

**Mémoire présenté devant l'Institut du Risk Management
pour la validation du cursus à la Formation d'Actuaire
de l'Institut du Risk Management
et l'admission à l'Institut des actuaires
le**

Par : Abdoulaye TOURE

Titre : Modélisation des obligations catastrophes paramétriques dans un marché incomplet
par une approche deep learning et une approche de modélisation par loi

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

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

Membres présents du jury de l'Institut des actuaires :

Membres présents du jury de l'Institut du Risk Management :

Secrétariat :

Bibliothèque :

Entreprise : GIE AXA

Nom :

Signature et Cachet

GIE AXA
21-25, avenue Maignon
75008 PARIS

Directeur de mémoire en entreprise :

Nom : LACHENAUD Grégoire

Signature :

Invité :

Nom :

Signature :

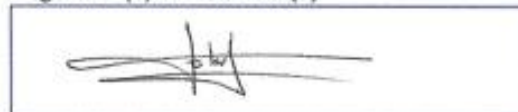
Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels

(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise



Signature(s) du candidat(s)



Résumé:

Mots clés:

CAT Bond, CAT Bond paramétrique, Evaluation CAT Bond, modélisation GEV, Modèle de Vasicek, Deep Learning, réseau de neurones récurrents (RNN), LSTM (Long Short-Term Memory), GRU.

Ce mémoire étudie des approches de valorisation d'une obligation catastrophe (ou CAT Bond) paramétrique.

Les pertes économiques dues aux catastrophes naturelles représentent un sujet majeur pour les pays émergents et en voie de développement dans leur avancée vers le développement, ainsi que pour les assureurs dans la recherche de protection contre des risques à pertes potentielles extrêmes.

Dans le même temps, les obligations catastrophes qui pourraient permettre aux pays ou aux (ré)assureurs d'avoir accès aux marchés pour transférer certains de leurs risques, du fait de leur structure, sont des produits complexes dans le sens que la mise en place d'une couverture de ces produits est très complexe avec les produits financiers traditionnels. Ce qui constitue un frein pour des investisseurs, qui bien qu'intéressés par leur haut rendement, redoutent la complexité des CAT Bond et des risques qui y sont associés.

Cette étude a pour objectif de présenter deux approches de valorisation des CAT Bond à déclenchement paramétrique dans un modèle à temps discret. La première approche sera réalisée à partir d'un modèle séquentiel des réseaux de neurones et l'autre selon une approche de modélisation par une loi de probabilité.

Cette étude a été réalisée en quatre grandes sections. La première partie présentera les CAT Bond dans leurs généralités. Ensuite un focus sera réalisé sur les réseaux de neurones séquentiels dans la seconde partie. Dans la troisième partie, les deux approches de modélisation seront étudiées. Enfin, dans la dernière partie, une application sera effectuée sur les risques liés aux tempêtes dans le sud-est des Etats-Unis. Ces deux approches permettront de faire une analyse comparative des deux modèles de valorisation. Afin de réaliser cette étude, les données d'analyse sont issues des données du NOAA.

ABSTRACT

Keywords:

CAT Bond, parametric CAT Bond, CAT Bond pricing, GEV modelling, Vasicek modelling, Deep learning, recurrent neural network, RNN, LSTM (long short-term memory), GRU.

This paper aims at studying the pricing of the parametric CAT Bond.

The economic losses generated by the Cat Nat events are very challenging for emerging countries whose economy development is at risk due to catastrophic events or for (re)insurers in their objective of transferring some of exposure with potential high losses to third parties. In addition to that, the catastrophes bond which were supposed to help facing the economic losses of CAT NAT events are not easily interpretable by investor not familiar to insurance risks. In fact, hedging strategies with primary financial product are complex.

This document aims to provide two approaches for pricing a parametric CAT Bond. The first approach is based on a sequential Deep Learning model using neural network and the second one is based on the probability distribution.

This paper can be split in four sections. In the first section, the CAT Bond product is briefly described. In the second section, the Deep Learning sequential model are presented. The third section is dedicated to the modelling approaches. Finally, in the last section, an application of the two approaches on the south-east US storms events is proposed. For the purpose of modelling the CAT Bond price, this study sourced its data from NOAA historical database.

1. Contexte et problématique de l'étude :

En 2019, selon les données du Swiss Re Institute (Sigma n° 2/2020, page 29), les pertes économiques dues aux catastrophes naturelles sont estimées à 146 milliards \$ dont environ 86 milliards \$ des pertes non couvertes par une assurance soit environ 59% du total des pertes économiques.

Cette marge importante de pertes économiques non couvertes par une assurance, est une preuve de la nécessité de trouver des alternatives de protection ou de financement contre les pertes économiques dues aux catastrophes naturelles. Pour faire face à cette situation, certains pays (par le biais du programme MultiCat de la Banque Mondiale) et des (ré)assureurs se sont intéressés aux obligations catastrophes.

Ces produits qui permettent à l'émetteur de transférer certains de ses risques vers les marchés financiers, ont des hauts rendements comparés aux obligations classiques. Pourtant, les investisseurs ne se bousculent pas sur le marché des CAT Bond. Cette situation est principalement due à la complexité du produit CAT Bond. En effet, la mise en place de stratégies de couverture, pour un CAT Bond, est pratiquement impossible avec des produits financiers traditionnels. De plus, le modèle de valorisation du prix d'un CAT Bond requiert un modèle de marché incomplet: ce qui implique que l'environnement risque neutre n'est pas unique pour ces produits et donc plusieurs possibilités de prix. Plusieurs approches ont été développées ces dernières années.

Dans le même temps, ces dernières années, des outils d'intelligence artificielle tels que le Deep Learning, ont acquis des avancées importantes qui pourraient contribuer à mieux comprendre les variables liées aux événements de catastrophes naturelles. Il serait donc intéressant d'étudier des possibilités de leur application aux obligations catastrophes.

2. Objectif de l'étude :

Dans cette étude, il sera proposé deux approches de valorisation du prix d'une obligation catastrophe. Dans un premier temps, une approche de modélisation par une loi de probabilité sera effectuée. Ensuite, une étude sera effectuée sur une modélisation avec un modèle de réseaux de neurones séquentiel.

3. Modélisations :

(i). Modèle de valorisation des CAT Bond :

Dans ce mémoire, le modèle utilisé pour la valorisation des CAT Bond, est un modèle à temps discret dans un marché incomplet. Ce modèle a été proposé par Cox et Pedersen [2000].

$$V_{CAT}(0) = E^{\mathbb{Q}} \left[\sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1})} * F_{CAT}(k) \right]$$

Dans le cas où les différents flux du cat bond ne dépendraient que des risques catastrophes, la formule du prix pourrait être simplifiée par l'expression donnée ci-dessous par Cox et Pedersen [2000] dans leur étude et démontrée dans ce mémoire :

$$V_{CAT}(0) = \sum_{k=1}^T E^{\mathbb{Q}} \left[\frac{1}{\prod_{j=1}^k (1 + r_{j-1})} \right] E^P [F_{CAT}(k)]$$

On notera $F_{CAT}(X_k) = F_{CAT}(k)$ dans le cadre de cette étude à partir de cette hypothèse.

Ce qui permettra d'isoler les termes des risques liés aux variables financières, qui seront modélisées dans un environnement risque neutre Q , et les variables liées aux risques catastrophes qui seront modélisées en utilisant une approche de modélisation par une loi de probabilité ou une approche de modélisation par un réseau de neurone séquentiel.

On notera

La valeur du prix du CAT Bond sera obtenue en utilisant des valeurs approchées :

- ✓ Dans l'approche de modélisation par une loi de probabilité, on pourrait estimer le prix du CAT Bond en utilisant la loi forte des grands nombres sous l'hypothèse d'existence de cette espérance, les variables X_k étant supposées i.i.d, les variables $r(k)$ étant supposées i.i.d, les variables X_k et $r(k)$ sont deux à deux indépendantes.

$$\text{On pose : } V_k = \frac{1}{\prod_{j=1}^k (1+r_{j-1})} F_{CAT}(X_k)$$

$$E^Q[V_k] \approx \lim_{m \rightarrow +\infty} \frac{1}{m} \sum_{i=1}^m V_k^{(i)} \text{ où } V_k^{(i)} \text{ représente le } i\text{-ème simulation de } V_k.$$

Ce qui impliquerait que :

$$V_{CAT}(0) \approx \lim_{m \rightarrow +\infty} \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1+r_{j-1}^{(i)})} F_{CAT}(X_k^{(i)})$$

Le prix sera obtenu par une méthode de Monte Carlo où $X_k^{(i)}$ et $r_k^{(i)}$ représentent les i -ème simulations des variables des risques catastrophes et des risques financiers.

