dots, n\}$$

La moyenne et la variance des montants cumulés projetés sont respectivement  $\mu_{i,j} = D'_{i,j}$  et  $\sigma_{i,j}^2 = D'_{i,j-1} \hat{\sigma}_j^2$ .

Comme à l'étape 2.7, nous simulons un montant cumulé, noté  $S_{i,j}$ , pour chaque année d'origine et période de développement soit par une approche non-paramétrique :

$$S_{i,j} = D'_{i,j} + r'_{i,j} \sqrt{D'_{i,j-1} \hat{\sigma}_j^2}$$

Soit par une approche paramétrique :

$$S_{i,j} \sim \mathcal{L}(D'_{i,j}, D'_{i,j-1} \hat{\sigma}_j^2)$$

Pour finir, nous en déduisons les provisions estimées  $\hat{R}_i$  en différentiant le montant cumulé estimé pour l'année d'origine  $i$  et la période de développement  $n$  et le montant cumulé sur la diagonale.

### 8.1.5 Les Modèles Linéaires Généralisés (GLM)

Les modèles linéaires généralisés (GLM) ont été introduits en 1972 par Nelder et al. [34]. Ils permettent de faire des hypothèses explicites quant à la nature des données et aux liens entre les variables explicatives  $X$  et expliquées (ou de réponse)  $Y$ .

## Modèle et hypothèses

Nous rappelons la définition de la famille exponentielle, utilisée dans la suite de notre exposé.

**Définition 8.1** (famille exponentielle). Soit  $Y$  une variable aléatoire ayant une distribution appartenant à la famille exponentielle, alors sa densité est donnée par :

$$f(y, \theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{a\left(\frac{\phi}{w}\right)} + c\left(y, \frac{\phi}{w}\right)\right)$$

avec

- $a(\cdot) > 0$  et  $b(\cdot)$  des fonctions deux fois dérivables et  $c(\cdot, \cdot)$  une fonction de normalisation de la densité.
- $\theta$  le paramètre dit *naturel*
- $\phi > 0$  le paramètre d'échelle aussi appelé paramètre de dispersion ou de nuisance
- $w$  un vecteur de poids évalués en fonction de la crédibilité de chaque observation

*Remarque.* La plupart des lois usuelles appartiennent à la famille exponentielle notamment les distributions gaussiennes, gamma, binomiale, Poisson...

Le modèle repose sur trois hypothèses ou trois composantes :

H1 : (composante aléatoire) Les composantes  $Y_{i,j}$  de la variable de réponse  $Y$  sont indépendantes entre elles et suivent une loi appartenant à la famille exponentielle définie par des paramètres différents.

H2 : (composante systématique) Le prédicteur linéaire  $\eta$  est donné par la relation

$$\eta = X\beta$$

avec

- $X$  la matrice des variables explicatives
- $\beta$  un vecteur de paramètres de la forme  $\beta = (u, o_1, \dots, o_n, d_1, \dots, d_n, c_1, \dots, c_k)$  avec  $u$  l'intercept et  $(o_i)_{i \in \{1, \dots, n\}}, (d_i)_{i \in \{1, \dots, n\}}$  et  $(c_i)_{i \in \{1, \dots, k\}}$  les effets liés à l'année d'origine (ligne), à la période de développement (colonne) et les effets calendaires respectivement.

H3 : (la fonction de lien) La relation fonctionnelle entre la variable expliquée  $Y$  et le prédicteur linéaire  $\eta$  est décrite par :

$$\mathbb{E}[Y] = \mu = g^{-1}(\eta) = b'(\theta)$$

où  $g$  est la fonction de lien différentiable et monotone.

*Remarque.*

1. Les lois de la famille exponentielle ont la propriété intéressante suivante : la variance peut s'écrire en fonction de la moyenne telle que :

$$\mathbb{V}(Y) = \frac{\phi V(\mu)}{w} = \frac{\phi b''(\theta)}{w}$$

où  $V(\cdot)$  est une fonction donnée, appelée *variance function*. Les formes usuelles associées à la loi choisie pour modéliser l'erreur sont présentées ci-dessous.

Loi	$V(x)$
<i>Normale</i>	1
<i>Poisson</i>	$x$
<i>Gamma</i>	$x^2$

2. La fonction de lien est dite canonique si  $g(\mu) = \theta$ . Grâce au logiciel de calcul, l'utilisation du lien canonique n'est plus primordiale.

*Remarque.*

Avantages : Ce modèle permet d'obtenir une expression de la variance. Nous pouvons ainsi déterminer des intervalles de confiance.

Inconvénients : Certaines hypothèses sont fortes et pas toujours vérifiées notamment l'indépendance des incréments. Par ailleurs, cette méthode compte un nombre important de paramètres qu'il peut être difficile d'estimer de manière consistante si le nombre de données est faible.

### Déroulement de la méthode

Les paiements incrémentaux  $C_{i,j}$  sont les variables expliquées supposées indépendantes.

1. Spécification de la matrice des variables explicatives  $X$  et du paramètre  $\beta$ .
2. Choix d'une fonction de lien et de la forme de l'erreur (de la distribution de la famille exponentielle).
3. Estimation des paramètres en maximisant la log-vraisemblance.
4. Prédiction des valeurs des incréments futurs  $(\hat{\mu}_{i,j})_{\substack{i,j \in \{1, \dots, n\} \\ i+j > n+1}}$  (la partie inférieure du triangle) à l'aide des paramètres estimés et de la fonction de lien.
5. Calcul des provisions en sommant les incréments estimés.

### Modèles linaires généralisés bayésiens

Comme pour le bootstrap, le but est d'obtenir une distribution des provisions estimées en se basant sur la méthode des modèles linaires généralisés. Le déroulement se fait en deux étapes; la première étape consiste à estimer par GLM les paramètres à partir de l'information passée (triangle des payés) et la deuxième étape s'intéresse à l'obtention d'une distribution de ces paramètres. La méthode est décrite par England et al. [17]. La deuxième étape est réalisée à partir d'une méthode de Monte-Carlo par chaînes de Markov (MCMC); celle implémentée dans le logiciel de provisionnement utilisé par la SCOR, *ResQ*, est l'algorithme de Métropolis Hastings proposé en 1953 par Metropolis et al.[33].

Cet algorithme permet d'obtenir un échantillon  $x$  d'une distribution  $\Pi$  que nous ne sommes pas en mesure de simuler directement. Cela n'est possible que si nous pouvons évaluer une fonction  $f$ , appelée distribution proposée, proportionnelle à la densité de  $\Pi$ , notée  $\pi$ . Le principe est le suivant :

- On choisit une valeur initiale  $x_0$  pour notre échantillon et une densité conditionnelle de probabilité  $Q(y | x)$  préférablement symétrique (i.e.  $Q(x | y) = Q(y | x)$ ). Cette distribution permet de proposer la valeur suivante de l'échantillon sachant la précédente  $x$ . La distribution très souvent utilisée est la loi gaussienne multivariée centrée en  $y$  qui présente l'avantage d'obtenir un échantillon suivant une marche aléatoire. Cette affirmation peut d'ailleurs être vérifiée en traçant le graphe des simulations.
- Pour chaque itération  $i$ ,  $i \in \{1, \dots, I\}$  :
  1. Simulation de la prochaine valeur  $x^*$  à partir de la distribution  $Q(x^* | x_i)$
  2. Calcul du ratio d'acceptation  $a = \frac{f(x^*)}{f(x_i)} = \frac{\pi(x^*)}{\pi(x_i)}$
  3. Acceptation ou rejet : si  $a \leq 1$  alors  $x_{i+1} = x^*$  car le nouveau candidat  $x^*$  est plus probable que  $x_i$ ; sinon  $x_{i+1} = x_i$
- Il faut effectuer un nombre suffisant de simulation  $I$  pour que la méthode converge vers la distribution  $\Pi$ . Une période de stabilisation est généralement nécessaire (nous supprimons les  $n$  premiers points de la distribution) notamment lorsque le point initial se trouve dans une région de faible densité, les premières simulations peuvent suivre une distribution très différente.

Nous appliquons cette méthode sur les paramètres du modèle linéaire généralisé. Une hypothèse souvent établie pour ces modèles est la gaussianité des paramètres asymptotiquement. Ainsi, la distribution  $f$  que nous choisissons est la loi gaussienne multivariée. L'algorithme de Metropolis-Hastings est appliqué, dans cette étude, sur chaque paramètre indépendamment. Cependant, il est aussi possible

de simuler tous les paramètres par bloc à chaque itération.

La méthode se décompose en deux étapes. La première étape de la simulation consiste en la réalisation de la méthode MCMC ci-dessous pour obtenir une estimation des paramètres et ainsi la valeur projetée des provisions futures. La deuxième étape permet de prendre en compte l'erreur de processus à partir des paramètres estimés. Les valeurs futures obtenues sont resimulées à partir d'une distribution choisie (généralement en fonction de la distribution de l'erreur choisie au préalable) et l'erreur de prédiction dépend de ce choix. Cette étape est identique à l'avant-dernière étape du bootstrap.

## 8.2 Validation des modèles

Dans cette partie, nous allons tester que les hypothèses de chaque modèle sont bien vérifiées pour le segment que nous souhaitons analyser. Comme nous travaillons sur les montants payés agrégés par année d'origine des sinistres, nous disposons de peu de données. Du fait de ce manque de données, certaines hypothèses des modèles de provisionnement sont difficiles à vérifier en pratique.

### 8.2.1 Chain-Ladder et Mack

Nous nous sommes appuyés sur l'article de Mack [30] pour la vérification des hypothèses des modèles de Chain-Ladder et Mack.

#### Hypothèse H1

L'hypothèse H1 suppose l'indépendance des paiements cumulés entre les différentes années de souscription ; i.e.  $(C_{i_1,1}, \dots, C_{i_1,n}) \Pi (C_{i_2,1}, \dots, C_{i_2,n}), \forall i_1, i_2 \in \{1, \dots, n\}$ . Cette hypothèse peut être violée s'il existe des effets calendaires dus par exemple à une modification importante de l'inflation. Le test réalisé pour vérifier cette hypothèse est donc un test d'absence d'effets calendaires.

Un changement conséquent à la hausse (à la baisse) affectant une année calendaire aura un impact sur la diagonale à la hausse (à la baisse) correspondante du triangle des payés et ainsi sur le triangle des facteurs de développement qui seront plus grands (plus petits) sur la diagonale précédente et plus petits (plus grands) sur cette diagonale, d'après la définition des facteurs de développement.

#### Déroulement du Test

Nous commençons par trier les facteurs de développement entre les différentes années de survenance pour une période de développement. A partir du rang obtenu  $r_{i,j}$ , nous formons deux groupes, ceux ayant un rang supérieur à la médiane des rangs et ceux ayant un rang inférieur ; le rang médian n'étant pas classé. Nous calculons ensuite pour chaque diagonale (excepté la première qui ne contient qu'un élément) le nombre de facteurs  $L_k$  appartenant au groupe des rangs élevés et ceux  $S_k$  appartenant au groupe des rangs bas.

Nous avons l'intuition que, en l'absence d'effets calendaires, le nombre de rangs élevés et le nombre de rangs bas devraient être à peu près égaux pour chaque diagonale. Dans le cas contraire, cela signifierait que les facteurs de développement sont significativement plus grands ou plus faibles que les précédents. Ainsi, la variable  $Z_k = \min\{L_k; S_k\}$  doit être proche de  $\frac{L_k+S_k}{2}$  pour ne pas rejeter l'hypothèse d'absence d'effets calendaires. L'espérance et la variance de  $Z_k$  sont données par :

$$\mathbb{E}[Z_k] = \frac{n}{2} - C_{n-1}^m \frac{n}{2^n}$$

et

$$\mathbb{V}[Z_k] = \frac{n(n-1)}{2} - C_{n-1}^m \frac{n(n-1)}{2^n} + \mathbb{E}[Z_k] - \mathbb{E}[Z_k]^2$$

avec  $n = L_k + S_k$  et  $m = \lfloor \frac{n-1}{2} \rfloor$

L'hypothèse est rejetée à 5% si :

$$Z = \sum_{k=1}^{n-1} Z_k \notin IC_{95\%} = [\mathbb{E}[Z] - 1.96\sqrt{\mathbb{V}[Z]}, \mathbb{E}[Z] + 1.96\sqrt{\mathbb{V}[Z]}]$$

avec -1.96 et 1.96 la valeur du quantile gaussien à 5% et 95% respectivement car, sous l'hypothèse d'absence d'effets calendaires, la variable  $Z$  suit asymptotiquement une loi normale.

### Résultats

Le tableau ci-dessous présente les valeurs obtenues pour les différentes variables et années calendaires.

TABLE 19 – Test d'absence d'effets calendaires

<b>k</b>	$L_k$	$S_k$	$Z_k$	<b>n</b>	<b>m</b>	$\mathbb{E}[Z_k]$	$\mathbb{V}[Z_k]$
<b>2</b>	1	1	1	2	0	0,5	0,25
<b>3</b>	0	2	0	2	0	0,5	0,25
<b>4</b>	2	2	2	4	1	1,25	0,438
<b>5</b>	4	1	1	5	2	1,563	0,371
<b>6</b>	4	2	2	6	2	2,062	0,621
<b>7</b>	6	1	1	7	3	2,406	0,554
<b>8</b>	8	0	0	8	3	2,906	0,804
<b>9</b>	8	0	0	8	3	2,906	0,804
<b>10</b>	5	5	5	10	4	3,770	0,986
<b>11</b>	3	7	3	10	4	3,770	0,986
<b>12</b>	4	8	4	12	5	4,646	1,168
<b>13</b>	1	9	1	10	4	3,770	0,986
<b>14</b>	3	10	3	13	6	5,034	1,10
<b>Total</b>			<b>23</b>			<b>35,083</b>	<b>9,317</b>

Pour que l'hypothèse d'absence d'effets calendaires ne soit pas rejetée, il faut que  $Z_k$  appartienne à l'intervalle de confiance  $IC_{95\%} = [29.10, 41.07]$ . La valeur estimée étant égale à 23, l'hypothèse est rejetée.

### Hypothèse H2

L'hypothèse H2 concerne l'indépendance des facteurs de développement individuels pour deux périodes de développement successives et la même année de survenance. Cette hypothèse est vérifiée à l'aide d'un test de corrélation sur les rangs pour ne pas avoir à émettre d'hypothèses supplémentaires sur la distribution des facteurs. Pour calculer les coefficients de corrélation de rangs de Spearman, nous reprenons le triangle des rangs établis pour le test précédent, noté  $r_{i,j}$  et nous calculons les rangs pour chaque période de développement en enlevant le dernier facteur, noté  $s_{i,j}$ . On veut tester :

$$H_0 : S_j = (s_{i,j})_{i=1,\dots,n-j-1} \text{ et } R_{j+1} = (r_{i,j+1})_{i=1,\dots,n-j-1} \text{ sont non corrélées}$$

$$\text{contre } H_1 : S_j \text{ et } R_{j+1} \text{ sont en relation monotone}$$

La statistique est

$$\rho_j = \frac{\sum_{i=1}^{n-j} (r_{i,j} - \bar{r}_{.,j})(s_{i,j-1} - \bar{s}_{.,j-1})}{\sqrt{\sum_{i=1}^{n-j} (r_{i,j} - \bar{r}_{.,j})^2 \sum_{i=1}^{n-j} (s_{i,j-1} - \bar{s}_{.,j-1})^2}} = 1 - 6 \frac{\sum_{i=1}^{n-j} (r_{i,j} - s_{i,j})^2}{(n-j)^3 - n + j}$$

où  $\bar{r}_{.,j}$  et  $\bar{s}_{.,j-1}$  représentent les moyennes empiriques des rangs sur les années de survenance.

Remarque.

- $\rho_j \in [-1, 1]$
- $\mathbb{E}(\rho_j) = 0$
- $\mathbb{V}(\rho_j) = \frac{1}{n-j-1}$

Pour n'avoir qu'une statistique de test globale et non une statistique pour chaque couple de périodes de développement, nous considérons la statistique suivante qui est une somme pondérée des  $\rho_j$  :

$$\rho = \frac{\sum_{j=2}^{n-2} (n-j-1)\rho_j}{\frac{(n-2)(n-3)}{2}}$$

avec comme espérance  $\mathbb{E}(\rho) = 0$  et variance  $\mathbb{V}(\rho) = \frac{2}{(n-2)(n-3)}$ .

$\rho$  suit asymptotiquement une loi normale sous  $H_0$ , donc l'hypothèse nulle est rejetée à 5% si :

$$\rho \notin IC_{95\%} = [\mathbb{E}[\rho] - 1.96\sqrt{\mathbb{V}[\rho]}, \mathbb{E}[\rho] + 1.96\sqrt{\mathbb{V}[\rho]}]$$

### Résultats

Le tableau des rangs est présenté en annexe et celui des résultats ci-dessous (cf tableau 20).

TABLE 20 – Statistiques de test et intervalles de confiance

$\rho_j$	0,769	0,427	0,382	0,673	0,617	0,452	-0,357	0,314	-0,5	-0,4	1	1
<b>n-j-1</b>	12	11	10	9	8	7	6	5	4	3	2	1
$\mathbb{E}(\rho_j)$	0	0	0	0	0	0	0	0	0	0	0	0
$\mathbb{V}(\rho_j)$	0,083	0,091	0,1	0,111	0,125	0,143	0,167	0,2	0,25	0,333	0,5	1
<b>IC</b>	$\pm 0,57$	$\pm 0,59$	$\pm 0,60$	$\pm 0,65$	$\pm 0,69$	$\pm 0,74$	$\pm 0,80$	$\pm 0,88$	$\pm 0,98$	$\pm 1,13$	$\pm 1,39$	$\pm 1,96$

$\rho$	0,399
$\mathbb{E}(\rho)$	0
$\mathbb{V}(\rho)$	0,0128
<b>IC</b>	$\pm 0,222$

Nous remarquons que l'hypothèse  $H_0$  n'est pas rejetée dans plusieurs cas si nous nous intéressons aux statistiques  $\rho_j$ . Cependant, l'hypothèse est rejetée quant à l'indépendance des facteurs de développement entre les deux premières périodes de développement. Ainsi, lorsque nous nous intéressons à la somme pondérée des statistiques,  $\rho$ , nous remarquons que l'hypothèse n'est plus vérifiée de manière globale. Cette conclusion s'explique par la pondération donnée aux différentes périodes de développement qui donne plus de poids aux premières années puisqu'il y a plus de données. Les dernières années ne sont pas assez développées pour que la théorie du test reposant sur des résultats asymptotiques de normalité de la statistique soit vérifiée.

Cette hypothèse s'intéresse aussi à la forme de l'espérance qui est linéaire. Nous testons cette hypothèse de manière graphique en étudiant la relation entre les paiements cumulés pour deux périodes de développement successives pour les différentes années de survenance.

Nous notons *régression j* le graphique représentant  $D_{i,j}$  en fonction de  $D_{i,j+1}$  pour chaque année d'origine. Pour chaque "régression", nous avons tracé la droite issue de la régression linéaire effectuée sur ces points (cf figure 17 ci-dessous).

Les résultats obtenus sur les régressions des facteurs de développement des périodes suivantes sont présentés en annexe B.1.2. Nous observons que l'hypothèse semble globalement vérifiée même s'il existe une ou deux périodes pour lesquelles l'existence d'une relation linéaire est mise en doute (ex : régression 5). Nous remarquons par ailleurs que l'existence de relations linéaires semble se vérifier davantage sur les périodes de développement les plus récentes. Cependant, cela peut s'expliquer par la diminution du nombre de données dans les années de développement les plus anciennes.

### Hypothèse H3

L'hypothèse H3 concerne la forme de la variance et n'est faite que dans le cas du modèle de Mack. Pour cela, nous regardons le caractère aléatoire des résidus de Mack,  $r_{i,j}^{Mack}$ , définis dans la description des modèles 8.1.4. Nous représentons donc  $r_{i,j}^{Mack}$  en fonction de  $D_{i,j}$  (cf figure 18 ci-dessous). L'idée sous-jacente est que si la variance supposée n'est pas bien spécifiée alors les résidus ne seront pas gaussiens car la variance ne prendra pas en compte tous les effets stochastiques.

Nous n'observons pas de tendance particulière des résidus, bien que sur les dernières périodes nous ne puissions pas réellement conclure par manque de données. Certains résidus prennent des valeurs un peu élevées ou basses mais ces valeurs extrêmes ne se présentent pas en grappe et ne viennent pas réfuter le caractère aléatoire des résidus. L'hypothèse H3 n'est donc pas rejetée.

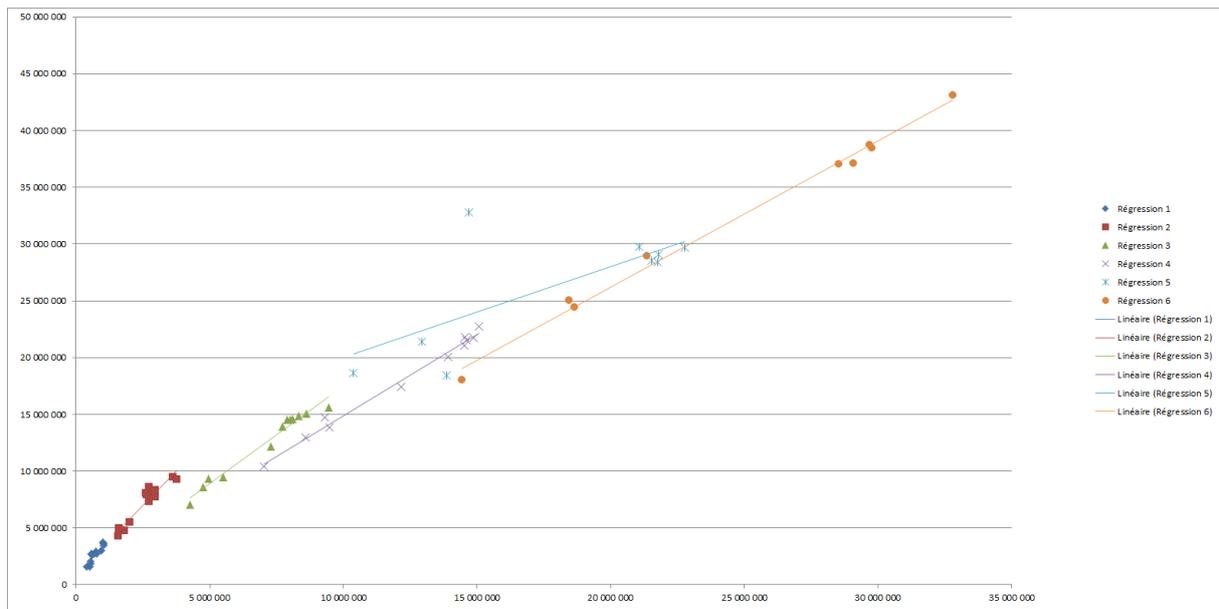


FIGURE 17 – Relation entre les paiements cumulés de deux périodes de développement successives

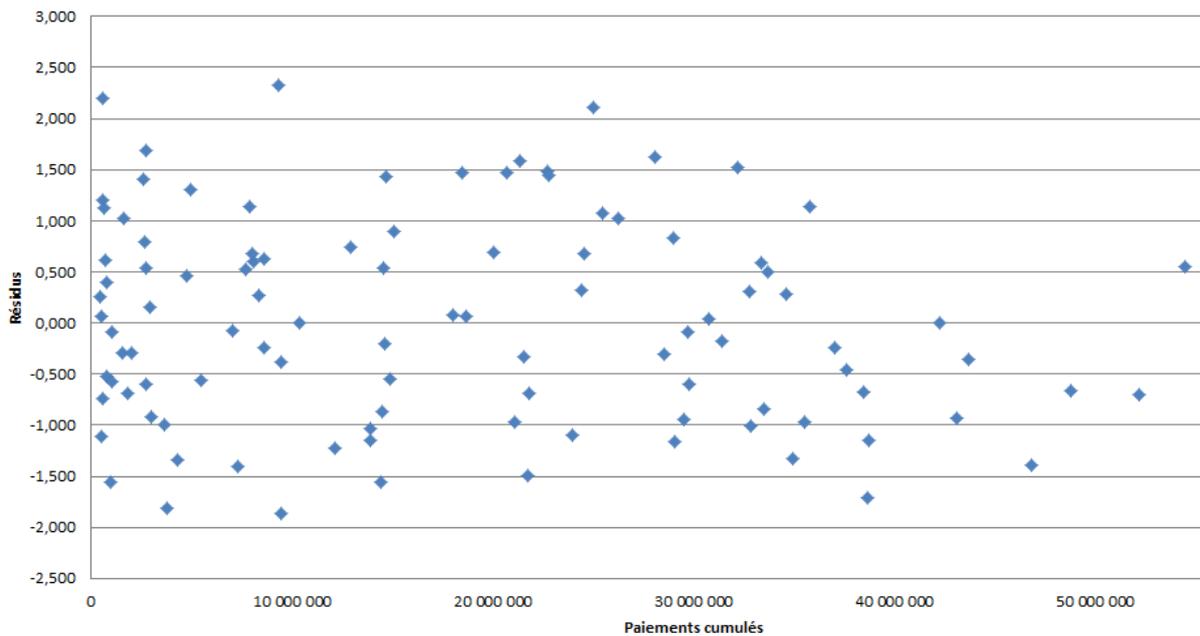


FIGURE 18 – Résidus de Mack en fonction des paiements cumulés

Finalement, nous avons pu observer que toutes les hypothèses n'étaient pas vérifiées en particulier l'hypothèse d'absence d'effets calendaires. Nous disposons de trop peu de données ici pour que les hypothèses soient pleinement satisfaites. Nous garderons tout de même ces modèles dans la suite de l'étude dans la mesure où ils sont couramment utilisés en pratique. Par ailleurs, quel que soit le modèle, nous ne disposons pas de suffisamment de données pour juger correctement de son adéquation.

### 8.2.2 GLM

Nous avons sélectionné pour notre étude trois formes de modèles linéaires généralisés qui représentent les modèles les plus utilisés. De plus, ils ont l'avantage de donner des distributions positives des provisions.

GLM	Lien	Distribution de l'erreur
1	log : $g(x)=\ln(x)$	Poisson sur-dispersé (paramètre de dispersion $\phi$ constant)
2	log : $g(x)=\ln(x)$	Gamma
3	identité : $g(x)=x$	Log-normale

Pour chaque modèle, nous avons décidé de réaliser un certain nombre de test sur les résidus de Pearson qui ont l'avantage d'être normalisés. Ainsi, nous allons vérifier que les résidus sont bien aléatoires de loi normale et que le choix du lien est approprié. Nous avons par ailleurs évalué les modèles pour plusieurs prédicteurs linéaires  $X\beta$ , chacun ayant une forme additive mais pas le même nombre de paramètres à estimer.

### Significativité des paramètres

Les résultats suivant sont présentés pour le modèle 'de base' comprenant tous les paramètres représentant les effets lignes et colonnes. En effet, pour chaque GLM, nous avons regardé deux autres configurations avec moins de paramètres estimés. Cependant, les résultats obtenus concernant l'adéquation des données et les résidus étant sensiblement identiques, nous n'avons traité, pour rendre l'exposé plus concis, que le modèle 'de base' pour valider les hypothèses. En pratique, les validations doivent être faites pour tous les modèles car l'ajout ou la suppression de paramètres peut changer considérablement l'adéquation.

- 1ère configuration : nous avons modélisé le même modèle en otant les paramètres non-significatifs. La significativité des paramètres a été testée à l'aide d'un test de Wald de significativité jointe (ou F-test). Ce test repose sur les hypothèses :

$$H_0 : \beta = 0 \text{ vs } H_1 : \beta \neq 0$$

Ce test est évalué par le logiciel *ResQ* et les paramètres non significatifs sont exclus du modèle.

- 2ème configuration : nous avons ajouté des effets calendaires. Le principe est de supposer la présence d'effets calendaires. A partir du modèle précédent n'ayant que les paramètres significatifs de lignes et colonnes, nous ajoutons un paramètre par effet calendaire (par diagonale) puis nous effectuons un F-test sur ce nouveau modèle afin de ne garder que les paramètres significatifs.

### L'adéquation des données

Pour tester l'adéquation du modèle à nos données, nous nous intéressons à deux statistiques :

- La statistique du  $\chi^2$  de Pearson : Cette statistique, définie par  $P = \sum_{i,j=1, i+j \leq n+1}^n \frac{(Y_{i,j} - \mu_{i,j})^2}{V(\mu_{i,j})}$ , permet de mesurer l'écart entre la valeur observée et la valeur estimée.
- La déviance : Cette statistique compare la log-vraisemblance du modèle saturé ; i.e. le modèle exact qui a autant de paramètres que d'observations, et la log-vraisemblance du modèle estimé. Elle est définie par  $D = -2 \sum_{i,j=1, i+j \leq n+1}^n (\mathcal{L}(Y_{i,j}, \hat{\beta}, \phi) - \mathcal{L}_{saturé}(Y_{i,j}))$

Sous l'hypothèse d'adéquation des données, ces statistiques suivent asymptotiquement, toutes deux, une loi du  $\chi^2$  à  $Nb-p$  degrés de liberté où  $Nb$  est le nombre d'observations et  $p$  le nombre de paramètres. Nous présentons les résultats de ces statistiques standardisées ( en divisant par le paramètre d'échelle) ci-dessous :

	GLM Poisson	GLM Gamma	GLM Log - normal
$\chi^2$ Pearson	91	91	91
Déviance	91.3	93.3	91

Le nombre de degrés de liberté étant 91, la statistique doit être dans l'intervalle de confiance à 95% (arrondi) [70; 130]. Cela est vrai pour chacune des statistiques donc l'hypothèse d'adéquation des estimations aux observations n'est pas rejetée.

Pour tester qualitativement cette hypothèse, nous comparons les valeurs observées avec les valeurs ajustées obtenues à partir des paramètres estimés. Si le lien est choisi de manière appropriée, il existe une approximation linéaire des valeurs observées après application de la fonction de lien. Nous traçons donc les valeurs observées transformées (en log) en fonction du prédicteur linéaire (cf graphe 19 pour le GLM Poisson) ; si une telle relation existe les points doivent être sur la première bissectrice du plan.

Les graphes pour les GLM Gamma et Log-normal sont présentés en annexe B.2.2 et B.2.3. La linéarité est clairement observée, ainsi les valeurs estimées par le modèle sont très proches des valeurs observées ce qui nous permet de justifier l'utilisation de ce modèle. Par ailleurs la droite de la régression linéaire est très proche de la première bissectrice comme le montre aussi les coefficients de la régression. Les mêmes conclusions sont établies sur les deux autres modèles.

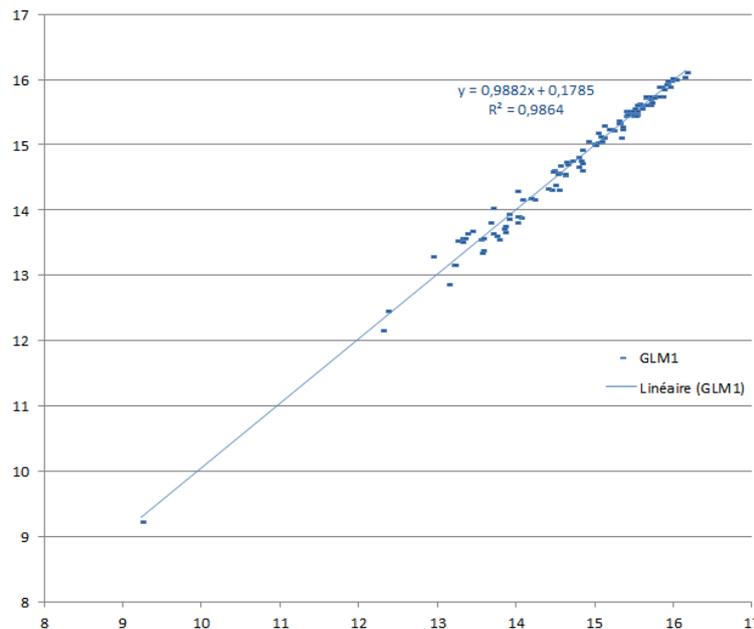


FIGURE 19 – Valeurs observées vs valeurs ajustées pour le *GLMPoisson*

### Les résidus

Les résidus correspondent à la différence entre l'observation  $Y_{i,j}$  et sa moyenne  $\mathbb{E}(Y_{i,j})$  qui est estimée par  $\mu_{i,j}$ . Pour normaliser ces résidus (prendre en compte l'hétéroscédasticité), nous utilisons les résidus de Pearson (cf bootstrap Chain-Ladder 8.1.3). Asymptotiquement, pour une fonction de variance (distribution de l'erreur) et un paramètre d'échelle bien choisis, ces résidus suivent une loi normale (cf annexe B.2.1). Nous présentons les résultats pour les prédicteurs linéaires comprenant tous les paramètres modélisant les effets lignes et colonnes.

Tout d'abord, nous regardons graphiquement le caractère aléatoire des résidus en présentant les résidus en fonction des paiements incrémentaux.

Les résidus des GLM Gamma et Log-normal sont présentés en annexe B.2.2 et B.2.3. Nous constatons que les résidus du GLM Poisson semblent aléatoires autour de 0. Par ailleurs, il n'y a pas trop de résidus extrêmes, i.e. de résidus à l'extérieur de l'intervalle de confiance à 99%  $[-2.576; 2.576]$ . Pour les deux autres modèles les résidus semblent présenter une tendance décroissante avec le montant des incréments. Cette constatation ne nous permet pas d'infirmer l'hypothèse sur les résidus ; c'est pourquoi nous avons comparé l'histogramme des résidus avec la densité gaussienne, réalisé les graphiques des quantiles empiriques en fonction des quantiles théoriques de la loi normale et, pour finir, effectué un test de Shapiro-Wilk pour avoir un outil de validation plus quantitatif.

Nous ne présentons que les résultats pour le GLM Poisson, les autres étant présentés en annexe B.2.2 et B.2.3.

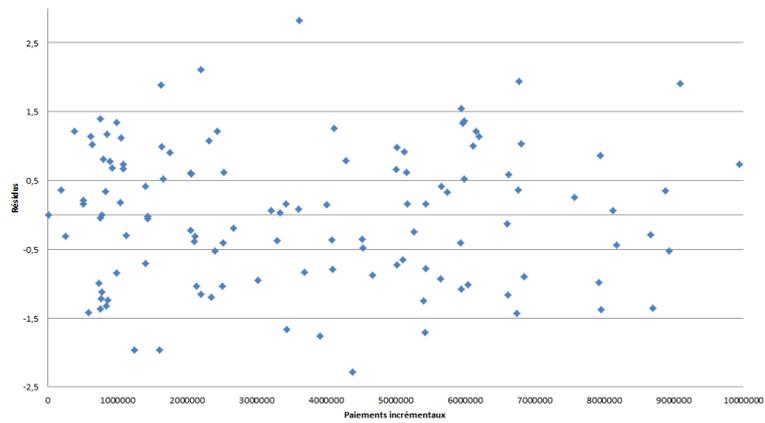


FIGURE 20 – Résidus de Pearson du modèle linéaire généralisé 1 (lien 'log' distribution de Poisson)

FIGURE 21 – Comparaison de l'histogramme des résidus du GLM Poisson avec la densité Gaussienne

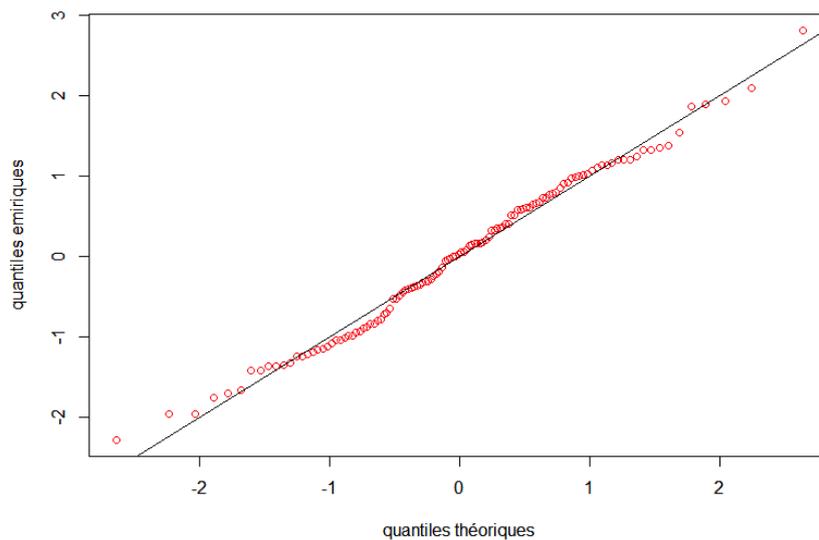
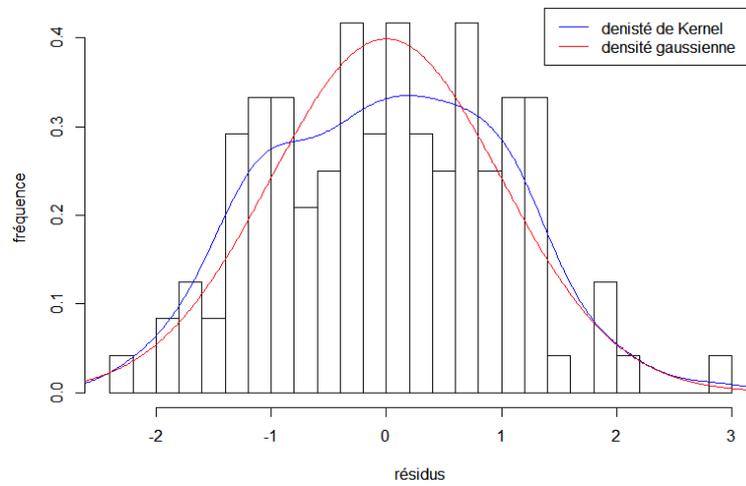


FIGURE 22 – Test quantiles-quantiles de normalité des résidus pour le GLM Poisson

Sur l'histogramme lié de la loi de Poisson, nous constatons que la normalité des résidus n'est pas évidente et la densité estimée à l'aide d'un estimateur à noyau n'est pas gaussienne et possède deux

modes. Concernant le GLM Gamma, les deux densités coïncident davantage et l’histogramme est plus proche de celui d’une variable gaussienne. Pour le dernier modèle, l’histogramme est assez éloigné de celui d’une variable gaussienne bien que les deux densités soient proches, la densité estimée étant décentrée et ayant une queue de distribution gauche plus lourde. Les graphiques quantiles-quantiles ne permettent pas de contredire l’hypothèse de normalité puisque les quantiles empiriques sont très proches de ceux de la loi normale excepté un ou deux quantiles extrêmes qui s’éloignent un peu des quantiles théoriques.

Le tableau ci-dessous présente la valeur de la statistique de Shapiro-Wilk pour chacun des modèles ainsi que la *p-value* associée. Les résultats confirment en partie les constatations graphiques. En effet, au vu du test, les résidus pour le modèle Gamma sont plus vraisemblablement gaussiens mais ce test nous indique aussi que l’hypothèse de normalité des résidus n’est rejetée pour aucun modèle.