- ✓ Dans le cas de l'approche Deep Learning, on utilisera l'expression provenant du théorème de l'espérance totale dans un cas discret pour estimer $E^P[F_{CAT}(X_k)]$:

$$E^P[F_{CAT}(X_k)] = \sum_{k=1}^Z \mathbb{P}(A_k^{[i]}) E^P[F_{CAT}(X_k) | A_k^{[i]}]$$

où $A_k^{[i]}$ représente l'événement « la variable X_k appartient à la i -ème tranche à la période k ».

On en déduit :

$$V_{CAT}(0) = \sum_{i=1}^Z \sum_{k=1}^T B(k) \mathbb{P}(A_k^{[i]}) E^P[F_{CAT}(X_k) | A_k^{[i]}]$$

$$\text{Avec } B(k) = E^Q \left[\frac{1}{\prod_{j=1}^k (1+r_{j-1})} \right]$$

(ii). Modélisation des risques financiers :

Avec l'hypothèse que le CAT Bond étudié est exposé à un seul risque financier, c'est-à-dire, le taux d'intérêt, le risque financier va être modélisé en utilisant un modèle stochastique de taux tel que le modèle de Vasicek. Ce choix a été fait dans une démarche de simplification. Après la calibration des paramètres du modèle Vasicek, une simulation de 100 000 trajectoires de 3 années sera effectuée dans le but de valorisation du prix du CAT Bond.

(iii). Modélisation des variables des risques catastrophes :

Approche de modélisation par une loi GEV

La variable des risques catastrophes qui sera la vitesse maximale annuelle du vent , va être modélisée en utilisant une loi GEV. Cette modélisation est effectuée après analyse de la stationnarité des données d’observations et la calibration des paramètres de la loi GEV. À la suite de cette modélisation, une simulation de 100 000 trajectoires 3 années de loi GEV, est réalisée dans le but d’estimer le prix avec la méthode de Monte Carlo.

Approche de modélisation par un modèle séquentiel LSTM

Dans cette approche, deux types de modélisations vont être effectués : un modèle LSTM de classification et un modèle de régression.

Après la construction de la base de données, un modèle LSTM de type séquence à séquence sera entraîné sur les données d’observations afin d’estimer les maximums annuels des variables à prédire.

4. Résultats :

Une application a été faite sur un CAT Bond paramétrique avec comme paramètre de déclenchement la vitesse maximale annuelle des rafales de vent dans le sud-est des Etats-Unis. L’échelle de Saffir-Simpson a été utilisée comme indice de déclenchement des différents seuils du CAT Bond. Le premier seuil qui déclenche le CAT Bond est le niveau Ouragan de catégorie 1 c’est-à-dire une vitesse comprise entre 118 km/h et 153 km/h soit 63 kts à 82 kts.

On définira donc la structure des seuils de paiements de sinistres comme décrite ci-dessous :

$$S(k) = \begin{cases} 0 & X \leq 21 \text{ kts} \\ 0 & 21 < X \leq 33 \text{ kts} \\ 0 & 33 < X \leq 63 \text{ kts} \\ \min(0.2P ; P_k) & 63 < X \leq 82 \text{ kts} \\ \min(0.4P ; P_k) & 82 < X \leq 95 \text{ kts} \\ \min(0.6P ; P_k) & 95 < X \leq 112 \text{ kts} \\ \min(0.8P ; P_k) & 112 < X \leq 136 \text{ kts} \\ \min(P ; P_k) & X > 136 \text{ kts} \end{cases}$$

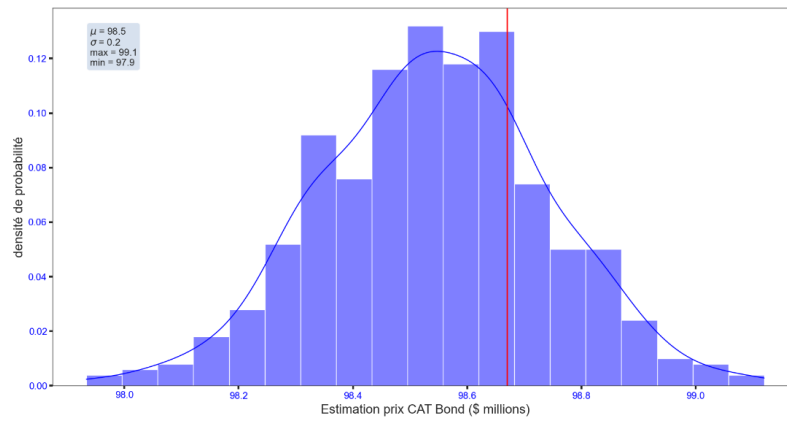
Avec P_k le principal restant à la fin de la période k.

Les risques financiers ont été modélisés avec le taux US 1an observé entre 2017 et 2019.

Pour un principal de 100 \$millions, on trouve :

Cas modélisation par une loi de probabilité : $V_{CAT}(0) = 98.67 \text{ \$ millions}$

En réalisant une centaine d’estimations de prix, on constate que les prix simulés varient entre 97.9 \$millions et 99.1 \$millions. La moyenne des prix est de 98.5 \$millions. La ligne rouge de la figure ci-dessous représente le prix $V_{CAT}(0)$



Cas Deep Learning :

Il a été étudié deux types de modélisations LSTM : un modèle classification et un modèle de régression.

- Dans le modèle de régression LSTM, Le prix du CAT Bond estimé varie entre 70 \$millions (cas où il y a deux déclenchements avec l’ajout du RMSE - racine carré de l’erreur quadratique moyenne - sur la sortie du modèle et 112.51 \$ millions (résultats du modèle).

		Résultats modèles	Cas an 1 + RMSE Test	Cas an 2 + RMSE Test	Cas an 3 + RMSE Test	Cas + RMSE sur toutes les années
LSTM - 3y (kts)	t+1	57.11	64.21	57.11	57.11	64.21
	t+2	51.38	51.38	58.68	51.38	58.68
	t+3	55.94	55.94	55.94	64.34	64.34
Flux coupons CAT Bond non actualisés (\$ millions)	t+1	6.50	5.20	6.50	6.50	5.20
	t+2	6.50	5.20	6.50	6.50	5.20
	t+3	6.50	5.20	6.50	5.20	3.90
	Principal	100.00	80.00	100.00	80.00	60.00
Prix CAT Bond (\$ millions)		112.51	90.01	112.51	92.53	70.03

- Dans le modèle de classification, les sorties du modèle sont les suivantes :

		Probabilités					
		Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6
LSTM - 3y	t+1	0.0009%	0.0032%	89.2923%	10.1607%	0.4848%	0.0581%
	t+2	0.0032%	0.0049%	87.6166%	12.0323%	0.2513%	0.0863%
	t+3	0.0026%	0.0085%	88.0841%	11.5603%	0.3220%	0.0225%

Les flux probables du CAT Bond issus du modèle sont donnés ci-dessous.

	Flux an 1	Flux an 2	Flux an 3	Principal
Flux probables CAT Bond non actualisés	6.363	6.203	6.110	93.999

Le prix du CAT Bond sera en appliquant la formule du modèle Deep Learning et des simulations des taux d’intérêts. On trouve $V_{CAT}(0) = 106.09$ \$ millions.

Cette étude bien qu’ayant des limites, a permis de modéliser le prix d’une obligation catastrophe en suivant deux approches de modélisation.

1. CONTEXT

In 2019, according to the Swiss Re Institute (Sigma n°2/2020, page 29), the economic losses related to natural catastrophe are estimated at 146 \$bn of which 86 \$bn were not covered by an insurance or protection. The uninsured losses represent up to 59% of the total economic losses due to CAT NAT event.

The size of uninsured losses shows how important it is for the insurance industry or countries seeking for protection against the CAT NAT economic losses, to find an alternative for their protection or risk transfer. In order to tackle this situation, countries (through the World Bank MultiCat program) and insurers have gained interest to the CAT Bond.

The CAT Bond is known as a high yield bond. Even though the CAT Bond has interesting yield compared to the classical bond, the investors still have some doubts on the product mainly due to the complexity of the product itself. In fact, it's complicated to put in place a hedge strategy to cover the risk taken in investing in the CAT bond by the traditional financial product. In addition to that, the valuation framework required an incomplete market framework which means the risk neutral environment is not unique meaning the price is also not unique. In the recent years, many valuation frameworks have been proposed.

The recent development of the artificial intelligence technique may contribute to a better understanding of the risk related to Cat Nat. This led to the interest on studying how this technique could be applied to the CAT Bond pricing.

2. Purpose of this paper

This paper aims at pricing a CAT Bond using two valuation approaches. In the first approach, the CAT Bond price will be modelled using a probability distribution approach. And in the second approach, the valuation will be based on a Deep Learning approach.