Modèle	Shapiro	
	Statistique	p-value
<b>1</b>	0.989	48.47%
<b>2</b>	0.993	72.59%
<b>3</b>	0.987	29.53%

Pour conclure, le caractère aléatoire et gaussien des résidus n’est pas toujours clair graphiquement mais en nous référant à l’aspect quantitatif nous ne pouvons pas infirmer cette hypothèse. Nous ne gardons pas le GLM Log-normal dont les hypothèses sur les résidus ne semblent pas vérifiées qualitativement. Les GLM et plus particulièrement les deux autres modèles semblent appropriés pour l’estimation des provisions pour ce segment ; nous les sélectionnons donc pour l’évaluation de l’erreur de modèle.

### 8.3 Evaluation de l’erreur de modèle

Parmi tous les modèles présentés ci-dessus, nous n’avons décidé de ne garder que ceux qui nous semblaient pertinents. Nous rappelons que l’ensemble des modèles doit contenir un nombre suffisant de modèles pour comprendre le ‘vrai’ modèle mais ne doit pas être alourdi par des modèles non pertinents.

Nous avons donc défini trois classes de modèles : les modèles MCMC GLM, les modèles de bootstrap Mack et les modèles de bootstrap Chain-Ladder. Parmi les modèles avec bootstrap, nous avons décidé de ne pas nous attarder sur l’erreur d’estimation puisque ce n’est pas le sujet de cette étude ; par conséquent l’estimation sera effectuée par ré-échantillonnage (approche non-paramétrique). La différence entre les modèles proviendra du choix de la distribution prédictive (erreur de processus) qui sera soit Gamma, Log-Normale, Normale ou ré-échantillonnée. Concernant les modèles GLM, nous n’avons pas gardé le modèle GLM Log-normal (lien ‘identité’ et loi log-normale) puisque ce modèle présentait les moins bons résultats aux tests de validation. Par ailleurs, nous avons retiré les modèles avec prise en compte des effets calendaires principalement car ces modèles sont très différents de ceux retenus dans la mesure où ils prennent en compte une troisième dimension à savoir l’évolution future des tendances calendaires. Par ailleurs, les effets calendaires doivent être mesurés sur les triangles désinflatés pour éviter d’ajouter un biais dû à l’inflation. Nous préférons donc ne considérer que les modèles ayant des paramètres représentatifs des effets lignes et colonnes, c-à-d. le modèle avec lien ‘log’ et distribution Poisson sur-dispersé et le modèle avec lien ‘log’ et distribution Gamma. Le graphe 23 présente la classification des modèles.

Comme pour les modèles d’actifs, nous utilisons deux approches différentes, pour mesurer l’erreur de modèle, une approche par pondération des classes et des modèles de chaque classe et une approche dite ‘mixte’ plus semblable à l’approche ‘worst case’ dans laquelle l’erreur de modèle est évaluée intra-classe par rapport au pire cas puis agrégée extra-classe à l’aide des pondérations de chaque classe.

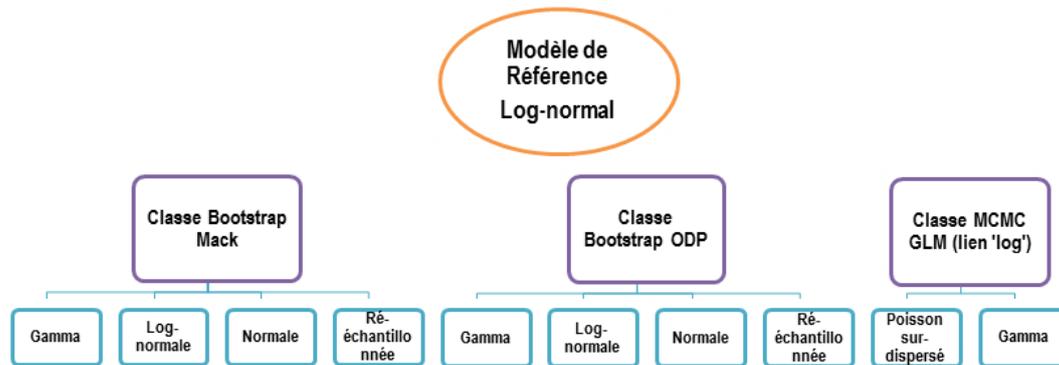


FIGURE 23 – Classification des modèles de provisionnement

### 8.3.1 Modèle de référence

Le modèle 'dit' *log-normal* qui n'est autre qu'une simulation des provisions totales à l'aide d'une loi log-normale sera le modèle de référence dans cette étude puisque c'est le modèle qui est souvent utilisé en pratique pour évaluer le coefficient de variation des provisions estimées conformément aux exigences réglementaires. Les paramètres de la loi log-normale sont estimés par le modèle de Mack. Ainsi, au vu de tests d'adéquation mentionnés ci-dessous, même si la distribution des provisions n'est pas identique à celle du bootstrap Mack, elle s'en approche fortement.

Nous présentons ci-dessous quelques statistiques mesurées sur la distribution des provisions obtenues à l'aide de ce modèle, en particulier le quantile à 99.5% qui servira de mesure de référence.

	Paramètres de la loi		Moyenne Mack	Ecart-type Mack	VaR 99.5%
	$\mu$	$\sigma^2$			
loi log-normale	19.53	0.21%	303.7M€	13.8M€	340.7M€

*Remarque.* Il est à noter que le risque sous-jacent ici est une sous-estimation des provisions due à un modèle mal spécifié. Ainsi, nous ne nous intéressons ici, non plus au quantile à 0.5% comme pour les actifs, mais au quantile à 99.5%.

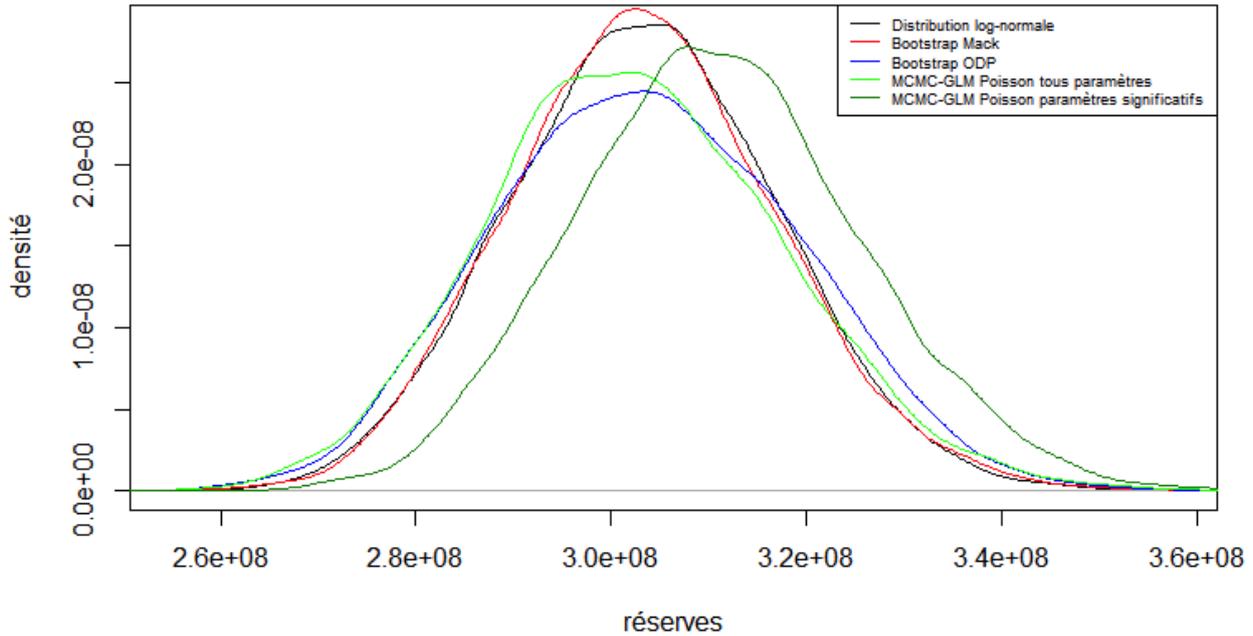
Ce modèle est choisi comme référence de par sa simplicité de mise en œuvre. En effet, ce modèle assure des résultats fiables au vu de nombreux tests de validation réalisés en interne et un temps de calcul optimal.

Pour vérifier le bien-fondé de ce modèle, nous représentons ci-dessous les densités estimées des provisions finales de cinq modèles : le modèle bootstrap Mack avec distribution prédictive Gamma, le modèle bootstrap ODP avec distribution prédictive Gamma, le MCMC-GLM poisson avec distribution prédictive Gamma comprenant tous les paramètres et celui ne comprenant que ceux significatifs et pour finir le modèle 'log-normal'.

Nous constatons que le modèle 'log-normal' est très proche du modèle de Mack mais qu'il diffère des autres modèles notamment au niveau de la queue droite de la distribution qui est plus fine. Les tests d'adéquation de Kolmogorov-Smirnov et d'Anderson-Darling entre les différentes distributions et la loi log-normale confirme ce résultat puisque seule la distribution issue du bootstrap Mack ne permet pas de rejeter l'hypothèse d'adéquation à la loi log-normale (cf annexe B.3.1). Cela coïncide avec nos attentes puisque les paramètres de la loi log-normale ont été obtenus à partir des paramètres du modèle de Mack (avant ré-échantillonnage).

Tous les modèles ont une moyenne assez similaire sauf le modèle de GLM où seuls les paramètres significatifs ont été retenus. Ce modèle ne comporte plus que 8 paramètres représentatifs des effets lignes et 6 pour les effets colonnes au lieu de 14 pour chaque type d'effet dans le modèle avec tous les paramètres. Ce modèle est sensé être plus adéquat que celui prenant en compte tous les paramètres et

FIGURE 24 – Densités de Kernel des distributions des provisions totales pour différents modèles



possède d'ailleurs une plus grande vraisemblance. Cependant, ce triangle des sinistres payés n'a pas été retraité, de l'inflation notamment, et certains effets peuvent ne pas être très bien modélisés à cause de ce biais. Il faut donc prendre des précautions concernant l'utilisation de ce modèle.

La plus grande disparité entre les résultats des deux modèles provient de la modélisation des provisions des deux dernières années d'origine puisque nous possédons très peu d'information sur les sinistres survenus ces années-là. Le premier modèle estime des paramètres non significatifs sur les dernières années tandis que le deuxième se base sur la tendance des années précédentes (qui peut être biaisée) pour déterminer ces paramètres. La modélisation des provisions des dernières années est un point difficile en provisionnement.

Le modèle GLM avec tous les paramètres n'est autres qu'un modèle ODP, les variations dans les résultats n'étant que le fruit des simulations ; c'est pourquoi, nous ne traiterons que du modèle GLM avec uniquement les paramètres significatifs qui apporte une véritable nouveauté.

### 8.3.2 L'approche mixte

Nous présentons les deux étapes de développement de cette approche ainsi que les résultats pour chaque classe.

Etape 1 : Evaluation de l'erreur de modèle définie comme suit :

$$AM = \frac{\max\{\mathcal{R}(M_{C_i,1}), \mathcal{R}(M_{C_i,2}), \dots, \mathcal{R}(M_{C_i,n})\}}{\mathcal{R}(M^*)} - 1$$

avec  $M_{C_i,j}$  le j-ième modèle de la classe  $C_i$  comprenant  $n$  modèles,  $M^*$  le modèle de référence log-normal et

$\mathcal{R} = VaR_{99.5\%}$  le quantile à 99.5% de la distribution des provisions estimées.

Ainsi nous obtenons les résultats ci-dessous :

		<i>Bootstrap ODP</i>	<i>Bootstrap Mack</i>	<i>MCMC GLM</i>
AM	en %	1.61%	0.31%	6.90%
	en valeur	5.5M€	1.1M€	23.5M€

TABLE 21 – Erreur de modèle *worst case* intra-classe

Etape 2 : Agrégation extra-classe en fonction des pondérations de chaque classe. Nous choisissons, dans un premier temps, la pondération équiprobable (1/3,1/3,1/3) en considérant que nous n'avons pas d'information sur la vraisemblance de chaque classe. L'erreur de modèle totale est donc de **2.941%** soit **10M€**.

Afin d'évaluer l'impact de la pondération des classes sur l'erreur globale, nous effectuons des tests de sensibilité. En fonction des poids des classes, l'erreur de modèle sera comprise entre 0.31% et 6.90%. Ainsi, l'erreur de modèle sera d'autant plus élevée (faible) que l'on donne du poids au modèle GLM (Mack).

Dans cet exemple, l'erreur de modèle est sensible au choix des pondérations des classes, il semble donc important d'avoir de l'information supplémentaire sur ces classes (par exemple des avis d'experts). Bien que les GLM puissent paraître plus adaptés que les modèles ODP puisqu'ils ne prennent en compte que les paramètres significatifs, nous avons vu que cela était à nuancer du fait du non-retraitement des données.

### 8.3.3 L'approche par pondération

Dans cette approche, seule la première étape diverge. Au lieu de calculer l'erreur de modèle par rapport au pire des cas, nous pondérons chaque modèle d'une classe. Cette technique vise à provisionner en tenant compte de la pertinence des modèles et non plus uniquement par rapport au modèle le plus prudent. La charge évaluée au titre du risque de modèle est donc inférieure à celle calculée par l'approche précédente.

L'erreur de modèle est ici obtenue par :

$$\rho(M_{C_i}) = \frac{\sum_{j=1}^n p_j \mathcal{R}(M_{C_i,j})}{\mathcal{R}(M^*)} - 1$$

Comme nous n'avons pas d'éléments permettant de favoriser une distribution prédictive par rapport à une autre pour les modèles de bootstrap, nous effectuons la pondération par rapport à l'erreur de prédiction obtenue (le CoV) et en particulier à l'erreur de processus (l'erreur d'estimation étant la même au sein d'une classe). Les erreurs de processus au sein d'une classe étant très proches et leur différence non significative, il n'est pas possible de différencier les modèles par rapport à cette mesure. Nous décidons donc de donner une pondération équiprobable pour les deux classes de bootstrap (cf annexe B.3.2).

Concernant les modèles GLM, la connaissance de la vraisemblance des modèles peut nous permettre de mettre en place une pondération notamment la pondération bayésienne vue précédemment. Cependant, nous ne voulons pas nécessairement effectuer une pondération bayésienne en absence d'éléments de choix de *l'a priori*. Nous décidons de mettre en place la pondération proposée par Burnham et al. [6]. Cette pondération est basée sur les critères d'information AIC (Akaike information criteria) proposés par Akaike [1] :

$$AIC = 2k - 2 \ln(L)$$

avec  $k$  le nombre de paramètres du modèle et  $L$  la vraisemblance.

*Remarque.* Le modèle préféré est celui qui a le plus faible AIC. L'intérêt de ce critère est de pénaliser les modèles ayant trop de paramètres. En effet, l'augmentation du nombre de paramètres améliore nécessairement l'ajustement aux données mais complexifie le modèle.

La pondération proposée par Burnham est la suivante :

$$p_{GLM_i} = \frac{\exp\left(-\frac{1}{2}(AIC_{GLM_i} - \min_{j \in \{1,2\}} AIC_{GLM_j})\right)}{\sum_{k=1}^2 \exp\left(-\frac{1}{2}(AIC_{GLM_k} - \min_{j \in \{1,2\}} AIC_{GLM_j})\right)}, i \in \{1, 2\}$$

Cette pondération correspond à la pondération bayésienne avec un *a priori* de 50% pour chaque modèle. En effet,  $\exp\left(-\frac{1}{2}(AIC_{GLM_i} - \min_{j \in \{1,2\}} AIC_{GLM_j})\right)$  est proportionnelle à la vraisemblance du modèle  $GLM_i$ . Nous pourrions modifier cette pondération en rajoutant un *a priori* (différent de 50%).

TABLE 22 – Critère d’Akaike et pondérations associées

	AIC	pondération
MCMC GLM Poisson	3 385,6	99,96%
MCMC GLM Gamma	3 401,7	0,033%

Cette approche nous semble surpondérer le GLM Poisson sur-dispersé au détriment du GLM Gamma. Elle ne prend en compte que la vraisemblance qui est nettement supérieure pour le modèle de GLM Poisson. Cependant, au vu de la validation des hypothèses, le modèle GLM Gamma semble aussi pertinent et adéquat pour modéliser les provisions.

Nous gardons ces pondérations dans un premier temps puis nous effectuerons une analyse de sensibilité pour chaque classe afin d’observer l’impact d’un changement de pondérations ainsi que les pondérations qui nous semblent les plus appropriées.

Nous présentons les erreurs de modèles pour chaque classe avec l’erreur globale en prenant une pondération équiprobable des classes.

TABLE 23 – Erreur de modèle intra-classe par pondération

		<i>Bootstrap ODP</i>	<i>Bootstrap Mack</i>	<i>MCMC GLM</i>	<b>Totale</b>
<b>Erreur</b>	<b>en %</b>	1.168%	0.034%	3.457%	1.553%
	<b>en valeur</b>	3.9M€	0.1M€	11.8M€	5.3M€

### Tests de sensibilité

Pour chaque classe, nous avons évalué la sensibilité des résultats à la pondération choisie.

Classe 1 (Bootstrap ODP) : Cette classe comporte quatre modèles. En fonction de la pondération attribuée à chaque modèle, l’erreur varie entre 0.669% et 1.611%, la borne supérieure étant obtenue dans le pire cas. Le modèle par défaut dans le logiciel *ResQ* est le modèle avec une distribution prédictive gamma. Ainsi, si nous choisissons de donner une pondération de 40% à ce modèle et de 20% aux autres, nous obtenons une erreur de modèle de 1.068%, ce qui est inférieur au cas équiprobable. Plus le poids donné à ce modèle est élevé plus l’erreur de modèle diminuera, c’est d’ailleurs le modèle le moins prudent des quatre. A contrario, le modèle le plus prudent (pire des cas) étant le modèle avec distribution prédictive normale, un poids fort donné à ce modèle augmentera l’erreur. Cependant, les provisions sont généralement modélisées par des distributions à queues plus lourdes que celles de la loi normale ce qui met en doute la pertinence de ce modèle.

Classe 2 (Bootstrap Mack) : L’erreur est globalement très faible dans cette classe composée de quatre modèles et peut même être négative. L’erreur est comprise dans l’intervalle [-0.301%,0.314%] en fonction des pondérations choisies. Au vu de la petite amplitude de cet intervalle, nous pouvons immédiatement conclure au faible impact de la pondération des modèles de cette classe. En opposition avec la classe précédente, la borne inférieure est obtenue en donnant un poids maximal au modèle avec la loi normale. Une erreur négative signifie que le modèle de référence est plus prudent que le modèle ‘log-normal’ référent. Pour être prudent, nous décidons de ne pas reprendre de capital au titre du risque de modèle dans ce cas mais plutôt de considérer l’erreur comme nulle. L’erreur maximale est obtenue pour la distribution prédictive log-normale. Si, comme dans la classe précédente, nous donnons un poids de 40% à la distribution gamma et 20% aux autres, nous obtenons une erreur encore une fois plus faible que le cas équiprobable de 0%.

Classe 3 (GLM) : L'erreur de modèle pour la classe GLM est comprise dans l'intervalle [3.456%,6.896%], la borne inférieure (supérieure) étant obtenue avec tout le poids mis sur le modèle de Poisson (de Gamma). Quelle que soit la pondération, l'erreur de modèle sera nettement supérieure à celle obtenue à l'intérieur des autres classes. Outre le choix important de la pondération de cette classe, le choix de la pondération des modèles doit être judicieux. Le cas de la pondération équiprobable donne une erreur de 5.176%. Cependant, ce choix ne prend pas en compte l'information sur la vraisemblance des modèles. Bien que la pondération obtenue par le critère d'information nous semble trop favoriser le modèle de Poisson, ce dernier reste plus vraisemblable. Nous décidons donc d'attribuer le poids 2/3 au modèle de Poisson et 1/3 à l'autre, pondération qui nous semble plus représentative de la réalité. Nous obtenons ainsi une erreur de 4.602%.