3. Modelling

(i). Valuation framework:

In this paper, the valuation framework will be a discrete time framework in an incomplete market. The model used in this paper is based on the approach developed by Cox and Pedersen [2000]. In this framework, the price is defined as follows:

$$V_{CAT}(0) = E^Q \left[\sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1})} * F_{CAT}(k) \right]$$

In case of CAT Bond cash flow depending only on variables related to catastrophe risks, the formula could be simplified as per below (proof in the section 3 Modelling):

$$V_{CAT}(0) = \sum_{k=1}^T E^Q \left[\frac{1}{\prod_{j=1}^k (1 + r_{j-1})} \right] E^P [F_{CAT}(k)]$$

We will assume $F_{CAT}(X_k) = F_{CAT}(k)$ as the implication of the previous hypothesis.

This new formula help isolating the financial risk variables from the catastrophe risks variables. The financial risk will be modelled in the risk neutral environment Q while the catastrophe variable could be modelled using the historical probability environment.

EXECUTIVE SUMMARY

The CAT Bond price expression could be approximated using different approximations approaches

- ✓ In the distribution approach, the strong law of the large numbers, under the assumption of the existence of the expectation, provides an estimate of the expectation as follows:

$$\text{Let note: } V_k = \frac{1}{\prod_{j=1}^k (1+r_{j-1})} F_{CAT}(X_k)$$

$$E^Q[V_k] \approx \lim_{m \rightarrow +\infty} \frac{1}{m} \sum_{i=1}^m V_k^{(i)}$$

where $V_k^{(i)}$ represents the simulation number i of the variable V_k .

It implies :

$$V_{CAT}(0) = \lim_{m \rightarrow +\infty} \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1+r_{j-1})} F_{CAT}(X_k^{(i)})$$

The CAT Bond price will be obtained using a Monte Carlo simulation approach.

- ✓ In the Deep Learning approach, the expectation could be computed using the expression of the total expectation theorem in a discrete framework:

$$E^{\mathbb{P}}[F_{CAT}(X_k)] = \sum_{i=1}^Z \mathbb{P}(A_k^{[i]}) E^{\mathbb{P}}[F_{CAT}(X_k) | A_k^{[i]}]$$

It implies:

$$V_{CAT}(0) = \sum_{i=1}^Z \sum_{k=1}^T B(k) \mathbb{P}(A_k^{[i]}) E^{\mathbb{P}}[F_{CAT}(X_k) | A_k^{[i]}]$$

$$\text{Where } B(k) = E^Q \left[\frac{1}{\prod_{j=1}^k (1+r_{j-1})} \right]$$

(ii). Financial risk modeling:

It has been considered an assumption of the financial risk is exposed only to interest rate risk in this study. The interest rate risk variable will be modelled using a stochastic interest model namely the Vasicek model. This choice is done for simplification purpose.

After calibrating the parameters of the Vasicek model, a simulation of 100 000 path of 3 years will be in order to evaluate the CAT Bond price.

(iii). Catastrophe risk variables modeling:

Modelling using a GEV (Generalized Extremum Value):

The variable of the risk related to catastrophe event will be modelled using a GEV as the variable will be the yearly maximum of the wind gust speed. This modelling will be done after performing a data analysis and calibrating the parameters of the GEV with the observations data. Following the modelling, 100 000 GEV simulations of 3 years will be run in the context of pricing with the Monte Carlo technique.

Modelling using Deep Learning LSTM approach:

In this Deep Learning approach, two types of Machine Learning approaches will be used: a classification model and a regression model.

After building the database based on the observations variables, a sequence-to-sequence LSTM model will be trained on observations data in order to predict the yearly maximum of the wing gust speed.

4. Application results :

An application has been done on a parametric CAT Bond with a trigger of the yearly wind gust in the South-East of USA. The Saffir-Simpson scale will be used as the index of the trigger parameter. The trigger threshold is set such as the first trigger corresponds to the Hurricane Category 1 on the Saffir-Simpson scale meaning a wind speed between 118 km/h et 153 km/h (63 kts to 82 kts).

The CAT Bond claims payments are defined as per below:

$$S(k) = \begin{cases} 0 & X \leq 21 \text{ kts} \\ 0 & 21 < X \leq 33 \text{ kts} \\ 0 & 33 < X \leq 63 \text{ kts} \\ \min(0.2P ; P_k) & 63 < X \leq 82 \text{ kts} \\ \min(0.4P ; P_k) & 82 < X \leq 95 \text{ kts} \\ \min(0.6P ; P_k) & 95 < X \leq 112 \text{ kts} \\ \min(0.8P ; P_k) & 112 < X \leq 136 \text{ kts} \\ \min(P ; P_k) & X > 136 \text{ kts} \end{cases}$$

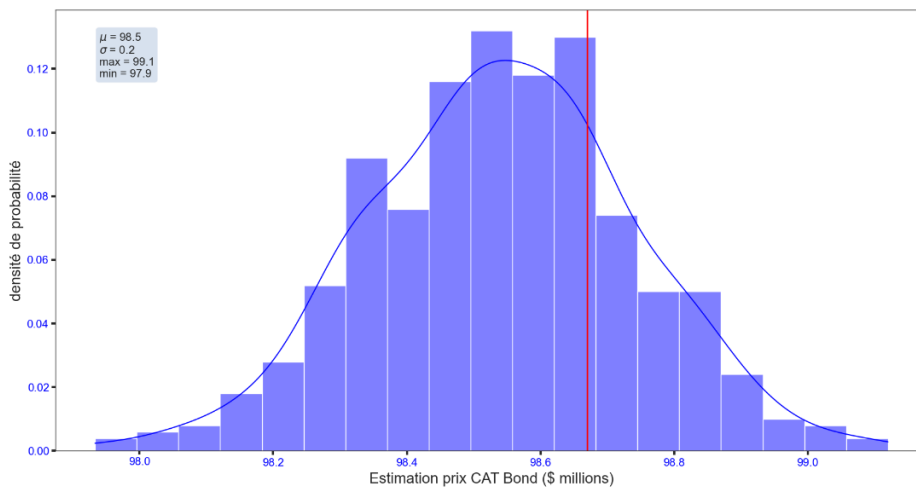
with P_k is the remaining principal at the end of period k.

The financial risk is modelled using the US 1an rate between 2017 and 2019.

For a principal of 100 \$millions, we obtained:

GEV approach: $V_{CAT}(0) = 98.67$ \$ millions

After running around 1000 price estimation process, the estimated price varies between 97.9 \$million and 99.1 \$millions. The red line in the graph below corresponds to $V_{CAT}(0)$



EXECUTIVE SUMMARY

Deep Learning approach:

- In the LSTM regression approach, the CAT Bond price varies between 70 \$millions (case with the addition of the average estimation error to the model output) and 112.51 \$ millions (case of the model output where no CAT Bond is triggered).

		Résultats modèles	Cas an 1 + RMSE Test	Cas an 2 + RMSE Test	Cas an 3 + RMSE Test	Cas + RMSE sur toutes les années
LSTM - 3y (kts)	t+1	57.11	64.21	57.11	57.11	64.21
	t+2	51.38	51.38	58.68	51.38	58.68
	t+3	55.94	55.94	55.94	64.34	64.34
Flux coupons CAT Bond non actualisés (\$ millions)	t+1	6.50	5.20	6.50	6.50	5.20
	t+2	6.50	5.20	6.50	6.50	5.20
	t+3	6.50	5.20	6.50	5.20	3.90
	Principal	100.00	80.00	100.00	80.00	60.00
Prix CAT Bond (\$ millions)		112.51	90.01	112.51	92.53	70.03

- In the LSTM classification approach, the model output

		Probabilités					
		Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6
LSTM - 3y	t+1	0.0009%	0.0032%	89.2923%	10.1607%	0.4848%	0.0581%
	t+2	0.0032%	0.0049%	87.6166%	12.0323%	0.2513%	0.0863%
	t+3	0.0026%	0.0085%	88.0841%	11.5603%	0.3220%	0.0225%

Based on the model output, the undiscounted cash flow per year is given by the table below:

	Flux an 1	Flux an 2	Flux an 3	Principal
Flux probables CAT Bond non actualisés	6.363	6.203	6.110	93.999

The price is computed based on the Deep Learning formula and the different simulated 100 000 paths of rates estimated in the financial modelling section and we get $V_{CAT}(0) = 106.09$ \$ millions.

Even with the limit related to this study, it has been proposed two approaches of pricing a CAT Bond: One based on distribution approach and the other one on Deep Learning approach.

LISTES DES TABLES	3
LISTES DES FIGURES	4
LISTE DES ABREVIATIONS	5
INTRODUCTION	6
PARTIE 1. GENERALITES SUR LES OBLIGATIONS CATASTROPHES.....	8
1.1. REASSURANCE TRADITIONNELLE.....	8
1.1.1. <i>Traité de réassurance proportionnelle</i>	8
1.1.2. <i>Traité de réassurance non proportionnelle</i>	9
1.2. COASSURANCE	9
1.3. REASSURANCE ALTERNATIVE OU ART (« ALTERNATIVE RISK TRANSFER »)	10
1.3.1. <i>Différents types des ART 'Alternative Risk Transfer'</i>	11
1.3.2. <i>Evolution du marché de réassurance alternative</i>	12
1.4. OBLIGATIONS CATASTROPHES OU « CAT BOND »	13
1.4.1. <i>Fonctionnement des CAT Bond</i>	13
1.4.2. <i>Différents types d'obligations catastrophes</i>	15
1.4.3. <i>CAT Bond par type de déclencheur</i>	17
1.5. LITTERATURE SUR LES MODELES EXISTANTS DE PRICING DES CAT BOND.....	18
1.5.1. <i>Modélisation par double processus stochastique de Poisson</i>	18
1.5.2. <i>Approche théorie d'équilibre des prix dans un marché incomplet</i>	20
PARTIE 2. GENERALITES SUR LE DEEP LEARNING ET LES RESEAUX DE NEURONES	23
2.1. DESCRIPTION DE L'APPROCHE RESEAU DE NEURONES	23
2.1.1. <i>Structure d'un réseau de neurones</i>	23
2.1.2. <i>Fonctions d'activation ou fonction de transfert</i>	25
2.1.3. <i>Fonctions d'objectif ou fonctions de coûts</i>	26
2.1.4. <i>Hyperparamètres d'un modèle de réseau de neurones</i>	27
2.1.5. <i>Optimisation du modèle d'estimation</i>	28
2.2. LES PRINCIPAUX TYPES DE RESEAUX DE NEURONES	29
2.2.1. <i>Le réseau de neurones perceptron (MLP)</i>	29
2.2.2. <i>Les modèles de réseau de neurones récurrents ou séquentiels</i>	30
2.2.3. <i>Réseau de neurones récurrents à mémoire long et court-termes</i>	31
2.2.4. <i>Réseau de neurones récurrents à portes</i>	33
PARTIE 3. MODELISATION CAT BOND	35
3.1. HYPOTHESES DE MODELISATION	35
3.1.1. <i>Hypothèse sur le modèle du marché</i>	35
3.1.2. <i>Autres hypothèses de modélisations</i>	35
3.1.3. <i>Le modèle de valorisation</i>	36
3.1.4. <i>Les hypothèses sur la structure du contrat de CAT Bond</i>	39
3.1.5. <i>Hypothèses de modélisation des risques financiers</i>	40
3.1.6. <i>Hypothèses de modélisation des risques catastrophes</i>	41
3.1.7. <i>Hypothèses sur les coupons et spread</i>	44
3.2. MODELISATION DES RISQUES FINANCIERS	45
3.2.1. <i>Méthodes de calibration des paramètres du modèle de taux d'intérêt</i>	45
3.2.2. <i>Simulations de trajectoires de taux d'intérêt</i>	48
3.3. MODELISATION CAT BOND : APPROCHE DE MODELISATION DES RISQUES CATASTROPHES PAR UNE LOI	49
3.3.1. <i>Méthodes de calibration des paramètres d'une loi GEV</i>	49
3.3.2. <i>Méthode de simulation de la GEV</i>	50
3.3.3. <i>Valorisation CAT Bond</i>	50

Tables des matières

3.5.	MODELISATION CAT BOND : APPROCHE DE MODELISATION DES VARIABLES DES RISQUES CATASTROPHES PAR UNE APPROCHE DEEP LEARNING.....	51
3.5.1.	<i>Description du modèle</i>	51
3.5.2.	<i>Construction de la base de données d'apprentissage</i>	52
3.5.3.	<i>Modèles de Deep Learning</i>	53
3.5.4.	<i>Valorisation du CAT Bond</i>	54
PARTIE 4.	APPLICATION A UN CAT BOND A DECLENCHEUR PARAMETRIQUE.....	55
4.1.	LES CARACTERISTIQUES DU CAT BOND ETUDIE	55
4.2.	APPLICATION DE LA MODELISATION DES RISQUES FINANCIERS	56
4.2.1.	<i>Données d'observations des taux d'intérêts</i>	56
4.2.2.	<i>Calibration du modèle de Vasicek sur les données US 1an</i>	58
4.2.3.	<i>Simulation de trajectoires de taux d'intérêts</i>	58
4.3.	DONNEES D'OBSERVATIONS DES VARIABLES CATASTROPHES	60
4.3.1.	<i>Description des variables de la base de données d'observations</i>	60
4.3.2.	<i>Analyse des données d'observations</i>	61
4.4.	APPLICATION DE L'APPROCHE DE MODELISATION PAR UNE LOI GEV	65
4.4.1.	<i>Analyse des vitesses maximales annuelles</i>	65
4.4.2.	<i>Test de stationnarité des observations</i>	66
4.4.3.	<i>Calibration des paramètres GEV</i>	67
4.4.4.	<i>Simulation de trajectoires de 3 années</i>	69
4.4.5.	<i>Evaluation du prix du CAT Bond</i>	69
4.4.6.	<i>Sensibilité du prix aux variations de seuil de déclenchement</i>	71
4.5.	APPLICATION DE L'APPROCHE DE MODELISATION DEEP LEARNING	73
4.5.1.	<i>Construction de la base de données des modèles</i>	73
4.5.2.	<i>Description des modèles</i>	74
4.5.3.	<i>Evaluation des modèles</i>	75
4.5.4.	<i>Estimation du prix du CAT Bond</i>	78
4.6.	ANALYSE DES RESULTATS	80
LIMITES ET CONTRAINTES DE L'ETUDE		81
CONCLUSION.....		82
ANNEXES		84
BIBLIOGRAPHIE		89

Liste des tables

Table 1 - Pertes économiques par région 2019.....	6
Table 2 - Echelle de Saffir-Simpson	55
Table 3- Caractéristiques du CAT Bond étudié.....	56
Table 4 - Résultats de la simulation de la trajectoire du taux d'intérêts	59
Table 5 - Variables des données d'observations	61
Table 6 - Vitesse maximale annuelle des rafales de vent (source GSOD)	65
Table 7 - Analyse descriptive de la vitesse maximale annuelle.....	65
Table 8 - Test de Dickey Fuller augmenté	66
Table 9 - Distribution de la simulation GEV sur 3 années	69
Table 10 - Statistique descriptive des flux financiers simulés.....	70
Table 11 - Sensibilité du prix du CAT Bond par rapport au seuil de déclenchement.....	72
Table 12-Résultats modèles LSTM de régression.....	76
Table 13 - Résultats des modèles de classification	77
Table 14 - Prix CAT Bond - LSTM régression.....	78

Figure 1 - Evolution des pertes économiques entre 1970 - 2019	6
Figure 2 - Evolution titres réassurance alternatives entre 1997 – 2020	10
Figure 3- Emission des titres de réassurance alternatives	12
Figure 4 - Structure d'un CAT Bond	15
Figure 5 - Répartition des émissions des CAT Bond par trigger	17
Figure 6 - Evolution des émissions par type de trigger	17
Figure 7 - Exemple de structure d'un réseau de neurone.....	23
Figure 8 - Transfert d'informations entre des neurones de deux couches (i) et (i+1)	24
Figure 9 - Structure d'un RNN	30
Figure 10 - Modèles LSTM (source : Wikipédia).....	33
Figure 11- Modèle GRU (source : Wikipédia).....	34
Figure 12 - Evolution des coupons et spread	44
Figure 13 - Modèles LSTM Seq to Seq	53
Figure 14 - Evolution journalière du taux US 1 an 2010-2020	56
Figure 15 - Relation entre deux observations journalières successives du taux US 1an entre 2017 et 2019.....	57
Figure 16 – Extrait de 100 trajectoires générées par la simulation de Vasicek	59
Figure 17 - Corrélation entre les variables des données d'observations	62
Figure 18 - Relation entre les variables.....	62
Figure 19 - Evolution des variables entre 2000 - 2014.....	64
Figure 20 - Vitesses maximales entre 1973 et 2019.....	66
Figure 21 – BoxPlot des vitesses maximales annuelles.....	66
Figure 22 – Alignement des données d'observations sur le modèle calibré	68
Figure 23 - Distribution des 1000 simulations de prix CAT Bond	70
Figure 24 - Passage en base 365 jours.....	73
Figure 25-LSTM Fenêtre d'historiques (1) et de prédictions (2)	74
Figure 26 - Résultats des modèles LSTM.....	77

Liste des abréviations

AOA : Absence d'opportunité d'arbitrage
ART : transfert de risque alternatif
CAT Bond : Obligations catastrophes
ConvLSTM : LSTM Convolutif
CNN : Réseau de neurones convolutif
 ddl_A : nombre de degré de liberté de A
EDS : Equation différentielle stochastique
EMV : Estimation par maximum de vraisemblance
IC : intervalle de confiance
GRU : Réseau de neurones récurrents à portes ou « Gated Recurrent Unit»
kt : vitesse en nœuds
LSTM : Réseau de neurones récurrents à mémoire long et court-termes ou “Long short term memory”
MAE : La moyenne des erreurs absolues
MCO : Moindre Carrés ordinaires
MC : Monte Carlo
MSE : La moyenne des erreurs quadratiques
RNN: Réseau de neurones récurrents
RMSE : La racine carrée des erreurs quadratiques moyennes
SCE : Variabilité expliquée par les paramètres du modèle dans le modèle de régression
SCT : Variabilité totale du modèle dans le modèle de régression
SCR : Variabilité expliquée par les résidus dans le modèle de régression
TVE : Théorie des valeurs extrêmes

En 2019, selon les données du Swiss Re Institute sur la Figure 1 et la table 1 ci-dessous (Sigma n° 2/2020, page 29), les pertes économiques dues aux catastrophes naturelles sont estimées à 146 milliards \$ dont environ 86 milliards \$ de pertes non couvertes par une assurance soit environ 59% du total des pertes économiques en 2019. De plus, les pertes moyennes annuelles non couvertes par une assurance sur ces 10 dernières années sont estimées à environ 137 milliards USD. Le continent asiatique représente, à lui seul, 45% de la perte totale en 2019 soit 65,9 milliards \$ avec uniquement 28% couvertes par une assurance.