Pour les modèles de Bootstrap, les cas extrêmes ne sont pas les plus vraisemblables et les poids équiprobables nous semblent cohérents et prudents.

Pour les GLM, l'erreur obtenue avec la pondération (2/3,1/3) est supérieure à l'erreur évaluée initialement mais cette pondération nous semble plus cohérente et nous sommes plus prudent en retenant ces poids.

L'erreur globale en modifiant les pondérations attribuées aux GLM par 2/3 et 1/3 devient de 1.935% soit 6.7M€. Par ailleurs, comme dans l'approche mixte, la modification des pondérations des classes va engendrer des niveaux de provisions différentes.

## 8.4 Conclusion

Nous observons, d'une part, que l'erreur de modèle est non négligeable pour les modèles de provisionnement sur ce segment. Par ailleurs, la variabilité de cette erreur est importante et dépend des pondérations des classes et des modèles intra-classes puisqu'elle varie entre -0.301% et 6.896%.

Cette erreur provient essentiellement de la classe MCMC GLM. Les provisions en *best estimate* obtenues pour les deux types d'erreur sont supérieures à celles des autres modèles de provisionnement. Il en est de même pour la VaR évaluée sur les distributions issues des GLM. De plus, nous observons qu'en fonction du choix des paramètres des modèles GLM (ajout ou non d'effets calendaires, choix du test de significativité des paramètres de ligne et colonne...) les provisions en *best estimate* vont être significativement différentes. Pour ces modèles GLM, le choix des paramètres et de la loi de la famille exponentielle est crucial. Cependant, le manque de données rend ce choix difficile à vérifier. Ces modèles ne sont d'ailleurs, à ce jour, pas couramment utilisés par la SCOR contrairement aux autres modèles de provisionnement étudiés.

Pour les modèles de provisionnement, l'ensemble de modèles est relativement conséquent et nous ne disposons que de peu de données pour tester leur adéquation. C'est pourquoi, nous avons décidé de garder des modèles dont toutes les hypothèses n'étaient pas vérifiées. L'avantage d'avoir beaucoup de modèles est d'assurer avec une plus grande probabilité que le 'vrai' modèle appartient à l'ensemble. D'un autre côté, si des modèles non pertinents et très prudents sont inclus dans l'ensemble, cela peut augmenter considérablement l'erreur de modèle, au moins dans l'approche *worst case* (cela pourrait être le cas des GLM). Dans l'approche par pondération, l'erreur de modèle peut aussi augmenter si les pondérations ne sont pas bien choisies. Pour les mêmes raisons de manque de données et donc de manque de résultats d'adéquation des modèles, des avis d'experts sont indispensables pour pondérer les différents modèles.

Etant dans l'incapacité de juger la pertinence des modèles MCMC GLM et au vu de l'impact qu'ils ont sur l'erreur, nous pourrions envisager de les retirer de l'ensemble de modèles. Une autre possibilité serait de travailler sur des données ayant une granularité plus fine, par exemple, les données par sinistres.

## 9 Application à des structures de dépendance

Dans cette section, nous allons évaluer l'erreur sur des modèles de dépendance de deux segments non-vie. Nous reprenons donc les mêmes modèles stochastiques de provisionnement que dans l'étude précédente (cf section 8). Nous agrégeons deux distributions de provisions estimées totales obtenues sur les triangles de payés des segments non-vie transport maritime, que nous avons étudié précédemment dans le cas d'une variable seule, et automobile. Les principaux risques couverts par l'assurance auto sont la responsabilité civile (RC), les dommages corporels et matériels. Pour le segment automobile, nous disposons, comme pour le segment transport maritime, du triangle des payés des sinistres survenus entre 1999 et 2013. Le développement est plutôt long (15 ans) du fait des risques à long terme couverts (RC et dommages corporels) qui peuvent donner lieu au versement de rentes.

Pour ce triangle de paiements, la méthode des GLM avec lien 'log' et erreur gamma ne semble pas adaptée aux données. C'est pourquoi, nous justifierons, par un des tests de validation du modèle, notre choix de le retirer de l'ensemble des modèles de provisionnement pour l'étude du segment automobile.

Par la suite, nous discuterons de la calibration des paramètres de la copule avant de mettre en œuvre les méthodes d'évaluation de l'erreur de dépendance pure d'une part et de dépendance globale d'autre part.

### 9.1 La non adéquation du modèle GLM avec erreur gamma

Comme nous l'avons mentionné, le modèle GLM avec lien 'log' et erreur gamma n'est pas adéquat pour les données du segment automobile. Nous avons pu constater que les provisions estimées en *best estimate* étaient beaucoup plus élevées de même que la volatilité de la distribution. Un des tests non validés pour ce modèle est le graphe confrontant les données observées et les données estimées par les GLM Poisson et Gamma.

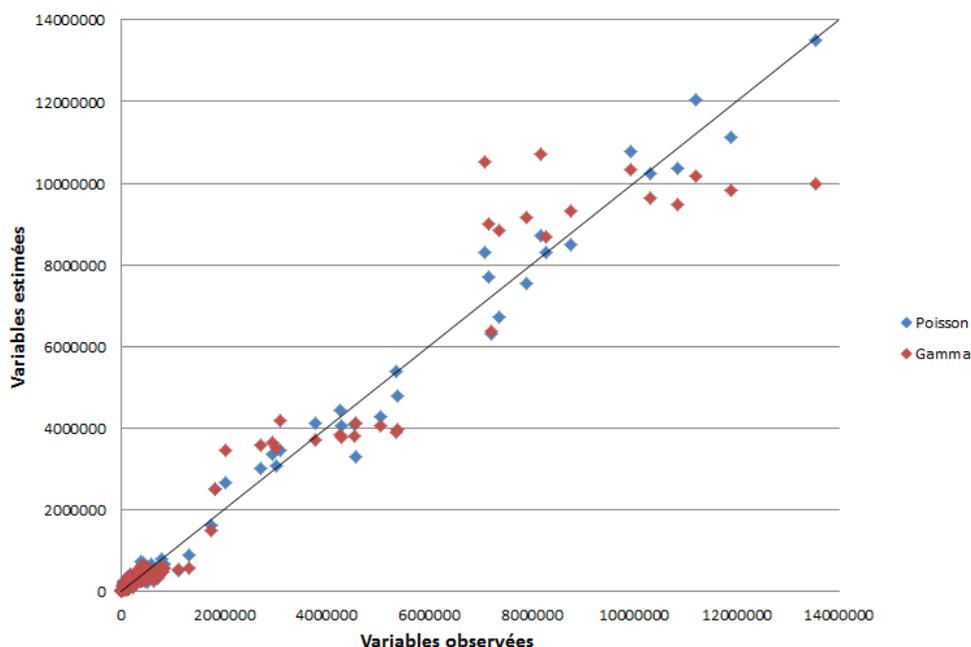


FIGURE 25 – Variables observées vs estimées pour les GLM

Comme nous l'avons vu dans l'étude du segment transport maritime, les variables estimées devaient être les plus proches possibles des variables observées pour que le modèle soit adéquat ; ainsi les points du graphe devraient être sur la première bissectrice (la droite noire). Les points bleus (issus du GLM Poisson) sont très proches de la bissectrice à une ou deux exceptions près qui ne s'en éloignent pas trop non plus. Les points rouges (issus du GLM Gamma) sont, quant à eux, plus dispersés et certains points très éloignés de la bissectrice, notamment pour les grandes sinistralités. Cela

permet de justifier de la mauvaise adéquation du modèle aux données et ainsi de la difficulté d'avoir des estimations des provisions fiables avec ce modèle.

Nous en déduisons donc que les variables expliquées (les paiements incrémentaux) ne suivent pas une loi gamma et choisissons de ne pas retenir ce modèle pour le provisionnement du segment automobile.

## 9.2 Choix de la copule et ses paramètres

### 9.2.1 Choix de la copule

Comme pour la dépendance d'actifs, la dépendance sera modélisée par des copules. L'utilisation des copules pour représenter la dépendance de segments non-vie est nécessaire puisque cette dernière n'est dans ce cas pas linéaire ; la copule gaussienne n'est, *a priori*, pas la mieux adaptée à ces risques. Sans tenir compte de nos *a priori*, nous avons sélectionné les cinq copules les plus utilisées, la copule gaussienne, la copule de Student, la copule de Clayton inversée, la copule de Gumbel et la copule de Frank. Ainsi, en l'absence d'idées préalables sur la forme de la dépendance, ces 5 copules nous permettent d'être en mesure de modéliser quasiment tous les types de dépendance en dimension 2. La copule de Clayton inversée correspond à la copule de Clayton pivotée de 180 degrés. Plus formellement, la copule de Clayton inversée appelée plus communément la copule de survie de Clayton est définie par :

$$\tilde{C}_\theta(u, v) = u + v - 1 + C_\theta(1 - u, 1 - v)$$

### 9.2.2 Choix des paramètres de copule

La calibration des paramètres de ces copules est plus délicate pour les modèles de provisionnement car nous disposons de très peu de données. En effet, la corrélation est évaluée sur les montants ultimes estimés de chaque année d'origine, soit la corrélation entre deux distributions de 15 points dans cette application (15 années d'origine). Le graphe 9.2.2 des rangs montre la difficulté, d'un point de vue graphique, d'émettre une conjoncture quant à la dépendance entre ces deux segments et à la copule modélisant cette dépendance. Cependant, nous avons *a priori* que la dépendance entre les deux segments sera faible de part la nature de leurs risques.

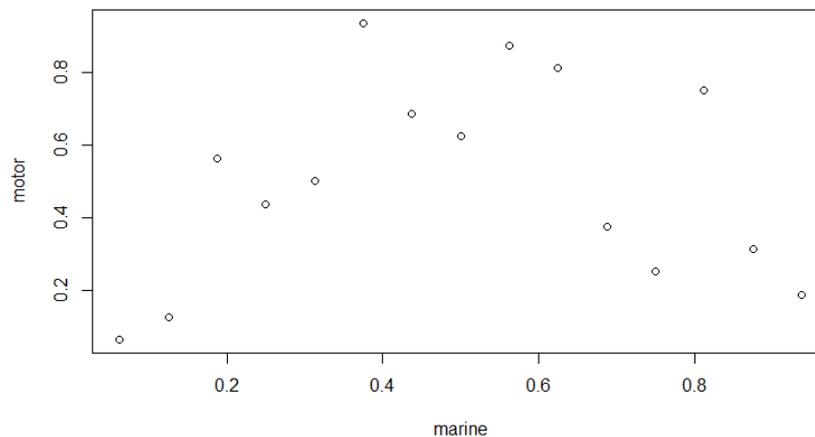


FIGURE 26 – Nuage de points des rangs

Nous avons retenu différentes méthodes, certaines quantitatives avec les problèmes sous-jacents liés au manque de données et d'autres issues des calibrations utilisées pour modéliser les dépendances dans le modèle interne de la SCOR. Comme il est difficile de juger de l'adéquation des copules estimées à nos données, nous nous intéressons aux dépendances 'extrêmes', à savoir l'absence de dépendance et la dépendance parfaite entre les deux segments. En effet, ces cas extrêmes nous permettent d'observer l'impact, sur le capital réglementaire, d'une potentielle sous-estimation ou surestimation de la dépendance due à un choix inadapté du paramètre de la copule.

Les méthodes présentées ci-dessous ont été implémentées avec des marges simulées à partir de chaque modèle de provisionnement. Cependant, comme les résultats donnaient des conclusions similaires pour les différentes marges, nous ne présentons, dans cette section, que les résultats évalués en utilisant le modèle 'log-normal' pour simuler les marges (rappel : ce modèle était le modèle de référence en provisionnement).

## CQEP

Tout d'abord, nous ne pouvons pas évaluer cette probabilité conditionnelle (la CQEP) dans ce cas, toujours à cause du manque de données. Nous avons donc retenu la valeur utilisée en interne pour la calibration de la dépendance entre ces deux segments qui est un CQEP de 0.10 pour un niveau de quantile  $\alpha$  de  $\frac{1}{100}$ . Cette valeur, indiquant une corrélation plutôt faible entre les deux segments, est obtenue par une méthode bayésienne élaborée dans le cadre de la mise en place du modèle interne (cf papier SCOR [2]). Cette méthode est basée sur un *a priori* et des avis d'experts reposant sur l'analyse des facteurs de risques qui impactent conjointement les segments étudiés. Elle pourrait incorporer les données historiques de sinistralité comme troisième source d'information mais, du fait du manque de données, le modèle actuel n'en prend, à ce jour, pas compte.

A partir du CQEP et du niveau du quantile, nous retrouvons aisément le paramètre de la copule à l'aide des formules définies dans la section 4.2. Nous présentons dans le tableau 24 ci-dessous les différents paramètres des copules ainsi que les quantiles à 99.5% des distributions agrégées. Les distributions agrégées sont obtenues en simulant les marges à partir du modèle 'log-normal' et en inversant empiriquement la copule selon la méthode décrite dans la section 4.3.

*Remarque.* Dans toute la suite de l'étude, la moyenne et l'écart-type des VaR sont obtenus sur 100 simulations de distributions de provisions finales estimées de taille 10 000. Si la différence entre les moyennes est inférieure à deux fois l'écart-type, les écarts entre les moyennes peuvent être associés uniquement aux erreurs de simulation. C'est le cas des résultats de cette section (cf tableaux ci-dessous) et cela laisse supposer, d'ores et déjà, que l'erreur de modèle sera faible.

TABLE 24 – Paramètres et quantiles à 99.5% par la méthode CQEP

Copule	Paramètre	Moyenne VaR	Ecart-type	CoV
<i>gaussienne</i>	0,435	355 483 398	858 082	0,241%
<i>Student à 5 degrés de liberté</i>	0,103	354 647 360	916 932	0,259%
<i>Clayton inversée</i>	0,209	355 072 863	917 721	0,258%
<i>Gumbel</i>	1,072	354 842 875	760 605	0,214%
<i>Frank</i>	11,100	356 135 513	854 399	0,240%

## $\rho$ de Spearman

Le  $\rho$  de Spearman (ou la corrélation des rangs) est défini comme le coefficient de corrélation linéaire de Pearson évalué sur les rangs des variables, empiriquement cela se traduit par :

$$\rho_n = \frac{\sum_{i=1}^n (r_i^X - \bar{r}^X)(r_i^Y - \bar{r}^Y)}{\sqrt{\sum_{i=1}^n (r_i^X - \bar{r}^X)^2 \sum_{i=1}^n (r_i^Y - \bar{r}^Y)^2}}$$

avec  $X$  la première covariable (ici marine),  $Y$  la deuxième covariable (automobile),  $n$  la taille de l'échantillon (ici 15),  $r$  le vecteur des rangs et  $\bar{r}$  la moyenne empirique des rangs.

*Remarque.* D'après la loi des grands nombres, pour  $n$  grand, cette quantité converge vers la corrélation théorique des rangs  $\rho = \mathbb{E}[(R^X - \mathbb{E}(R^X))(R^Y - \mathbb{E}(R^Y))]$ . Or, les données sont insuffisantes ici pour avoir une telle propriété.

Plutôt que d'évaluer s'il existe une relation linéaire entre les deux variables, la corrélation de Spearman s'intéresse à l'existence d'une relation monotone entre les deux variables. Le  $\rho$  de Spearman varie entre -1 et 1 comme la corrélation linéaire, une corrélation négative (positive) indique une tendance décroissante (croissante) entre les deux variables. Cette corrélation évalue donc la quantité d'information que les deux covariables ont en commun sans faire d'hypothèses sur la distribution jointe.

Dans le cas de notre étude, nous calculons cette quantité sur les vecteurs des paiements ultimes estimés pour chaque année d'origine. Ainsi, nous cherchons à savoir si le comportement des ultimes est le même d'un segment à l'autre ; c-à-d. si, lorsque les ultimes baissent (augmentent) entre deux années d'origine sur un segment, les ultimes de l'autre segment subissent aussi une baisse (augmentation). Nous ne nous intéressons pas vraiment aux ordres de grandeur des variations (qui peuvent manquer de stabilité du fait des estimations) mais plutôt à l'année d'occurrence. C'est pourquoi, nous préférons la corrélation de Spearman à la corrélation linéaire.

Le coefficient de corrélation obtenu est :  $\rho_n = 0.1$ . Ainsi, la corrélation entre nos deux vecteurs est faible ce que nous avons déjà observé avec le CQEP.

L'inversion du  $\rho$  de Spearman, afin d'obtenir le paramètre de chaque copule, s'effectue par la méthode des moments. Cependant, pour une copule donnée, la corrélation des rangs théoriques des covariables n'est pas toujours explicite et requière souvent l'utilisation de méthodes d'approximation numériques (réalisées par la fonction *iRho* sur R). C'est le cas pour toutes les copules étudiées excepté la copule gaussienne dont la relation est la suivante :

$$\rho_n = \sin\left(\frac{\theta\pi}{2}\right)$$

avec  $\theta$  le paramètre de la coupule gaussienne tel que  $\Sigma = \begin{pmatrix} 1 & \theta \\ \theta & 1 \end{pmatrix}$

Le tableau ci-dessous répertorie les paramètres obtenus pour chaque copule à l'aide de cette méthode ainsi que le capital réglementaire requis en modélisant les marges par le modèle 'log-normal'.

TABLE 25 – Paramètres et quantiles à 99.5% par la méthode du  $\rho$  de Spearman

Copule	Paramètre	Moyenne VaR	Ecart-type	CoV
<i>gaussienne</i>	0,105	354 204 381	816 154	0,230%
<i>Student à 5 degrés de liberté</i>	0,105	354 645 104	876 384	0,247%
<i>Clayton inversée</i>	0,143	354 807 713	833 223	0,235%
<i>Gumbel</i>	1,073	354 620 756	880 694	0,248%
<i>Frank</i>	0,603	354 103 831	872 669	0,246%

## Dépendances extrêmes

Nous étudions deux cas extrêmes de dépendance, le cas où il n'y a pas de dépendance et le cas où la dépendance est de 100%, c-à-d. qu'il existe une relation parfaite entre les deux covariables. Nous n'étudierons pas la dépendance négative qui ne semble pas appropriée à notre étude. L'intérêt de ces deux cas est d'évaluer le risque maximal potentiel engendré par une mauvaise paramétrisation des copules. C'est particulièrement important dans cette étude où nous ne sommes pas en mesure d'émettre d'avis quant à la validité des méthodes de calibration présentées précédemment du fait du manque de données.

Les tableaux 26 et 27 récapitulent les mesures de risque ainsi que les erreurs de simulation dans le cas des dépendances extrêmes pour chaque copule.

TABLE 26 – Paramètres des copules et quantiles à 99.5% en l'absence de dépendance

Copule	Paramètre	Moyenne VaR	Ecart-type	CoV
<i>gaussienne</i>	0	353 757 223	842 216	0,238%
<i>Student à 5 degrés de liberté</i>	0	354 037 615	923 832	0,261%
<i>Clayton inversée</i>	1E-10	353 778 826	923 318	0,261%
<i>Gumbel</i>	1	354 125 795	764 661	0,216%
<i>Frank</i>	0	353 630 738	895 973	0,253%

TABLE 27 – Paramètres des copules et quantiles à 99.5% pour une dépendance parfaite

Copule	Paramètre	Moyenne VaR	Ecart-type	CoV
<i>gaussienne</i>	1	357 932 600	1 120 546	0,313%
<i>Student à 5 degrés de liberté</i>	1	357 902 512	1 149 434	0,321%
<i>Clayton inversée</i>	100	357 885 528	1 234 792	0,345%
<i>Gumbel</i>	100	357 918 854	1 244 034	0,348%
<i>Frank</i>	100	357 454 268	1 102 335	0,308%

Nous observons, d'une part, que, à niveau de dépendance fixé, les montants de capital risque sont assez proches d'une copule à une autre. Par ailleurs, la potentielle surestimation ou sous-estimation du capital réglementaire si nous considérons, à tort, qu'il y a ou qu'il n'y a pas de dépendance, n'est pas significative. En effet, en fonction du niveau de dépendance, le capital réglementaire variera d'environ 4M, soit moins de 1% du capital risque. Ainsi cette erreur est négligeable dans notre cas. Cela s'explique en partie par le poids très fort d'une covariable par rapport à l'autre dans l'évaluation du capital risque à 99.5%. La variation du quantile entre plusieurs niveaux de dépendance sur d'autres segments non-vie ayant des montants de provisions comparables pourrait être bien plus importante. Cet impact est donc à vérifier avant de retenir une méthode de calibration, particulièrement en l'absence de critères de décision dû au manque de données.