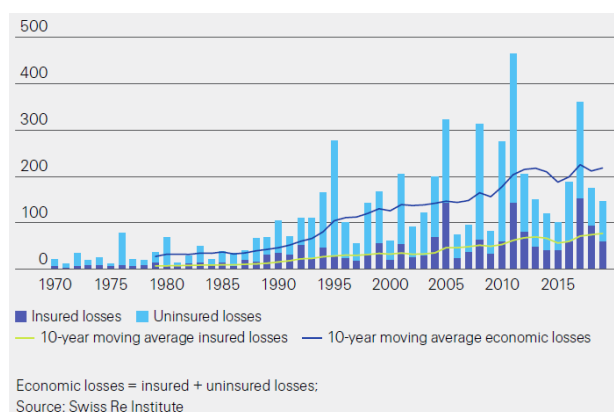


Figure 1 - Evolution des pertes économiques entre 1970 - 2019

Region	Pertes assurées (en USD milliard)	%	Pertes économiques totales (en USD milliard)	%
North America	27,2	45,60%	44,7	30,60%
Latin America & Caribbean	5,2	8,70%	11,9	8,20%
Europe	5,4	9,00%	13,6	9,30%
Africa	0,8	1,40%	5,3	3,60%
Asia	18,3	30,60%	65,9	45,10%
Oceania/Australia	2,5	4,10%	4,1	2,80%
Seas / Space	0,4	0,70%	0,4	0,30%
World	60	100,00%	146	100,00%

Source : Swiss Re Institute

Table 1 - Pertes économiques par région 2019

Face aux pertes colossales que peuvent engendrer les catastrophes naturelles, la recherche d'une protection financière contre ces risques est devenue un sujet majeur pour la stabilité économique des pays émergents et en voie de développement.

La Banque Mondiale, dans le but d'aider les pays émergents et en voie de développement à s'assurer contre les risques catastrophes, a lancé en 2009 le programme MultiCat¹. Le projet Multicat est un programme d'élaboration d'obligations catastrophes pour soutenir l'économie de certains pays émergents et en voie de développement.

En plus de la difficulté à trouver un financement pour se couvrir contre ces risques, les émetteurs des obligations catastrophes font face à une autre difficulté sur le marché financier. En effet, pour les investisseurs, la difficulté de mettre en place des stratégies de couvertures des obligations catastrophes avec des produits financiers traditionnels, constituent un frein à l'entrée sur le marché de CAT Bond pour certains investisseurs. Cela est principalement dû à la problématique complexe de la réplique des risques liés aux événements catastrophes avec des instruments financiers traditionnels.

De plus, du fait de cette structure des obligations, la modélisation d'un prix nécessite une approche de modélisation dans un marché incomplet comme évoqué par Cox et Pedersen (2000). Ce qui suggère qu'il existe plusieurs environnements risque neutre donc plusieurs possibilités de prix.

Depuis l'avènement des obligations catastrophes, plusieurs approches de valorisation des obligations catastrophes ont été proposées dans différentes études.

¹ <https://www.banquemonde.org/fr/news/press-release/2009/10/19/world-bank-launches-multicat-program>

Introduction

Ces dernières années, l'intelligence artificielle a eu des avancées considérables. En partant de l'application dans le domaine de la reconnaissance faciale à la conduite assistée, les avancées des Machine Learning ont contribué à plusieurs avancées technologiques. De nos jours, plusieurs domaines de l'assurance utilisent déjà ses avancés pour améliorer leurs outils de projection et d'identification des risques. C'est dans ce contexte que nous allons nous intéresser à l'application du Deep Learning pour l'estimation des risques liés aux catastrophes naturelles dans le but d'élaborer une approche de valorisation des obligations catastrophes ou « CAT Bond » en plus d'une modélisation par une loi de probabilité.

Afin de pouvoir répondre à cette problématique, cette étude sera structurée en quatre grandes parties. Dans la première partie, nous aborderons les généralités sur les mécanismes de transferts de risques en assurance.

Dans la seconde partie, une présentation générale des méthodes d'apprentissage approfondis sera effectuée.

Dans la troisième section, nous proposerons des approches de modélisation des obligations catastrophes à déclenchement paramétrique.

Enfin, une application de ces modélisations sera effectuée sur les risques de tempêtes dans le sud-est des Etats-Unis.

Partie 1. Généralités sur les obligations catastrophes

Les obligations catastrophes sont une forme de transfert de risque assurantiel vers les marchés financiers. Dans cette partie, nous allons premièrement faire un rappel des différentes approches de transfert de risque. Ensuite nous ferons un focus sur les obligations catastrophes : on étudiera leur mode de fonctionnement, l'évolution du marché, la littérature sur la modélisation du prix des obligations catastrophes.

Dans les sections suivantes, les principales approches de transfert de risques abordées seront :

- La réassurance traditionnelle
- La coassurance
- La réassurance alternative

1.1. Réassurance traditionnelle

La réassurance traditionnelle fait principalement intervenir deux acteurs :

- Un assureur cherchant à transférer une partie de ses risques vers une autre entité
- Un réassureur disposé à prendre en charge une partie des risques de l'assureur

Dans la suite de cette partie, les principaux types de réassurance traditionnelle seront présentés. On distingue deux principaux groupes qui sont le traité de réassurance proportionnelle et le traité de réassurance non-proportionnelle.

1.1.1. Traité de réassurance proportionnelle

Dans un traité de réassurance proportionnelle, l'assureur et le réassureur définissent la proportion de sinistres pris en charge par l'assureur et le réassureur dans le cas de survenance d'un élément déclencheur de paiements de sinistres.

Dans ce type de traité, l'assureur cède la proportion de primes relatives à la part de risque pris par le réassureur. On en distingue plusieurs formes :

- **Traité QP – Traités en quote-part ou « quota-Share »**

Dans un traité quote-part, le réassureur et l'assureur se partagent un pourcentage de primes et de sinistres. Le traité définit la proportion pris en charge par le réassureur. Par exemple, dans un traité QP 30%, l'assureur cédera 30% des primes au réassureur contre une prise en charge de 30% des sinistres par le réassureur.

- **Traité XP – Traités en excédent de capitaux (ou excédent de pleins)**

Dans un traité de type XP, contrairement au modèle du traité QP, il est défini un niveau de capital assuré à partir duquel la réassurance proportionnelle va s'appliquer.

Ainsi pour chaque contrat d'assurance inclus dans le traité, si le capital assuré dépasse le seuil de rétention défini dans le traité, le réassureur prendra en charge la proportion du ratio défini dans le contrat en fonction de l'excédent du capital assuré. Le ratio se calcule comme suit :

$$ratio_{contrat\ i} = \frac{Capital\ assuré_{contrat\ i} - seuil\ de\ rétention}{Capital\ assuré_{contrat\ i}}$$

En contrepartie, l'assureur cédera au réassureur la proportion de prime correspondant à ce même ratio sur le contrat dépassant le seuil de rétention.

1.1.2. Traité de réassurance non proportionnelle

- Traité XS ou XL – Traité en excédent de sinistre par risque

Le fonctionnement d'un traité XS est un peu similaire à l'approche de franchise d'un contrat d'assurance avec l'assureur dans le rôle de l'assuré et le réassureur dans celui de l'assureur. Dans ce type de traité, il est défini une priorité et une portée.

Le contrat est généralement identifié avec la notation '*portée XS priorité*'.

Les engagements du réassureur commencent à partir des sinistres atteignant la priorité et jusqu'à la portée. La priorité peut être définie comme le seuil à partir duquel le réassureur interviendra dans la prise en charge des sinistres et la portée comme la limite de ses engagements du réassureur c'est-à-dire la charge maximale qu'il serait amenée à prendre en charge dans le cadre du contrat.