### Méthode retenue

Nous avons décidé de retenir la méthode du CQEP. Ce choix s'explique par divers critères.

- Très peu de méthodes ont pu être mises en place à cause du peu de données. Les méthodes d'adéquation par le maximum de vraisemblance ou le Tau de Kendall n'ont pas donné de résultats significatifs et n'ont ainsi pu être retenues.
- La méthode du  $\rho$  de Spearman présente généralement de moins bon résultats que les autres méthodes quantitatives en terme d'erreur quadratique moyenne (cf Kojadinovic et al. [26]).
- Dans notre application, les quantiles à 99.5% sont plus prudents par la méthode du CQEP que par la méthode du  $\rho$  de Spearman qui estime une dépendance plus faible.
- Les dépendances retenues pour le modèle interne prennent en compte plusieurs sources d'information et font l'objet d'une double validation (en interne et par le régulateur).

### 9.3 Erreur de dépendance pure

Nous commençons par évaluer l'erreur de dépendance pure liée uniquement au choix de la copule. Nous évaluons cette erreur en choisissant le modèle 'log-normal' pour modéliser les marges. Les résultats obtenus pour des marges simulées à partir des autres modèles de provisionnement et non présentés sont très similaires.

La copule référente dans cette étude est la copule de Clayton inversée, copule ayant une dépendance de queue inférieure nulle et une dépendance de queue supérieure positive (inverse de la Clayton). Cette copule est prudente, dans le sens où elle prend bien en compte la dépendance de queue supérieure, c-à-d. la dépendance entre les événements extrêmes. Par ailleurs, les études sur les dépendances en

provisionnement non-vie ont montré que cette copule était la plus adaptée pour agréger ces segments, c'est d'ailleurs celle qui a été retenue pour l'agrégation des risques non-vie.

L'erreur de modèle a été évaluée une fois de plus par les deux approches, l'approche *worst case* et l'approche par pondération. Nous n'utilisons pas une pondération bayésienne dans la mesure où nous ne disposons pas d'information sur l'adéquation de chacune des copules mais seulement d'un *a priori*. Nous traiterons de deux pondérations, la pondération équiprobable et la pondération mettant la moitié du poids sur la copule de Clayton inversée et le reste de façon équiprobable entre les quatre autres copules.

TABLE 28 – Deux pondérations des copules

	<i>gaussienne</i>	<i>Student</i>	<i>Clayton inversée</i>	<i>Gumbel</i>	<i>Frank</i>
<i>poids 1</i>	20%	20%	20%	20%	20%
<i>poids 2</i>	12.5%	12.5%	50%	12.5%	12.5%

Pour l'approche par pondération, la méthodologie consiste à réaliser, comme auparavant, des simulations par bootstrap entre les différentes copules pour obtenir la distribution jointe puis la distribution des provisions.

Le tableau ci-dessous présente les statistiques des quantiles à 99.5% issues du bootstrap entre les cinq copules.

TABLE 29 – VaR à 99.5% pour l'approche par pondération

<b>VaR 99,5%</b>	<b>Moyenne</b>	<b>Ecart-type</b>	<b>CoV</b>
<i>poids 1</i>	355 372 874	850 292	0.239%
<i>poids 2</i>	355 187 736	853 019	0.240%

La VaR moyenne obtenue avec les deux poids est considérée comme identique au vu de l'écart-type. Ainsi, donner un poids fort à la copule de Clayton inversée ne semble pas avoir d'impact ce qui est justifié par le fait que la mesure de risque de cette copule est quasiment égale à la moyenne des mesures de risque de chaque copule. Pour finir, les erreurs de dépendance sont présentées ci-dessous. L'erreur dans le pire des cas est évaluée à partir des résultats du tableau 24 ci-dessus.

TABLE 30 – Erreur de dépendance pure par les deux approches

	<b>Référence</b>	355 072 863
<b>Erreur</b>	<i>AM</i>	0.299%
	<i>Bootstrap poids 1</i>	0.084%
	<i>Bootstrap poids 2</i>	0.032%

Nous observons que l'erreur est quasiment nulle quelle que soit l'approche ou la pondération choisie. Ainsi dans notre étude, le choix de la copule n'a que très peu d'impact.

## 9.4 Erreur de dépendance globale

Dans cette section, nous évaluons l'erreur de dépendance globale; c-à-d. l'erreur de spécification de la copule et l'erreur du choix du modèle de provisionnement permettant de simuler les marges. Le modèle de référence est ici le modèle 'log-normal' pour simuler les marges et la copule de Clayton inversée pour modéliser la dépendance. Le quantile de référence est donc le même que pour l'erreur de dépendance pure : 355.072M.

## Approche *worst case*

Pour l'approche par le pire des cas, nous avons évalué les quantiles à 99.5% sur les distributions jointes des provisions obtenues pour chaque copule et chaque modèle de marges. Ces valeurs sont présentées en annexe B.4.

Nous commençons par évaluer l'erreur *worst case* en valeur absolue pour chaque classe de la même manière que pour un risque seul. Puis, nous évaluerons l'erreur globale après pondération équiprobable des classes.

TABLE 31 – Erreur de modèle globale par le pire des cas pour chaque classe

	Bootstrap ODP	Bootstrap Mack	GLM MCMC
AM	1.597%	0.414%	6.774%

Comme dans l'évaluation de l'erreur pour un risque individuel, l'erreur est élevée pour la classe GLM. Par ailleurs, les ordres de grandeur des erreurs sont très proches de ceux d'un risque non-vie vu individuellement. La copule ne semble pas rajouter d'erreur ce qui ne fait que confirmer la partie précédente sur la dépendance pure. Cela s'explique aussi par la prédominance des provisions pour le segment marine par rapport au segment auto. Ainsi l'impact de la branche auto est moindre ainsi que celui de la dépendance, d'autant que celle-ci est faible entre les deux risques. Des résultats différents seraient observés sur des risques ayant une plus forte dépendance et dont les provisions estimées seraient du même ordre de grandeur.

En utilisant la pondération équiprobable des classes, nous obtenons une erreur en *worst case* de 2.93%. Cette erreur varie entre 0.414% et 6.774% en fonction des pondérations des classes choisies. Etant donné que le modèle GLM gamma n'a pas été retenu pour estimer les provisions pour le segment auto, nous décidons de donner un poids de 20% à cette classe et de 40% aux deux autres. L'erreur est alors diminuée de 0.8 points de pourcentage.

## Approche par pondération

Pour évaluer l'erreur de modèle globale par pondération, nous assignons un poids à chaque marge et à chaque copule. Nous simulons par bootstrap les distributions des provisions pour chaque segment en prenant en compte le poids de la classe et le poids de chaque modèle intra-classe simultanément. Nous simulons par ailleurs une distribution jointe par bootstrap parmi les cinq copules. Puis, nous utilisons les vecteurs ordonnés pour évaluer les marges de la distribution jointe en inversant empiriquement les fonctions de répartition (cf section 4.3). Pour obtenir la distribution des provisions totales estimées, il ne reste qu'à sommer les deux marginales. Le capital risque à 99.5% est évalué sur cette distribution.

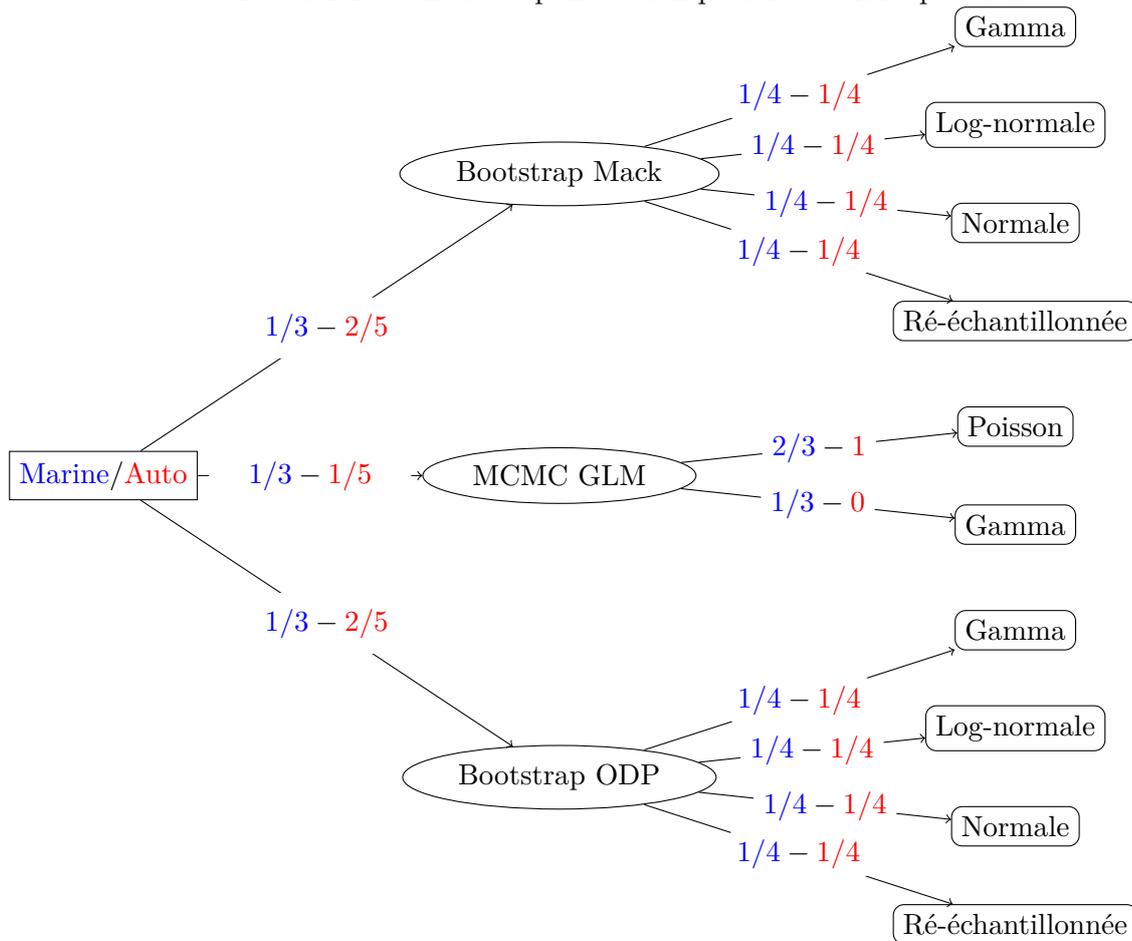
Pour la simulation des marges, nous avons retenu les pondérations qui nous semblaient les plus appropriées, répertoriées dans l'arbre de pondération 27 ci-dessous avec en bleu les pondérations pour le segment marine et en rouge les pondérations du segment auto.

Concernant, les copules, nous avons effectué les simulations avec les deux poids sélectionnés dans la section précédente (cf table 28). Ainsi le tableau 32 donne la valeur moyenne du quantile à 99.5%, obtenue sur 100 itérations de distributions agrégées des provisions de taille 10 000 et obtenues après pondérations des copules et des marges, et l'erreur de modèle associée.

TABLE 32 – VaR à 99.5% pour l'approche par pondération

VaR 99,5%	Moyenne	Erreur
<i>poids 1</i>	366 726 214	3.282 %
<i>poids 2</i>	366 707 989	3.277%

FIGURE 27 – Arbre de pondération pour les deux risques non-vie



Contrairement au cas de l'approche *worst case*, les pondérations des modèles intra-classes et extra-classes sont intégrées dès la simulation des marges et les pondérations extra-classes sont différentes entre les deux segments. Cette dernière subtilité n'était pas possible dans l'approche *worst case* puisque nous prenions en compte le poids des classes qu'après obtention de la distribution agrégée des provisions. Ainsi, l'approche par pondération permet de gagner en précision mais cela rend l'erreur associée plus élevée que celle par le pire des cas après pondération des classes.

## 9.5 Conclusion

L'erreur de modèle globale sur nos deux segments est, comme pour l'étude d'un risque isolé, assez élevée quelle que soit l'approche d'évaluation. En effet, dans le cas de l'approche *worst case*, l'erreur est de 2.9% tandis qu'elle est de 3.3% dans l'approche par pondération.

D'une part, il est important de souligner que ces deux erreurs ne sont pas directement comparables puisque les poids utilisés pour pondérer les classes dans l'approche par pondération sont plus fins que dans l'approche *worst case*. En effet, si les erreurs étaient obtenues avec la même pondération des classes, l'erreur dans l'approche par pondération serait, par définition, plus faible que dans l'approche *worst case*.

D'autre part, l'erreur est relativement proche de celle obtenue pour un risque isolé. L'erreur de modèle provient effectivement essentiellement de la modélisation des marges et non de celle des copules. Comme nous l'avons mentionné dans l'étude d'un risque seul, les modèles GLM sont la source principale de cette erreur.

Au vu de ce dernier constat, les pistes d'amélioration pour l'évaluation du risque de modèle sur des risques non-vie agrégés sont donc sensiblement les mêmes que celles soumises dans l'étude des risques non-vie vus individuellement.

## 10 Analyse et interprétations des résultats sur les modèles de provisionnement

Dans cette partie, l'erreur de modèle n'est pas nulle sur les segments étudiés. Nous observons, par ailleurs, une plus grande disparité des value-at-risk obtenues entre les différents modèles de provisionnement ; cela explique les variations significatives de l'erreur selon les pondérations retenues pour les marges.

L'erreur provient, par ailleurs, essentiellement des marges et non de la modélisation de la dépendance, le choix de la pondération des copules n'ayant que peu d'impact sur l'erreur globale de dépendance. Comme nous l'avons mentionné, le modèle de marges qui augmente de façon importante l'erreur est le modèle MCMC GLM.

La principale difficulté rencontrée sur les modèles de provisionnement non-vie est le manque de données. En effet, les sinistres sont étudiés après agrégation par année de survenance. Les données sont particulièrement manquantes sur les deux dernières années d'origine. La vérification des hypothèses des modèles et de leur adéquation aux données s'avère donc dure à évaluer en pratique. C'est pourquoi, nous n'avons retenu que des modèles couramment utilisés pour ce type de risques et ayant été approuvés par des jugements d'experts. En réassurance, les MCMC GLM sont actuellement peu appliqués du fait de la nécessité d'avoir une quantité importante de données.

Nous retiendrons principalement de cette étude sur deux risques non-vie particuliers que le choix de l'ensemble de modèle pour les marges est très important. Pour améliorer cette étude, nous pourrions évaluer l'erreur en enlevant les GLM qui présentent des résultats très différents des autres modèles sans que nous ne soyons en mesure d'évaluer leur pertinence. L'autre solution serait de travailler sur des données par sinistres pour pouvoir améliorer l'analyse de l'adéquation des modèles.

Une autre considération qui concerne indirectement le choix de l'ensemble de modèle est le choix des pondérations. En effet, plus le nombre de modèles sélectionnés dans l'ensemble est élevé, plus la probabilité que cet ensemble de modèle contienne le 'vrai' modèle augmente. Cependant, l'augmentation de la taille de l'ensemble rend aussi plus probable la possibilité d'avoir des modèles non pertinents. Dans ce cas, l'erreur de modèle pourra être potentiellement élevée pour l'approche *worst case*. Pour l'approche par pondération, l'erreur peut être bien évaluée dans la mesure où les poids des modèles choisis dépendent bien de l'adéquation des modèles. Cependant, comme la pertinence des modèles n'est pas quantifiable par des outils statistiques dans ces études à cause de l'absence de données suffisantes, l'intégration d'avis d'experts pour le choix des pondérations permettrait de se rapprocher de la 'vraie' erreur de modèle et non de celle biaisée par des modèles inadaptés.

Pour finir, il serait intéressant d'élargir notre étude à d'autres segments et généraliser l'approche afin de pouvoir évaluer le risque de modèle sur un portefeuille complet.

# Conclusion

A ce jour, il n'existe pas de consensus sur la définition du risque de modèle et des différentes sources d'erreur associées. De même, aucun standard ne répertorie les bonnes pratiques pour évaluer ce risque. Cette étude est centrée sur l'évaluation de l'erreur de spécification du modèle, qui est une des sources d'erreur de modèle. Dans la littérature, deux approches ont notamment été proposées pour mesurer ce risque. La première approche, l'approche dite *worst case*, évalue l'erreur pas rapport au pire des cas. La deuxième approche, l'approche par pondération, évalue l'erreur en pondérant chacun des modèles en fonction de leur adéquation. Une première étude avait été menée à la SCOR pour mettre en application la première approche sur des risques isolés. Ainsi, le premier objectif de cette étude était de reprendre cette étude et de l'améliorer en mettant en application la deuxième approche, qui avait déjà été implémentée dans un autre mémoire d'actuariat. Le deuxième objectif était l'évaluation de l'erreur de modèle sur un portefeuille de risques en intégrant la dépendance sous-jacente. Etant donné la nouveauté que représentait cette partie, nous nous sommes focalisés sur l'agrégation de deux risques.

La partie applicative de l'étude s'est portée sur deux risques de modèle, le risque de modélisation d'actifs et le risque de modèles de provisionnement non-vie. Nous avons ensuite agrégé deux actifs ainsi que deux risques non-vie afin d'évaluer l'erreur de modèles de dépendance. L'agrégation a été modélisée par des copules afin de prendre en compte les dépendances non linéaires. L'erreur de dépendance correspond, dans notre étude, à l'erreur de spécification de la copule. L'erreur globale de modèle évalue donc simultanément l'erreur de dépendance et l'erreur de marges.

Sur les actifs étudiés, l'erreur de modèle est non significative pour l'approche bayésienne que ce soit pour les modèle d'actifs isolés ou sur les modèles globaux avec prise en compte de la dépendance. Pour l'approche *worst case*, l'erreur de modèle était aussi nulle dans la mesure où le modèle référent était le plus prudent. Par ailleurs, l'implémentation des méthodes a été réalisée à la fois sur les rendements logarithmiques mensualisés et sur les rendements logarithmiques journaliers des actifs. Bien que l'erreur de modèle était très similaire pour les deux fréquences, la modélisation des rendements mensuels est préférée du fait de sa simplicité et sa stabilité.

Sur les lignes d'affaire non-vie sélectionnés, l'erreur pour un risque seul ou agrégé était plus élevée. Comme nous l'avons mentionné, cette erreur est issue des modèles utilisés pour les marges et des modèles GLM en particulier. En effet, les données utilisées pour le provisionnement sont des montants agrégés de sinistres payés (dans cette étude) ou encourus, ce qui limite énormément leur nombre. Ainsi, les résultats statistiques basés sur la théorie des grands nombres ne sont pas viables ici et la validation des modèles utilisés est plus délicate. Une piste d'amélioration de ces résultats serait d'exclure de l'analyse les modèles dont l'adéquation est mise en doute ou encore d'appliquer l'approche sur des données par sinistre afin d'avoir une granularité plus fine et d'être en mesure d'évaluer la pertinence des modèles.

Nous avons vu, dans chaque partie applicative, de manière plus détaillée les limites des modèles retenus. Nous nous concentrons ici sur les limites des deux approches d'évaluation.

L'approche *worst case* est très prudente et peut nécessiter un ajout de capital important pour prendre en compte une potentielle erreur de modèle. L'approche par pondération évalue de manière plus précise l'erreur en tenant compte de l'adéquation des modèles. Cependant, le choix de la pondération peut s'avérer compliqué. Nous avons présenté, en particulier, l'approche par pondération bayésienne qui n'est applicable qu'en connaissance des vraisemblances des modèles et demande le choix d'un poids *a priori*. En l'absence de telles informations, les pondérations sont bien souvent obtenues à partir d'avis d'experts.

Les deux approches présentent des limites similaires à cause de leurs hypothèses fortes communes. Une première hypothèse suppose que le 'vrai' modèle appartient à l'ensemble de modèle sélectionné.

Or, cette hypothèse théorique ne peut pas être vérifiée en pratique d'une manière précise. Le jugement d'expert de l'actuaire prend donc une place importante dans ce cas, permettant de faire une évaluation qualitative de l'adéquation d'un modèle basée sur la connaissance et l'expertise du risque analysé ou du portefeuille. Nous pouvons donc supposer que l'ensemble des modèles retenus contient très probablement le 'vrai' modèle. Une deuxième hypothèse concerne le choix du modèle de référence puisque les deux approches évaluent l'erreur par rapport à une référence. Il est donc nécessaire de choisir un modèle pertinent au préalable. Le modèle généralement retenu est celui qui a fait l'objet de nombreuses validations et qui donne un bon compromis entre adéquation et simplicité d'implémentation. Pour finir, la mise en place de ces méthodes s'avère coûteuse en temps puisqu'elle nécessite l'implémentation et la validation de chaque modèle avant l'évaluation de l'erreur.

Au vu de ces constatations, plusieurs pistes d'amélioration de l'approche sont envisageables. D'une part, notre étude se focalise sur le risque de spécification et non sur les autres risques inclus dans le risque de modèle, notamment le risque d'estimation. L'idée serait d'améliorer la prise en compte de l'erreur en évaluant simultanément ces deux erreurs pour la dépendance. La SCOR a, en particulier, développé un modèle appelé Probex (cf article SCOR [2]) qui permet d'intégrer l'erreur d'estimation lors de la calibration des copules. D'autre part, afin de pouvoir généraliser notre étude et inclure plusieurs risques par exemple, il serait nécessaire d'optimiser le temps de calcul. Pour cela, nous pourrions envisager la mise en œuvre de méthodes numériques permettant, par exemple, d'approximer les intégrales.

Finalement, dès lors que les différentes sources d'erreur de modèle sont intégrées dans la mesure du risque et que le temps de calcul est optimisé, l'approche pourrait être généralisée pour pouvoir calculer l'erreur de modèle sur un portefeuille entier.

# Bibliographie

## Références

- [1] H. Akaike. Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, pages 267–281, 1973.
- [2] P. Arbenz and D. Canestraro. Probex : A new method for the calibration of copula parameters from prior information, observations and expert opinions. *SCOR Papers*.
- [3] P. Barrieu and G. Scandolo. Assessing financial model risk. *European Journal of Operational Research*, 2013.
- [4] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81 :637–654, 1973.
- [5] T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31 :307–327, 1986.
- [6] K.P. Burnham and D.R. Anderson. *Model selection and multimodel inference - A practical information-theoretic approach*. Springer, 2002.
- [7] A.J.G. Cairns. A discussion of parameter and model uncertainty in insurance. *Insurance : Mathematics and Economics*, pages 313–330, 2000.
- [8] A. Charpentier. Chapitre 3 : Mesures de risque. *Mathématique de l'assurance non-vie*, 2, 2005.
- [9] R. Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3) :519–547, 2006.
- [10] M. Crouhy, D. Galai, and R. Mark. Model risk. *Journal of Financial Engineering*, 7 :267–288, 1998.
- [11] Comité de Bâle sur le contrôle bancaire. Convergence internationale de la mesure et des normes de fonds propres. *Banque des règlements internationaux*, 2004.
- [12] E. Derman. Model risk. *Quantitative Strategies Research Notes*, Goldman Sachs, 1996.
- [13] F.X. Diebold, T.A. Gunther, and A.S. Tay. Evaluating density forecasts with applications to financial risk management. *International Economic Review*, 39(4) :863–883, 1998.
- [14] D. Draper. Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society*, 1994.
- [15] S. Dussaule. L'erreur de modèle dans le dispositif solvabilité ii : problématiques et méthodologies. *Mémoire d'actuariat présenté à l'IA.*, 2014.
- [16] P.D. England and R.J. Verrall. Stochastic claims reserving in general insurance. *Institute of Actuaries and Faculty of Actuaries*, 2002.
- [17] P.D. England and R.J. Verrall. Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1(2) :221–270, 2006.
- [18] Commission Européenne. Quantitative impact study 5 - technical specifications. *EIOPA*, 210.
- [19] R.A. Fisher. Proc. camb. philos. soc. 26 :528–535, 1930.
- [20] C. Genest, K. Ghoudi, and L-P. Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3) :543–552, 1995.
- [21] C. Genest and J. Mackay. Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *Canadian Journal of Statistics*, 14(2) :145–159, 1986.
- [22] C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas : a review and a power study. *Insurance : Mathematics and Economics*.
- [23] M.R. Hardy. A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2) :41–176, 2001.
- [24] K. Itô. Stochastic integral. *Proc. Imperial Acad. Tokyo*, 20 :519–524, 1944.

- [25] J. Kerkhof, B. Melenberg, and H. Schumacher. Model risk and regulatory capital. *Available at SSRN 301531*, 2002.
- [26] I. Kojadinovic and J. Yan. Comparison of three semiparametric methods for estimating dependence parameters in copula methods. *Insurance : Mathematics and Economics*, 47(1) :52–63, 2010.
- [27] T. Lallement. Le risque de modèle. *Mémoire d'actuariat présenté à l'IA.*, 2014.
- [28] D.X. Li. On default correlation : A copula function approach. *Journal of Fixed Income*, 2000.
- [29] T. Mack. Distribution-free calculation of the standard error in the chain-ladder reserving method. *ASTIN Bulletin*, 23(2) :213–225, 1993.
- [30] T. Mack. Measuring the variability of chain ladder reserve estimates. *Casualty Actuarial Society*, 1993.
- [31] B. Mandelbrot. B. mandelbrot interviewé par a. kahn : "il était inévitable que des choses très graves se produisent". *Le monde*, 2009.
- [32] R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2) :125–144, 1976.
- [33] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), 1953.
- [34] J. Nelder and R. Wedderburn. Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, 135(3) :370–384, 1972.
- [35] Board of Governors of the Federal Reserve System et OCC. Supervisory guidance on model risk management. *OCC Bulletin*, 2011.
- [36] Comptroller of the Currency Administrator of National Banks. Risk modeling. *OCC Bulletin*, 2000.
- [37] F. Planchet and P.E. Therond. L'impact de la prise en compte des sauts boursiers dans les problématiques d'assurance. *Journal of Economic Literature Classification*, 2005.
- [38] R. Rebonato. *Managing Model Risk*, volume 2. in Handbook of Risk Management, 2001.
- [39] SGP&C Reserving. Coefficient of variation best practice. *Reserves Variability*, June 2015.
- [40] F. Salmon. Recipe for disaster : The formula that killed wall street. *Wired Magazine*, 2009.
- [41] A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 2 :229–231, 1959.
- [42] F. Wang and A. Sakuth. Model risk. *MSA400 Financial risk*, 2012.

## Table des figures

1	Cartographie du risque de modèle . . . . .	9
2	Cours boursier britannique entre 2001 et 2014 . . . . .	25
3	Rendements logarithmiques mensualisés du cours britannique entre 2001 et 2014 . . . . .	26
4	Moyenne et volatilité des rendements logarithmiques mensuels par année de 2001 à 2013 . . . . .	32
5	Graphique Quantile-Quantile des rendements logarithmiques mensuels 2001-2013 . . . . .	33
6	Autocorrélation des carrées des rendements logarithmiques journaliers sur 2001-2013 . . . . .	34
7	Autocorrélation des carrées des rendements logarithmiques mensuels sur 2001-2013 . . . . .	34
8	Prévisions des rendements logarithmiques par le modèle de Garch en journalier . . . . .	34
9	Prévisions des rendements logarithmiques par le modèle de Garch en mensuel . . . . .	35
10	Prévisions du cours boursier pour l'année 2014 par le modèle de BS-RS . . . . .	35
11	PIT modèle de Black-Scholes à changements de régime . . . . .	38
12	Autocorrélogrammes modèle de Black-Scholes à changements de régime . . . . .	38
13	Nuage de points des rangs sur les indices mensualisés US et UK . . . . .	48
14	Comparaison des VaR à 0,5% pour différentes marges et copules . . . . .	53
15	Nuage de points des rangs sur les indices journaliers US et UK . . . . .	55
16	Sinistres cumulés par période de développement pour chaque année d'origine . . . . .	59
17	Relation entre les paiements cumulés de deux périodes de développement successives . . . . .	70
18	Résidus de Mack en fonction des paiements cumulés . . . . .	70
19	Valeurs observées vs valeurs ajustées pour le <i>GLMPoisson</i> . . . . .	72
20	Résidus de Pearson du modèle linéaire généralisé 1 (lien 'log' distribution de Poisson) . . . . .	73
21	Comparaison de l'histogramme des résidus du GLM Poisson avec la densité Gaussienne . . . . .	73
22	Test quantiles-quantiles de normalité des résidus pour le GLM Poisson . . . . .	73
23	Classification des modèles de provisionnement . . . . .	75
24	Densités de Kernel des distributions des provisions totales pour différents modèles . . . . .	76
25	Variables observées vs estimées pour les GLM . . . . .	80
26	Nuage de points des rangs . . . . .	81
27	Arbre de pondération pour les deux risques non-vie . . . . .	87

## Liste des tableaux

1	Tests d'adéquation . . . . .	39
2	Test d'indépendance . . . . .	40
3	Vraisemblance de chaque modèle . . . . .	41
4	Probabilités <i>a posteriori</i> . . . . .	41
5	VaR à 0.5% sur 100 itérations de distributions de taille 20 000 . . . . .	42
6	Erreur de Modèle sur des distributions de taille 20000 avec estimation sur 100 itérations	42
7	Distributions <i>a posteriori</i> . . . . .	43
8	Quantiles à 0.5%, Chocs à 0.5% et Erreur de modèle associée . . . . .	43
9	VaR à 0,5% et erreur de modèle sur 200 itérations des distributions de taille 50000 . .	44
10	Erreur de modèle pour l'approche <i>worst case</i> . . . . .	44
11	Synthèse des erreurs par les deux approches . . . . .	45
12	VaR à 0,5% et erreur de modèle sur 100 itérations de distributions de taille 20 000 . .	46
13	Valeurs du CQEP low tail pour différents niveaux . . . . .	49
14	Paramètres estimés pour différentes copules et différents niveaux . . . . .	49
15	Paramètres estimés par pseudo maximum de vraisemblance . . . . .	50
16	VaR à 0,5% sur 100 itérations et le modèle de Garch pour les marges . . . . .	52
17	AM en supposant successivement chaque copule comme copule de référence . . . . .	52
18	Log-vraisemblances, pondérations et erreur de modèle sur 100 itérations de distributions de taille 10 000 . . . . .	54
19	Test d'absence d'effets calendaires . . . . .	68
20	Statistiques de test et intervalles de confiance . . . . .	69
21	Erreur de modèle <i>worst case</i> intra-classe . . . . .	76
22	Critère d'Akaike et pondérations associées . . . . .	78
23	Erreur de modèle intra-classe par pondération . . . . .	78
24	Paramètres et quantiles à 99.5% par la méthode CQEP . . . . .	82
25	Paramètres et quantiles à 99.5% par la méthode du $\rho$ de Spearman . . . . .	83
26	Paramètres des copules et quantiles à 99.5% en l'absence de dépendance . . . . .	84
27	Paramètres des copules et quantiles à 99.5% pour une dépendance parfaite . . . . .	84
28	Deux pondérations des copules . . . . .	85
29	VaR à 99.5% pour l'approche par pondération . . . . .	85
30	Erreur de dépendance pure par les deux approches . . . . .	85
31	Erreur de modèle globale par le pire des cas pour chaque classe . . . . .	86
32	VaR à 99.5% pour l'approche par pondération . . . . .	86

# Appendices

## A Etudes des modèles d'actifs

### A.1 Les modèles d'actifs

#### A.1.1 Black-Scholes

Le tableau suivant présente les paramètres estimés par le modèle de Black-Scholes.

TABLE 33 – Paramètres estimés Black-Scholes

	$\mu$	$\sigma$
2001	-0.01044	0.0395
2001-2002	-0.01631	0.0499
2001-2003	-0.00610	0.0494
2001-2004	-0.00231	0.0440
2001-2005	0.00120	0.0412
2001-2006	0.00290	0.0388
2001-2007	0.00324	0.0372
2001-2008	-0.00065	0.0424
2001-2009	0.00168	0.0556
2001-2010	0.00247	0.0549
2001-2011	0.00211	0.0534
2001-2012	0.00261	0.0518
2001-2013	0.003498	0.0507

Les paramètres varient en fonction de l'historique de données considéré. Cela met en doute l'hypothèse de constance des paramètres.

Le graphe suivant présente la moyenne et l'écart-type des rendements logarithmiques obtenus sur les observations journalières de chaque année.

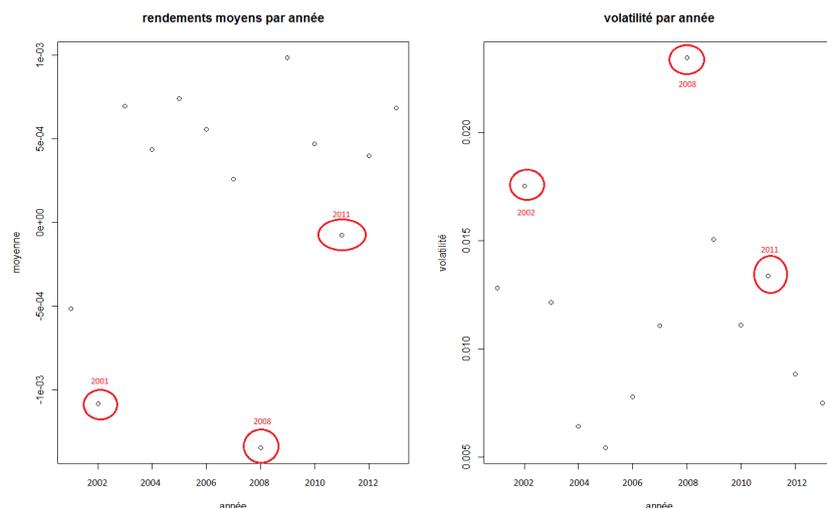


FIGURE 28 – Moyenne et volatilité des rendements logarithmiques journaliers par année de 2001 à 2013

Comme pour les données mensuelles, nous observons clairement trois périodes de crise : en 2002, en 2008 et en 2011.

### A.1.2 Merton

Le tableau suivant présente les paramètres estimés par le modèle de Merton.

TABLE 34 – Paramètres estimés Merton

	$\mu$	$\sigma^2$	$\lambda$	$\sigma_X^2$
2001	-0.00972	0.0015	3.76e-02	0.0004
2001-2002	-0.01512	0.0024	-9.54e-05	0.0802
2001-2003	-0.00493	0.0023	5.518e-02	0.0005
2001-2004	-0.00146	0.0017	5.61e-02	0.0035
2001-2005	0.00195	0.0015	3.28e-02	0.0054
2001-2006	0.00363	0.0015	3.63e-04	0.0533
2001-2007	0.00387	0.0013	1.17e-02	0.0093
2001-2008	0.00011	0.0015	5.36e-02	0.0050
2001-2009	0.00282	0.0023	3.52e-02	0.0224
2001-2010	0.00359	0.0022	3.70e-02	0.0206
2001-2011	0.00314	0.0021	3.97e-02	0.0193
2001-2012	0.00357	0.0019	3.99e-02	0.0185
2001-2013	0.00435	0.0017	5.57e-02	0.0152

Les paramètres des variances sont tous positifs, ce qui valide l'utilisation de la méthode des moments pour l'estimation des paramètres.

### A.1.3 Garch

Le tableau suivant présente les paramètres estimés par le modèle de Garch.

TABLE 35 – Paramètres estimés Garch(1,1)

	$\lambda$	$\omega$	$\alpha$	$\beta$
2001	-1.064e-02	1.259e-05	1.00e-08	1.00
2001-2002	-1.535e-02	9.075e-04	1.45e-01	4.88e-01
2001-2003	-3.950e-03	5.872e-04	1.33e-01	6.23e-01
2001-2004	1.046e-02	2.234e-05	5.31e-01	5.53e-01
2001-2005	1.043e-02	5.174e-05	4.76e-01	5.68e-01
2001-2006	9.126e-03	7.453e-05	3.13e-01	6.52e-01
2001-2007	8.159e-03	8.585e-05	2.69e-01	6.78e-01
2001-2008	6.522e-03	6.041e-05	3.43e-01	6.91e-01
2001-2009	6.894e-03	2.545e-05	3.53e-01	7.31e-01
2001-2010	7.419e-03	5.596e-05	3.70e-01	6.84e-01
2001-2011	7.292e-03	7.882e-05	3.62e-01	6.74e-01
2001-2012	7.106e-03	8.156e-05	3.29e-01	6.86e-01
2001-2013	7.223e-03	1.0384e-04	2.63e-01	7.19e-01

Les paramètres semblent relativement stable en augmentant l'historique de données.

## A.2 PIT

### A.2.1 Black-Scholes

Les graphes suivants représentent l'histogramme, la fonction de répartition empirique et l'autocorrélogramme obtenus pour le modèle de Black-Scholes.

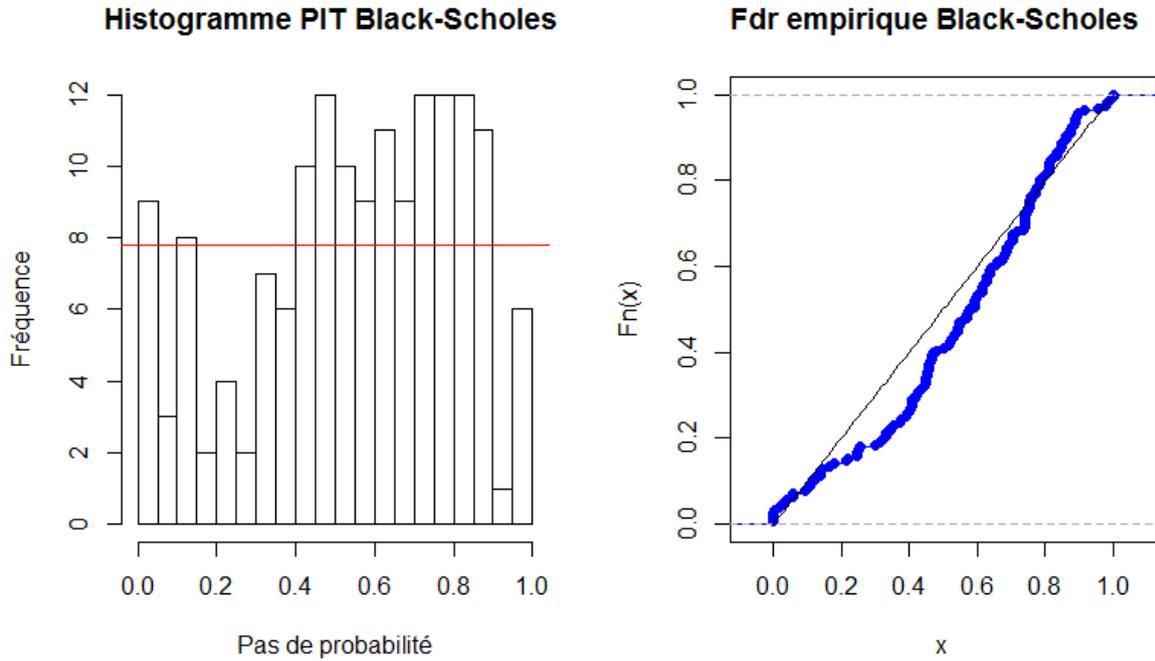


FIGURE 29 – PIT modèle de Black-Scholes

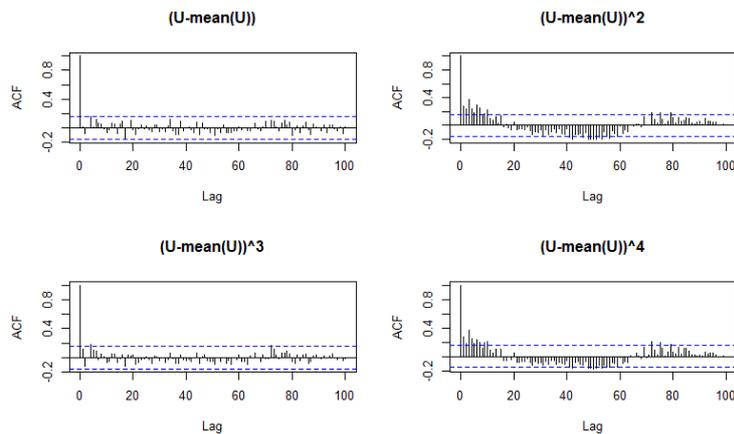


FIGURE 30 – Autocorrélogrammes modèle de Black-Scholes

L'uniformité des variables PIT est remise en question par les deux premiers graphes, de même que l'absence de corrélation des données aux ordres 2 et 4 par le troisième.