Le contrat est défini par type de risque, ce qui implique que plusieurs sinistres sur un même risque pourront potentiellement être couverts par le contrat.

- Traité Cat XS ou Cat XL – Traité en excédent de sinistre par événement

Le traité Cat XS est pratiquement similaire au traité XS à la seule différence que le contrat Cat XS concerne la charge totale des sinistres liés à un même événement.

Il est généralement utilisé dans le cas de la réassurance des risques extrêmes telles que des catastrophes naturelles, des crash aériens, etc.

- Traité SL – Traité en excédent de perte (Stop Loss)

Le fonctionnement du traité 'Stop Loss' est pratiquement similaire au traité XS.

Les différences avec les traités XS sont les suivantes :

- La couverture concerne un ensemble d'activités qui peuvent être les pertes totales d'une société, une branche d'activités, etc.
- La priorité et la portée peuvent être définies comme des ratio (exemple : ratio S/P)

1.2. Coassurance

La coassurance est un principe de partage de risque entre plusieurs assureurs. Dans cette approche, chaque assureur a des engagements directs avec l'assuré c'est-à-dire qu'en cas de défaut d'un assureur, les engagements des autres assureurs ne seront pas impactés même si les contrats respectifs concernent le même risque.

Cette approche de transfert de risque est différente de l'approche de réassurance dans laquelle le réassureur n'a pas d'engagement envers l'assuré. Un défaut du réassureur expose l'assureur à la prise en charge de la part des sinistres que devait couvrir le réassureur.

1.3. Réassurance alternative ou ART (« Alternative Risk Transfer »)

La réassurance alternative ou transfert de risques assurantiels vers les marchés financiers a commencé à émerger à la suite des événements catastrophes lié à l'ouragan Andrew sur la côte Sud-Est des Etats-Unis en 1992 qui a été l'un des événements catastrophes les plus destructeurs dans l'histoire des Etats-Unis. Les pertes liées à cet ouragan ont été estimées à plus de 27 milliards \$. Le marché de l'assurance des risques liés aux catastrophes naturelles se devait de trouver des ressources additionnelles pour se protéger contre ces pertes extrêmes.

Depuis les premières émissions de produits financiers liés aux événements catastrophes, le marché ne cessent de croître après un léger ralentissement entre 2007 et 2011 comme le montre la figure ci-dessous. Alors qu'en 1997, les ART ne représentaient que \$ 785 millions en termes de capital couvert, ils sont à plus de \$ 45 milliards en termes de couverture des risques catastrophes. On note aussi une émission moyenne annuelle de plus de \$ 10 milliards depuis 2016.

En 2019, ils sont évalués à plus de 40 milliards \$ en couverture des risques catastrophes, selon les chiffres du site spécialisé www.artemis.com.

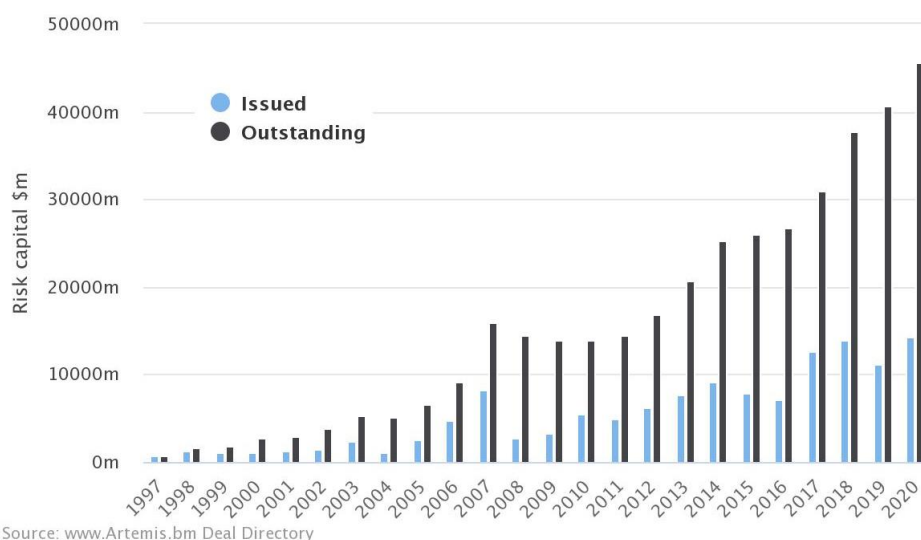


Figure 2 - Evolution titres réassurance alternatives entre 1997 – 2020

1.3.1. Différents types des ART 'Alternative Risk Transfer'

L'avènement des ART dans le marché financier s'est aussi accompagné de la création de plusieurs types de produits financiers. Les principaux produits de réassurance alternative seront décrits dans cette section.

- **Industry Loss Warranties (ILWs) ou Garantie sur la perte de l'industrie :**

Ce type de réassurance alternative fonctionne comme des options en finance. En effet, dans ce type d'actif, le sous-jacent est la perte totale de l'industrie assurance dans le secteur défini dans le contrat lors de type d'événements spécifiés par le contrat.

- **Les obligations catastrophes ou « CAT Bond » :**

Dans le cas des obligations catastrophes, l'assureur/réassureur transfère aux marchés financiers ses expositions à des risques catastrophes spécifiques à travers l'émission d'une obligation. Cette obligation est mise en place par l'intermédiaire d'un fonds commun de créances ou SPV créé spécialement pour la gestion du CAT bond. La création de cette entité a pour but de séparer les actifs de l'émetteur des actifs découlant du contrat CAT Bond. En contrepartie de paiement de coupons périodiques aux investisseurs, l'assureur va recevoir une protection contre des risques Cat Nat pour un montant maximum défini comme le principal de l'obligation. A la différence des obligations traditionnelles, le principal investi par les investisseurs est à risque du fait uniquement des pertes dues aux événements catastrophes et non du défaut de l'assureur.

Si le facteur de déclenchement de la prise en charge des sinistres par le CAT bond est atteint, le principal sera diminué des paiements effectués à l'assureur et les paiements des coupons par l'assureur seront fonction du principal restant.

A la maturité, les investisseurs recevront le montant restant du principal.

Le fonctionnement des obligations catastrophes sera plus élaboré dans la section 1.4 de ce document.

- **Les Swaps de risques catastrophes :**

Le fonctionnement d'un contrat de swap de risque catastrophe est pratiquement similaire à celui d'un swap en finance. Dans ce modèle de transfert de risques, deux (ré)assureurs se transfèrent des risques sur ces catastrophes auxquels ils sont exposés.

Les deux jambes du swap sont assujetties à des seuils de déclenchement qui peuvent se baser sur des données paramétriques liées à chaque événement.

Lors de la mise en place du swap de risques de catastrophes :

- Il n'y a pas d'échange de flux de paiements
- Les déclencheurs de chaque jambe sont définis de sorte à avoir une parité dans les estimations des pertes sur la période du contrat.

- **Les Put catastrophe « Cat-E-Puts » :**

Dans le cas d'un contrat Cat-E-Puts, l'assureur obtient contre le paiement d'une prime à l'investisseur, le droit et non l'obligation d'émettre des « Preferred Shares » à un prix prédéfini à la survenance d'événements catastrophes. Ce qui lui permettra de lever du capital à un prix supérieur aux valeurs du marché au moment de la catastrophe où ses actions pourraient potentiellement baisser. Cependant, cette forme de transfert de risque expose l'assureur à un risque de contrepartie du fait qu'il n'y ait pas de collatéral dans ce type de contrat. L'investisseur qui a acheté l'option pourrait être en défaut au moment de l'émission des actions de préférences.

- Réassurance avec collatéral :

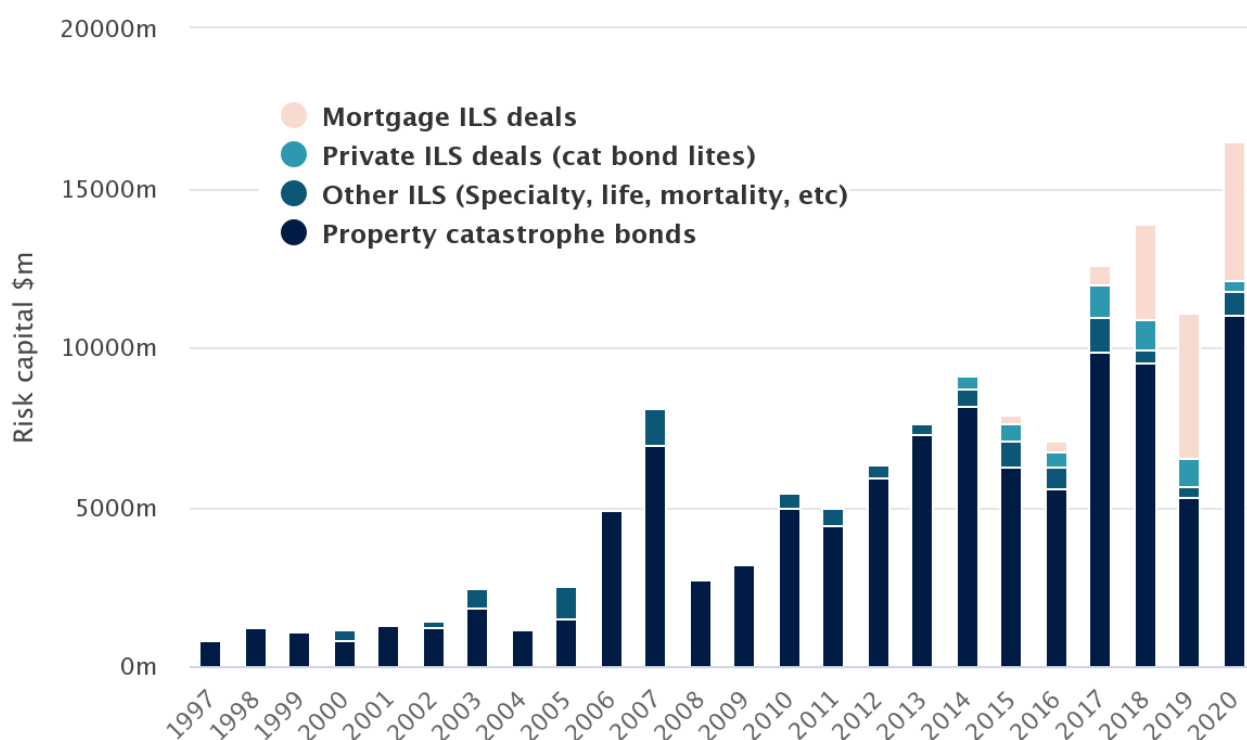
La réassurance avec collatéral est une émission de titre de créance pour couvrir un contrat de réassurance. La valeur totale des titres est limitée à la couverture maximale du contrat de réassurance.

- Le Sidecar :

Le sidecar a un fonctionnement presque similaire avec la réassurance proportionnelle traditionnelle. Le réassureur est remplacé dans cette approche par un SPV créé avec une durée limitée. Dans le cas du sidecar, un SPV est créé par un sponsor (assureur ou réassureur). Le sponsor partage les primes et sinistres avec le SPV comme dans le cas de la réassurance proportionnelle quote-part. Le capital du SPV vient des investisseurs du marché financier et des primes de réassurance versées par le sponsor. En cas de survenance de sinistres, le capital du SPV en plus des intérêts provenant des investissements du capital serviront aux paiements.

1.3.2. Evolution du marché de réassurance alternative

La réassurance alternative liée aux événements catastrophes naturelles, représente plus des trois quarts des titres émis dans le cadre d'un transfert vers le marché financier. Le graphique ci-dessous résume l'évolution des émissions de titres (source : artemis.com). Les titres liés au CAT Bond, depuis la création, constituent en moyenne plus de 75% des émissions des titres sur le marché en termes de capital.



Source: www.Artemis.bm Deal Directory

Figure 3- Emission des titres de réassurance alternatives

Après avoir revu les différents types de réassurance alternative, nous allons nous intéresser, dans les sections suivantes, à l'obligation catastrophe et à son mode de fonctionnement.

1.4. Obligations catastrophes ou « CAT Bond »

L'obligation catastrophe ou CAT Bond est une obligation qui permet à une entité cédante du risque appelé « sponsor » d'émettre des obligations sur le marché financier dont les coupons et le principal dépendront de la survenance de sinistres de type catastrophes naturelles. Dans cette partie, nous allons d'abord, décrire le principe de fonctionnement d'une obligation catastrophe. Ensuite nous décrirons les principaux types de CAT Bond. Enfin, nous allons analyser l'évolution du marché de CAT Bond et les principaux risques couverts.

1.4.1. Fonctionnement des CAT Bond

Un CAT Bond fait généralement intervenir 3 principaux acteurs :

- **Un sponsor** : C'est l'entité qui cherche à s'assurer contre des événements catastrophes. Le sponsor est généralement l'entité qui cherche à transférer un risque auquel elle est exposée aux marchés financiers. Il peut donc s'agir d'un assureur ou un réassureur principalement mais aussi d'un pays qui cherche à trouver un financement pour faire face aux pertes économiques liées à la survenance des catastrophes naturelles. Il émet une obligation catastrophe par le biais de la création d'une entité ad-hoc indépendante dédiée à la gestion du CAT bond.
- **Le SPV (Special Purpose Vehicle)**: Le SPV est la structure mise en place dans le cadre de la création du CAT Bond. C'est l'entité qui sera en charge de la gestion des différents flux financiers entre les investisseurs et le sponsor liés aux CAT Bond. Il sera l'intermédiaire qui va garantir le fait que le sponsor n'intègre pas les actifs du CAT Bond dans ses propres actifs et y avoir recours en cas d'événements non liés aux risques couverts. Pour garantir les paiements des coupons aux investisseurs, le SVP va investir le capital dans des actifs sans risque ou peu risqué. En cas de survenance d'un événement catastrophe répondant aux critères de déclenchement du CAT bond, le SPV va garantir le paiement à l'assureur des montants requis selon les termes du contrat.
- **Les investisseurs** : Ce sont les acteurs du marché financier qui acceptent le risque en investissant le capital et recevront en retour des intérêts venant des placements du principal et de primes payées par le sponsor. Dans une optique de recherche de diversification de leurs portefeuilles, les investisseurs acceptent le paiement de capitaux pour la couverture des risques liés aux événements catastrophes et ils reçoivent en contrepartie des coupons issus des CAT Bond qui sont des actifs à haut rendement comparés aux obligations ordinaires. Le capital investi est à risque car en cas d'atteinte du seuil de déclenchement du paiement, le capital investi par l'investisseur est diminué des paiements effectués au sponsor. Pour le risque pris, ils reçoivent de l'assureur un spread additionnel que l'on pourrait définir comme une prime de risque. De plus, la décorrélation entre les obligations catastrophes et les produits financiers classiques est une des principales raisons de l'intérêt des investisseurs pour ce type de produits.

Dans un contrat CAT Bond, il existe plusieurs flux financiers entre la période d'émission et la maturité du contrat. Nous pouvons diviser cette période en 3 grandes sous-périodes :

- **A l'émission :**

- Le sponsor met en place un SPV pour l'émission de titre CAT Bond
- Le SPV reçoit des investisseurs le principal, c'est-à-dire le capital couvert.
- Le SPV va investir le principal dans un actif sans risque garantissant le paiement de la part des coupons revenant des revenus générés par l'investissement du principal.
- Le sponsor reçoit la protection contre les risques couverts par le contrat du CAT Bond.

- **Entre la date d'émission et la maturité :**

Après l'émission du titre, le fonctionnement des CAT Bond va dépendre principalement de la survenance d'une catastrophe naturelle couverte par le contrat.

- S'il n'y a aucun événement catastrophe durant la période active du contrat :
L'assureur paie une prime périodique au SPV qui sera reversée aux investisseurs.
L'investisseur reçoit un coupon qui est composé des intérêts du placement du principal et de la prime payée par l'assureur.

- S'il y a des événements catastrophes :
 - Cas où les paramètres des événements catastrophes n'atteignent pas le seuil de déclenchement du contrat :

On aura les mêmes flux que dans le cas où il n'y a aucun événement catastrophe.

- Cas où le seuil de déclenchement est atteint :
L'assureur reçoit les paiements du SPV selon les termes du contrat dans la limite du principal. L'investisseur reçoit des coupons diminués en proportion des paiements déduits du principal.

Dans le cas où les paiements à l'assureur atteignent le principal, cela mettra fin au contrat avec l'assureur ne payant plus de prime et les investisseurs perdant le principal et les prochains coupons.

A la maturité :

- S'il n'y a eu aucun événement catastrophe déclenchant les paiements, les investisseurs reçoivent le principal plus le dernier coupon
- S'il y a eu des événements catastrophe :
L'investisseur reçoit le principal diminué des paiements effectués à l'assureur.