### A.2.2 Merton

Les graphes suivants représentent l'histogramme, la fonction de répartition empirique et l'autocorrélogramme obtenus pour le modèle de Merton.

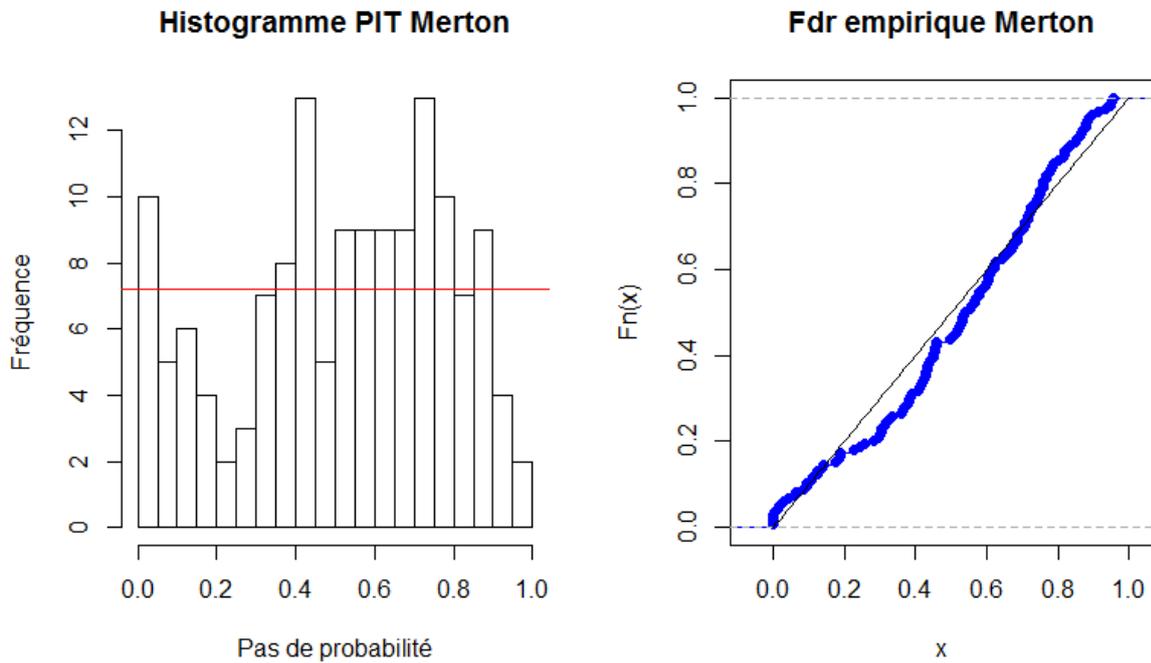


FIGURE 31 – PIT modèle de Merton

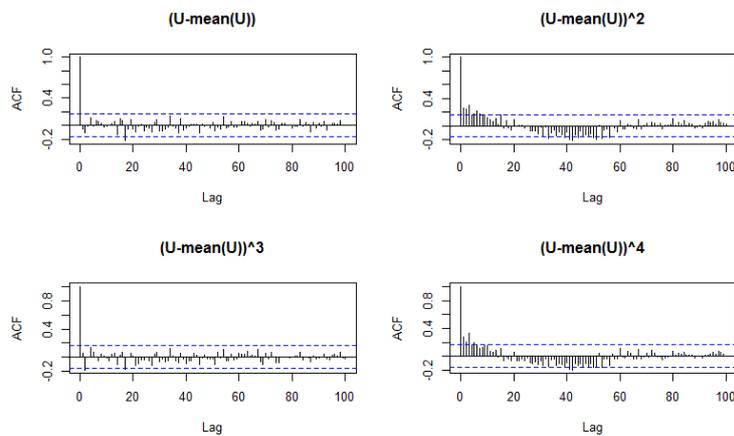


FIGURE 32 – Autocorrélogrammes modèle de Merton

L'uniformité des variables PIT est remise en question par les deux premiers graphes, de même que l'absence de corrélation des données aux ordres 2 et 4 par le troisième.

Le graphes suivant représentent l’histogramme, la fonction de répartition empirique et l’autocorrélogramme obtenus pour le modèle de Garch.

### A.2.3 Garch

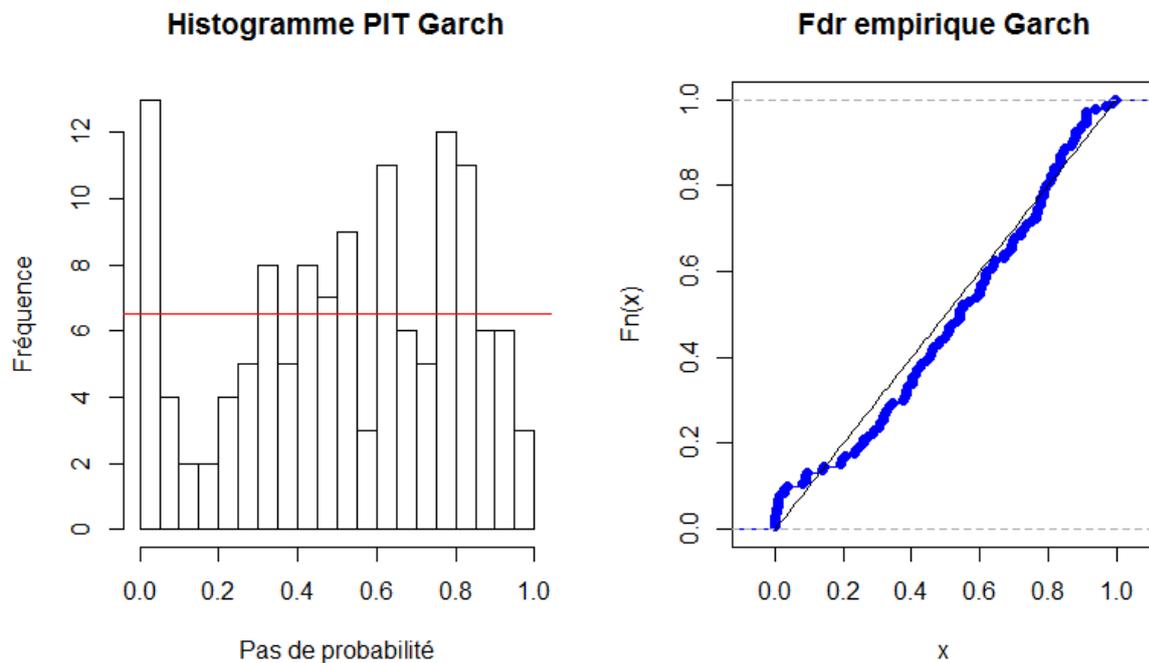


FIGURE 33 – PIT modèle de Garch

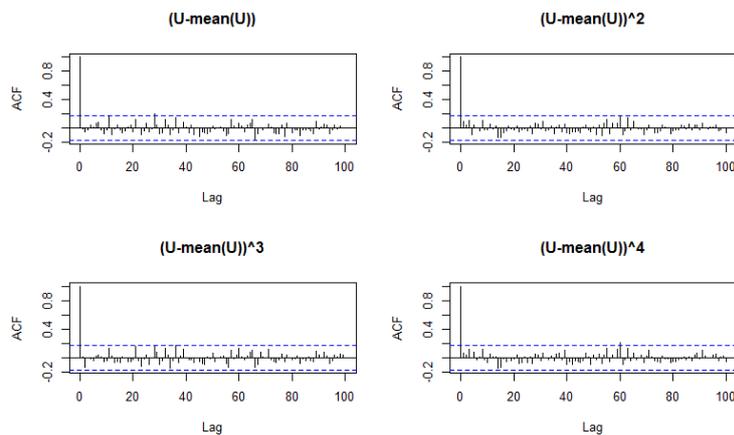


FIGURE 34 – Autocorrélogrammes modèle de Garch

L’hypothèse d’uniformité des variables PIT paraît davantage justifiée que pour les autres modèles d’actifs. Par ailleurs, le caractère indépendant des variables semble vérifié.

### A.3 Statistiques descriptives des rendements journaliers

Le graphe suivant présente l'évolution des rendements logarithmiques journaliers entre 2001 et 2014.

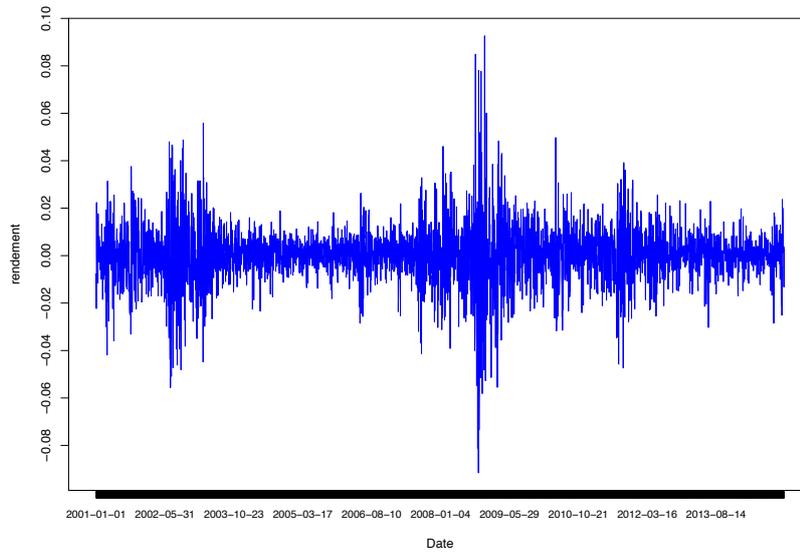


FIGURE 35 – Rendements logarithmiques du cours britannique entre 2001 et 2014

Le graphique quantile-quantile suivant montre, comme en base mensuelle, la présence d'évènements extrêmes correspondant aux périodes de crise du XXIème siècle.

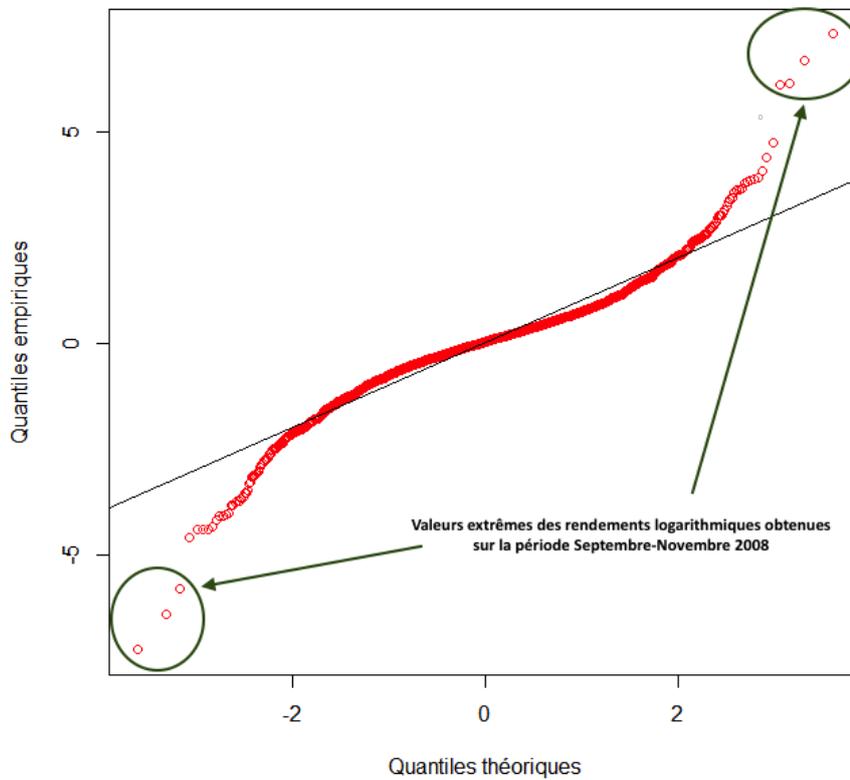


FIGURE 36 – Graphique Quantile-Quantile des rendements logarithmiques mensuels 2001-2013

## A.4 Erreur de dépendance pure sur données journalières d'actifs

### A.4.1 Estimation des paramètres

Les paramètres des copules sont estimés en maximisant la pseudo-vraisemblance grâce à la fonction  $R_{fitCopula}$ .

TABLE 36 – Paramètres estimés par pseudo maximum de vraisemblance

degrés	Normal	Student				Clayton	Gumbel	Frank
		2	2.56	3	4			
paramètre	0.555705	0.53594	0.56032	0.5717	0.5856	0.95791	1.61832	4.03377
statistique	0.13378	0.03875	-	0.046464	0.087806	0.78173	0.2296	0.39923
p-value	0.0005	0.02348	-	0.00749	0.0005	0.0005	0.0005	0.0005
log-vrais	589.1	772.8364	780.5	778.2335	765.3928	544.8	650.6	554.3

La vraisemblance est bien supérieure pour les copules de Student, cela justifie l'utilisation de la copule de Student à 2.56 degrés de liberté comme référence.

### A.4.2 Pondération des copules

Les pondérations sont obtenues à partir des pondérations de l'application en mensuel dans la mesure où les vraisemblances sont trop élevées pour mettre en place correctement l'approche bayésienne. Cependant, les modèles ayant la plus grande vraisemblance se sont vus attribuer la plus grande pondération.

TABLE 37 – Log-vraisemblances et pondérations

Degrés	Normal	Student				Clayton	Gumbel	Frank
		2	2,56	3	4			
log-vrais	589,10	772,84	780,50	778,23	765,39	544,80	650,60	554,30
Poids 1	13,673%	22,286%	23,315%	23,215%	17,506%	0,000%	0,005%	0,000%
log-vrais	589,10	-	780,50	-	-	544,80	650,60	554,30
Poids 2	31,589%	-	68,402%	-	-	0,000%	0,010%	0,000%

### A.4.3 Approche *worst case* pour le modèle de Garch

Le tableau suivant récapitule les quantiles à 99.5% obtenus pour chaque copule en fixant les marges et utilisés pour mesurer l'erreur dans le pire des cas.

TABLE 38 – VaR à 0,5% sur 100 itérations et le modèle de Garch pour les marges

Degrés	Garch				
	Normal	Student			
		2	2.56	3	4
Moyenne	-36 156 775	-37 556 050	-37 545 597	-37 379 705	-37 470 901
Ecart-type	759 328	871 117	911 754	943 180	869 327
CoV	2.100%	2.320%	2.428%	2.523%	2.320%
	Clayton	Gumbel	Frank		
Moyenne	-38 946 705	-34 586 133	-33 680 518		
Ecart-type	1 080 496	712 489	626 126		
CoV	2.774%	2.060%	1.859%		

#### A.4.4 Approche Bayésienne

Les quantiles à 99.5% issus du bootstrap de 8 ou 5 copules respectivement avec pondération bayésienne sont présentés dans les tableaux ci-dessous, ainsi que l'erreur de modèle associée.

TABLE 39 – Erreur de modèle bayésiennes avec les poids 1 sur des distributions de taille 10000 avec 100 itérations

VAR à 0.5%	Moyenne	%	Ecart-type	CoV	Erreur
<b>BS</b>	-36 020 370	-36.020%	639 697	1.776%	-0.567%
<b>Merton</b>	-36 784 558	-36.785%	685 040	1.862%	-0.652%
<b>Garch</b>	-37 374 148	-37.374%	897 966	2.403%	-0.457%
<b>BSRS</b>	-34 453 175	-34.453%	633 860	1.840%	-0.516%

TABLE 40 – Erreur de modèle bayésiennes avec les poids 2 sur des distributions de taille 10000 avec 100 itérations

VAR à 0.5%	Moyenne	%	Ecart-type	CoV	Erreur
<b>BS</b>	-35 706 705	-35.707%	596 217	1.670%	-1.433%
<b>Merton</b>	-36 459 376	-36.459%	618 896	1.697%	-1.530%
<b>Garch</b>	-36 967 538	-36.968%	818 601	2.214%	-1.540%
<b>BSRS</b>	-34 133 888	-34.134%	589 808	1.728%	-1.438%

L'erreur est négative dans tous les cas puisque le modèle de référence (copule de Student à 2.56 degrés de liberté et marges de Garch) est le plus prudent.

#### A.4.5 Marges et Copules

Le graphe suivant présente les quantiles à 99.5% pour chaque copule et chaque marge. La copule

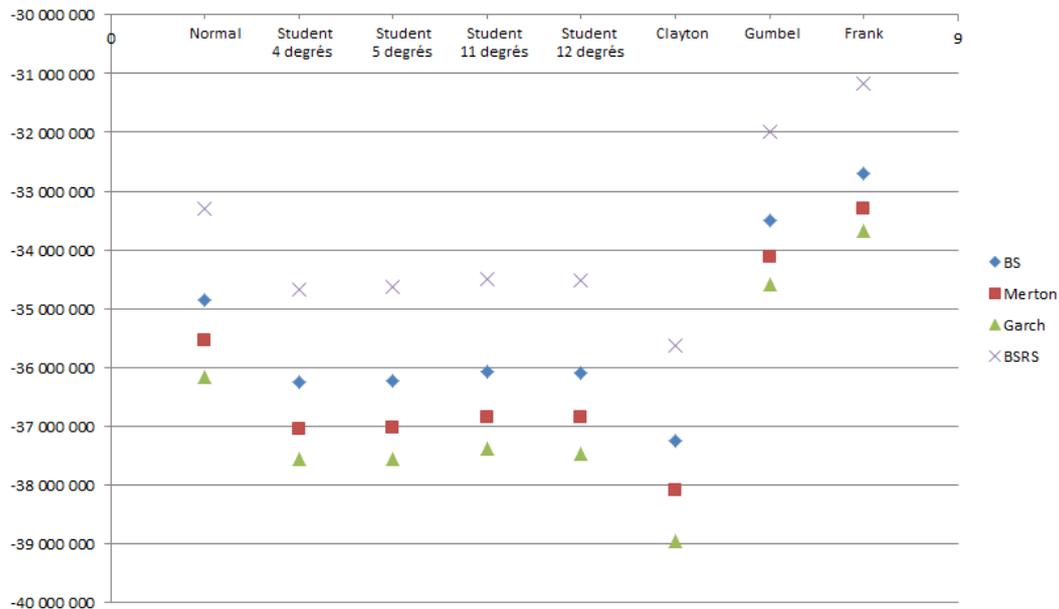


FIGURE 37 – Comparaison des VaR à 0,5% pour différentes marges et copules

la plus prudente est la copule de Clayton avec marges de Garch.

## B Etudes modèles de provisionnement

### B.1 Validation des hypothèses

#### B.1.1 Corrélation des rangs

Le premier tableau présente les rangs des coefficients de développement évalués par année de développement. Le deuxième tableau donne les rangs obtenus en enlevant la dernière observation de chaque période de développement.

TABLE 41 – Rangs des coefficients de développement avec et sans le dernier de chaque période de développement

$r_{i,1}$	$r_{i,2}$	$r_{i,3}$	$r_{i,4}$	$r_{i,5}$	$r_{i,6}$	$r_{i,7}$	$r_{i,8}$	$r_{i,9}$	$r_{i,10}$	$r_{i,11}$	$r_{i,12}$	$r_{i,13}$	$r_{i,14}$
2	6	1	7	6	1	5	7	6	1	4	3	2	1
8	7	4	5	2	6	6	6	4	5	3	2	1	
4	4	9	9	8	8	8	1	3	2	2	1		
9	11	12	11	10	9	7	2	5	3	1			
11	13	5	10	9	7	2	3	2	4				
13	12	8	6	5	5	1	4	1					
14	9	10	3	3	4	4	5						
12	10	11	8	4	3	3							
10	8	6	4	1	2								
1	2	7	2	7									
5	5	3	1										
6	3	2											
7	1												
3													

$s_{i,1}$	$s_{i,2}$	$s_{i,3}$	$s_{i,4}$	$s_{i,5}$	$s_{i,6}$	$s_{i,7}$	$s_{i,8}$	$s_{i,9}$	$s_{i,10}$	$s_{i,11}$	$s_{i,12}$
2	5	1	6	6	1	4	6	5	1	3	2
7	6	3	4	2	5	5	5	3	4	2	1
3	3	8	8	7	7	7	1	2	2	1	
8	10	11	10	9	8	6	2	4	3		
10	12	4	9	8	6	2	3	1			
12	11	7	5	5	4	1	4				
13	8	9	2	3	3	3					
11	9	10	7	4	2						
9	7	5	3	1							
1	1	6	1								
4	4	2									
5	2										
6											

Ce tableau nous permet ensuite de vérifier l'hypothèse d'indépendance des facteurs de développement individuels pour les modèles Chain-Ladder et Mack.

## B.1.2 Régression linéaires des paiements cumulés

Ce graphe représente les relations entre les paiements cumulés de deux périodes de développement successives pour les dernières périodes de développement.

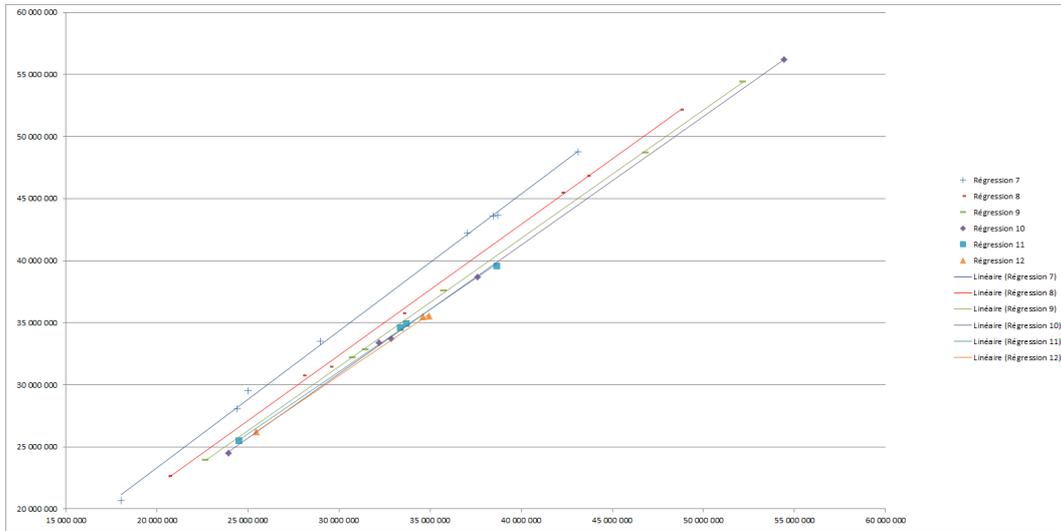


FIGURE 38 – Régressions linéaires des paiements cumulés de deux périodes de développement successives

Nous observons que l'hypothèse de linéarité des paiements cumulés est plus difficile à vérifier sur les dernières périodes de développement sûrement à cause du manque de données.

## B.2 GLM

### B.2.1 résidus de Pearson

Les résidus de Pearson sont définis par :

$$r_{i,j}^p = \frac{Y_{i,j} - \mu_{i,j}}{\sqrt{V(\mu_{i,j})}}$$

ainsi l'espérance et la variance conditionnelle à l'information détenue pour l'année d'origine  $i$  et la période de développement  $j$  sont :

$$\mathbb{E}(Y_{i,j} | \mathcal{F}_{i,j}) = \mu_{i,j} \quad \Rightarrow \quad \mathbb{E}(r_{i,j}^p | \mathcal{F}_{i,j}) = 0$$

$$\mathbb{V}(Y_{i,j} | \mathcal{F}_{i,j}) = \frac{\phi}{w} V(\mu_{i,j}) \quad \Rightarrow \quad \mathbb{V}(r_{i,j}^p | \mathcal{F}_{i,j}) = \frac{\mathbb{V}(Y_{i,j} | \mathcal{F}_{i,j})}{V(\mu_{i,j})} = \frac{\phi}{w}$$

### B.2.2 GLM 2 : lien 'log' distribution gamma

Nous présentons le graphe des résidus de Pearson ainsi que l'histogramme et le graphe quantile-quantile. Le premier graphe permet de vérifier le caractère aléatoire des résidus.

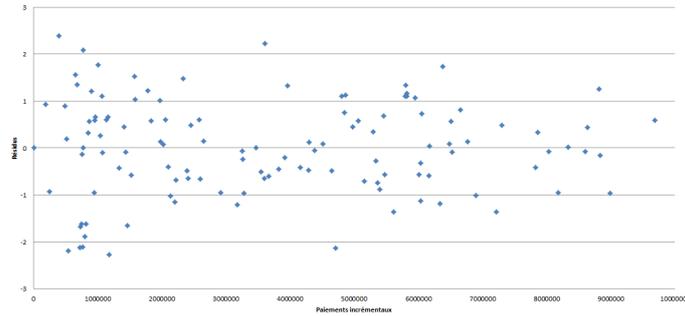


FIGURE 39 – Résidus de Pearson du modèle linéaire généralisé Gamma

Cette hypothèse est difficile à vérifier mais nous n'observons pas ici de tendance particulière. L'histogramme et le QQ-plot servent à vérifier la normalité des résidus.

FIGURE 40 – Histogramme des résidus du GLM Gamma

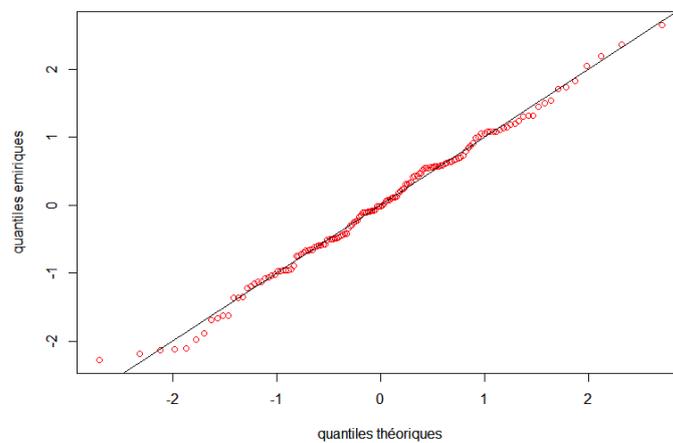
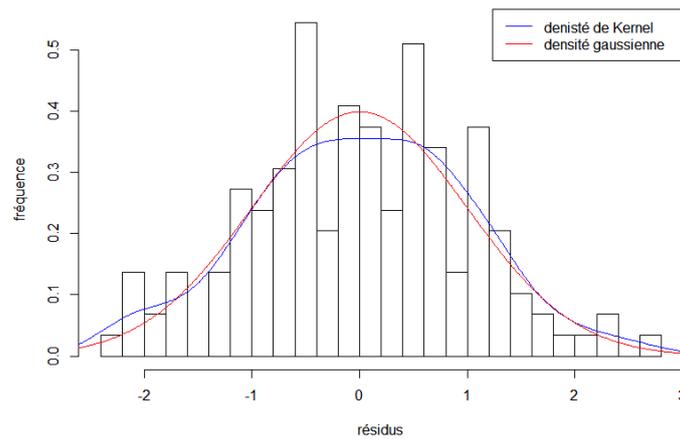


FIGURE 41 – QQ-plot des résidus du GLM Gamma

Au vu des résultats, nous ne pouvons pas rejeter cette hypothèse.

Les valeurs observées sont ensuite confrontées aux valeurs estimées par le modèle :

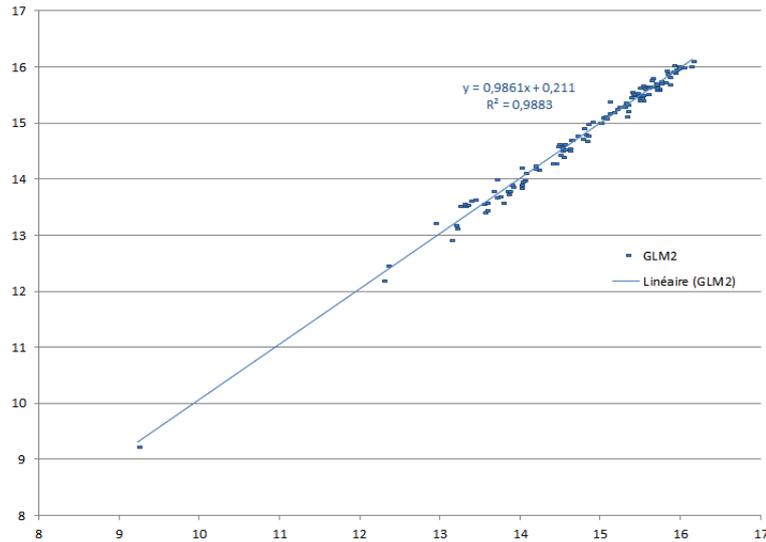


FIGURE 42 – Valeurs observées vs valeurs ajustées pour le GLM Gamma

Le graphique montre très clairement que les valeurs estimées sont très proches des valeurs observées (cf la linéarité de la régression des unes par rapport aux autres).

### B.2.3 GLM 3 : lien 'identité' distribution log-normale

Nous effectuons les mêmes graphes que dans la section précédente en choisissant une erreur log-normale.

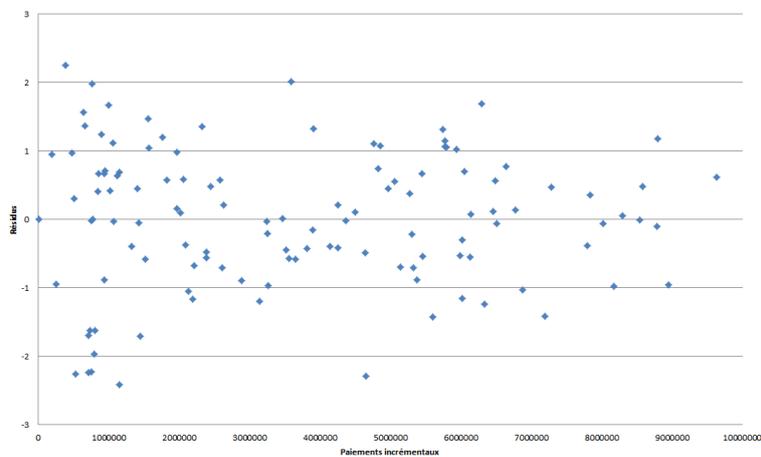


FIGURE 43 – Résidus de Pearson du modèle linéaire généralisé Log-normal

Les résidus ne semblent pas complètement aléatoires et il semblerait qu'il y ait un caractère décroissant des résidus avec le montant des paiements incrémentaux.

FIGURE 44 – Histogramme des résidus du GLM Log-normal

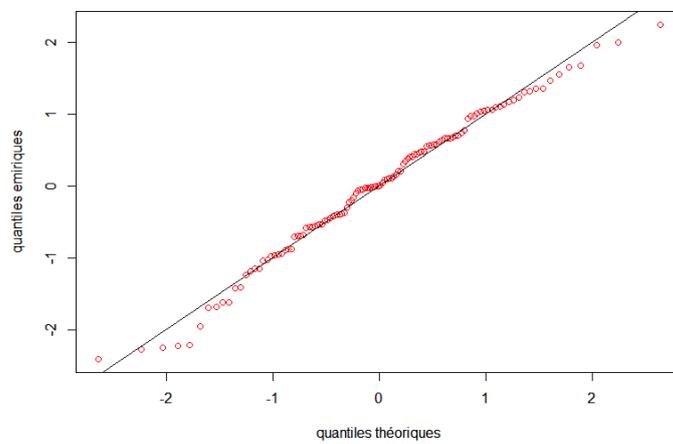
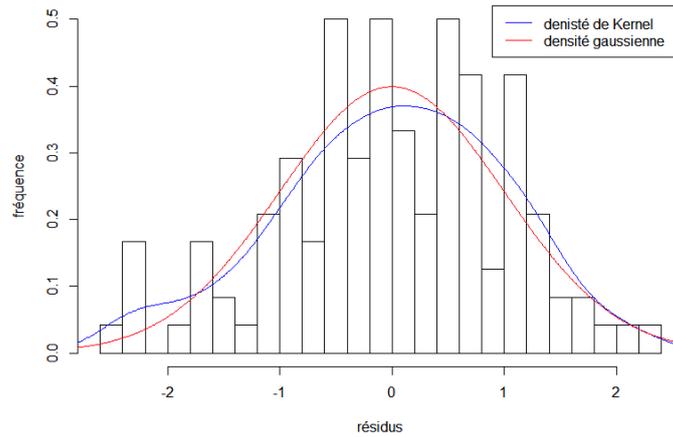


FIGURE 45 – QQ-plot des résidus du GLM Log-normal

De même, l'hypothèse de normalité des résidus semble rejetée puisque la distribution présente des valeurs extrêmes notamment dans la queue inférieure de la distribution.

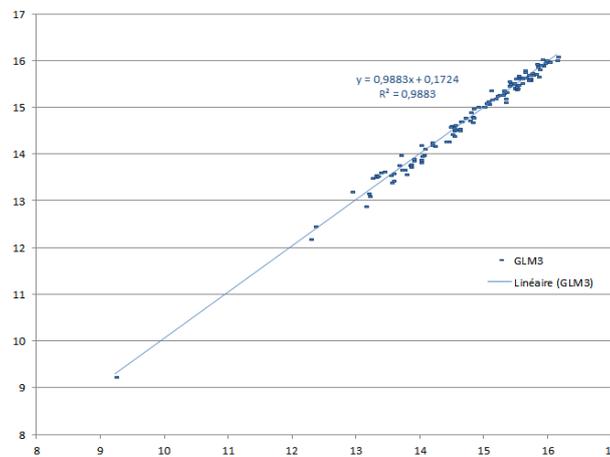


FIGURE 46 – Valeurs observées vs valeurs ajustées pour le GLM Log-normal

Cependant, les valeurs estimées semblent coïncider avec les valeurs observées.

## B.3 Evaluation de l'erreur de modèle

### B.3.1 Modèle de référence

Pour valider l'utilisation du modèle 'log-normal' comme modèle de référence, nous avons effectué des tests d'adéquation, en l'occurrence, le test de Kolmogorov-Smirnov et le test d'Anderson-Darling. Pour ce faire, nous avons confronté la distribution obtenue à partir de ce modèle avec celle obtenue avec les autres modèles (Bootstrap ODP et Mack avec erreur Gamma et MCMC GLM avec erreur Poisson).

TABLE 42 – p-value des tests d'adéquation

	<i>Bootstrap ODP</i>	<i>Bootstrap Mack</i>	<i>MCMC GLM</i>
<b>KS</b>	2.0e-9	0.7576	<2.2e-16
<b>AD</b>	6.0e-8	0.4672	6.0e-8

Finalement, la distribution log-normale est très proche de la distribution issue du bootstrap Mack (les p-values sont très nettement supérieures à 5%).

### B.3.2 L'approche par pondération

Pour les modèles des classes bootstrap Mack et bootstrap ODP, la vraisemblance n'est pas calculable. Nous avons donc cherché un autre moyen pour évaluer les pondérations. Ainsi, nous avons divisé l'erreur de prédiction (CoV) en erreur d'estimation et erreur de processus. Cependant les différences entre les erreurs n'étant pas significatives, nous avons finalement retenu la pondération équiprobable dans chaque classe de modèles.

TABLE 43 – Erreur de prédiction (CoV) et pondérations des classes bootstrap ODP et Mack

	<b>Distribution de l'erreur</b>	<b>Gamma</b>	<b>Log-normale</b>	<b>Normale</b>	<b>Rééchantillonnée</b>
<i>ODP</i>	CoV	5.27%	5.19%	5.20%	5.25%
	Erreur d'estimation	5.11%	5.11%	5.11%	5.11%
	Erreur de processus	0.16%	0.08%	0.09%	0.14%
	Pondération	25%	25%	25%	25%
<i>Mack</i>	CoV	4.55%	4.57%	4.57%	4.61%
	Erreur d'estimation	2.41%	2.41%	2.41%	2.41%
	Erreur de processus	2.14%	2.16%	2.41% 2.16%	2.20%
	Pondération	25%	25%	25%	25%

## B.4 Evaluation de l'erreur dépendance globale pour les risques non-vie

Afin d'évaluer l'erreur de modèle globale, nous avons besoin des quantiles à 99.5% obtenus à partir de chaque copule et chaque marge qui sont présentés dans le tableau ci-dessous.

TABLE 44 – Quantile à 99.5% pour chaque copule et chaque marge

<b>Bootstrap Mack</b>			
<i>Gamma</i>	<i>Log-normale</i>	<i>Normale</i>	<i>Ré-échantillonner</i>
354 080 815	356 182 207	353 543 156	355 010 327
353 101 569	355 265 134	352 258 590	353 831 860
353 695 102	355748973	352 967 983	354 609 496
352 586 523	354 547 532	352 057 488	353 412 375
354 688 495	356 559 650	354 087 268	355 852 468
<b>Bootstrap ODP</b>			
<i>Gamma</i>	<i>Log-normale</i>	<i>Normale</i>	<i>Ré-échantillonner</i>
358 757 208	359 667 179	360 149 726	360 044 893
357 821 907	358 522 862	359 020 912	358 977 142
358 439 201	359 014 286	359 687 987	359 832 519
357 131 124	358 027 527	358 583 919	358 569 509
359 245 334	360 162 333	360 741 964	360 692 373
<b>GLM MCMC</b>			
	<i>Poisson</i>	<i>Gamma</i>	
	366 299 766	378 339 991	
	365 509 810	377 649 410	
	365 926 370	378 097 603	
	364 730 668	376 884 818	
	366 984 972	379 125 275	

*Remarque.* Il est à noter que, n'utilisant pas le modèle GLM gamma pour le segment auto, nous avons donc simulé les marges par deux modèles différents dans ce cas, le modèle GLM gamma pour la marge correspondant au segment marine et le modèle 'log-normal' pour le segment auto.