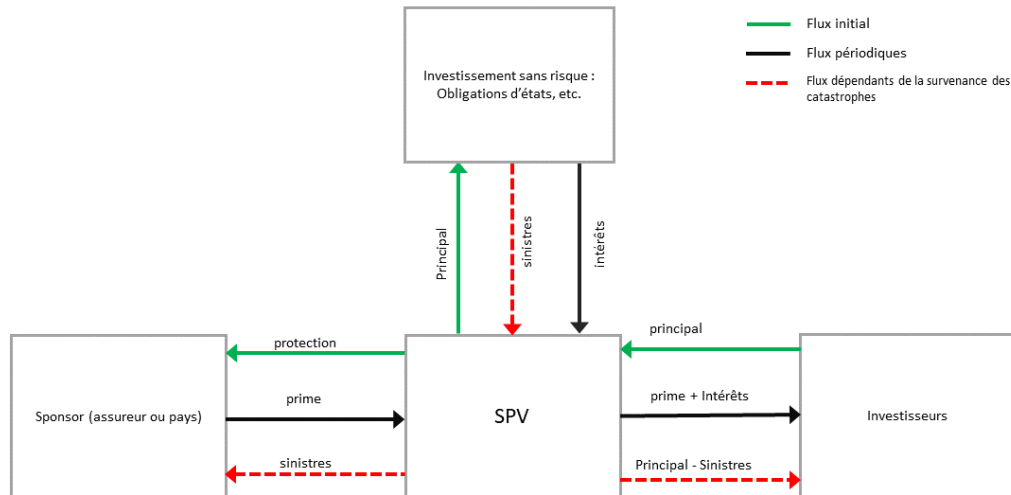


Figure 4 - Structure d'un CAT Bond

Dans la section suivante, nous décrivons les principaux types de CAT Bond

1.4.2. Différents types d'obligations catastrophes

Les CAT Bond peuvent être classés en plusieurs groupes selon la nature du déclenchement des flux dédiés à la protection de l'assureur.

- CAT Bond indemnitaire :

Dans le modèle de CAT Bond indemnitaire, le déclenchement des paiements est lié aux pertes réelles du sponsor dues aux risques inclus dans le contrat du CAT Bond.

Le principal risque de base associé à ce type de contrat est l'aléa moral, c'est-à-dire la possibilité pour le sponsor d'être plus large dans le paiement des sinistres à proximité de la zone de déclenchement du CAT bond.

Ce type de CAT Bond est celui qui diminue le risque de base pour l'assureur. En effet, la survenance d'un événement déclenchant le CAT Bond garantit, pour le sponsor, la prise en charge de la totalité des pertes enregistrées dans la limite du principal. En revanche, il est aussi celui avec le moins de transparence pour les investisseurs étant exposé à l'aléa moral.

On pourrait noter des similarités entre les 'CAT Bond' avec les traité Cat XS de la réassurance traditionnelle. Le réassureur est remplacé dans ce cas par les investisseurs.

- CAT Bond à déclenchement basé sur les pertes d'un modèle

Ce modèle de CAT Bond est pratiquement similaire au modèle indemnitaire. La différence vient du fait que les paiements ne dépendent pas des pertes réelles du sponsor mais de celles estimées par un modèle.

Le principal risque de base associé à ce type de contrat est l'écart entre les pertes réelles du sponsor et les pertes du modèle basé sur des données paramétriques de l'événement catastrophe. Le sponsor pourrait avoir éventuellement des pertes au-delà du seuil de déclenchement sans pour autant déclencher le CAT Bond. Ce qui pourrait mettre à risque la solvabilité du sponsor et ainsi générer un risque de défaut sur les différents paiements des coupons.

- CAT Bond indiciel

Dans le modèle indiciel, le facteur de déclenchement est lié à un indice assurantiel tel que, par exemple, l'indice des pertes de l'industrie dues à un événement catastrophe. Il existe plusieurs indices servant de base pour l'élaboration de ce type de CAT bond tel que le PCS Loss Index.

Le principal risque de base associé à ce type de contrat est lié à l'impact réel d'un événement sur le sponsor comparé aux restes de l'industrie. Il pourrait être éventuellement plus exposé à un risque spécifique par rapport à l'industrie sans pour autant déclencher le CAT bond.

- CAT Bond paramétrique

Dans le modèle paramétrique, le facteur de déclenchement est lié à des données paramétriques définies par le contrat. Ces facteurs peuvent être par exemple l'échelle de Richter, l'échelle de Mercalli, le niveau de pluie, une température minimale sur un nombre de jours de successifs, etc.

Dans ce type de contrat, le risque de base vient principalement du fait que l'assureur peut subir des pertes importantes sans pour autant déclencher les paiements des CAT Bond dus aux données paramétriques non atteintes.

Ce type de CAT Bond fournit le plus de transparence pour l'investisseur. En cas de survenance d'un événement catastrophe, l'investisseur peut avoir une estimation exacte de l'impact sur son portefeuille contrairement au modèle indemnitaire où les pertes peuvent être cumulées jusqu'à la fin de paiements des sinistres liés à l'événement par l'assureur.

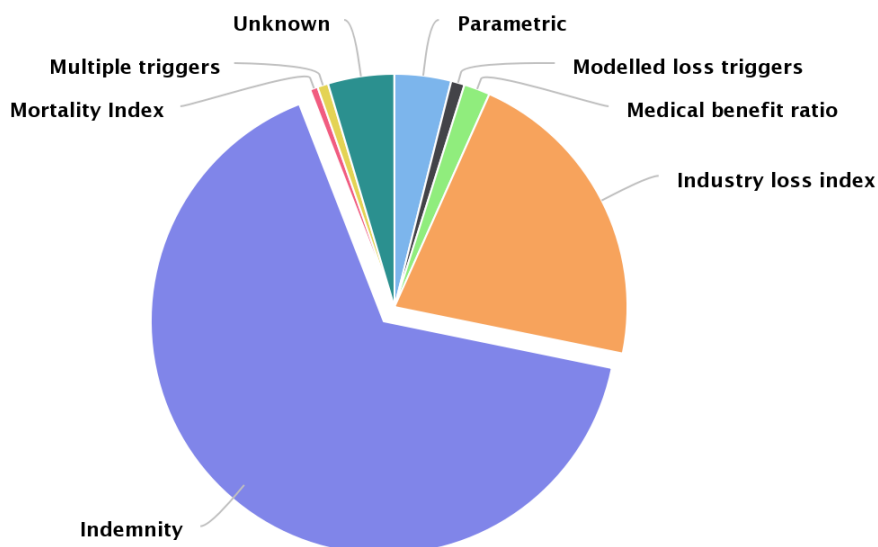
- CAT Bond hybrides

Ces CAT Bond sont le plus souvent la combinaison de deux ou plusieurs des modèles de déclenchement cités plus haut. L'avantage de ces types de CAT Bond est qu'ils peuvent limiter les principaux risques de base pour l'assureur et aussi permettre une meilleure transparence pour l'investisseur en limitant l'aléa moral avec des conditions supplémentaires qui viennent des données paramétriques.

1.4.3. CAT Bond par type de déclencheur

Les CAT Bond indemnitaires représentent plus de la moitié des obligations émises comme nous pouvons le voir sur le graphique ci-dessous. La majorité des CAT Bond émis par les assureurs sont de type indemnitaires. Le risque de base sur ce type de CAT Bond est négligeable comparé aux autres types.

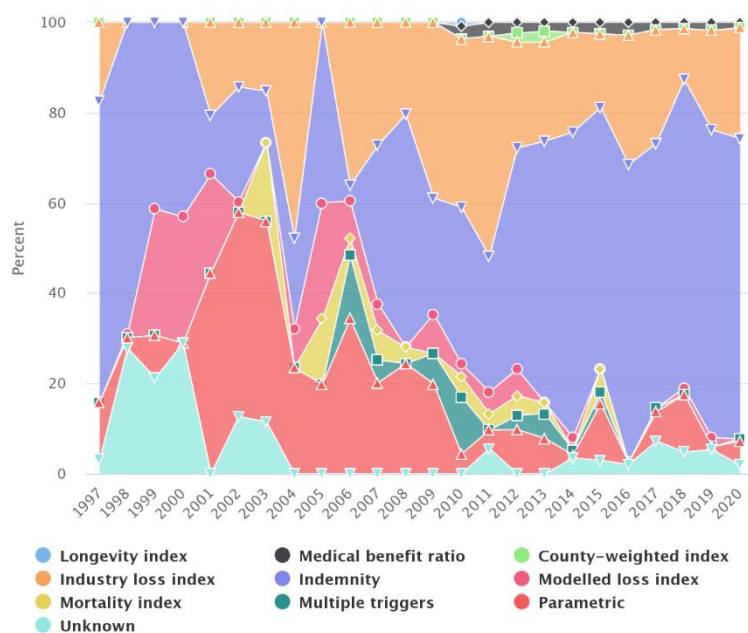
Du point de vue des investisseurs, les CAT Bond paramétriques représentent celui avec le moins d'aléa moral étant donné que les indices sont généralement publics.



Source: www.Artemis.bm Deal Directory

Figure 5 - Répartition des émissions des CAT Bond par trigger

La figure 5 présente l'évolution des émissions de titres par trigger. Les émissions des CAT Bond paramétriques ont ralenti à partir de 2004 qui coïncident aussi avec la période de croissance de part de marché des obligations indemnitaires. On note aussi un ralentissement des émissions d'obligations indicielles. Cette tendance est expliquée par l'attrait des obligations indemnitaires pour les (ré)assureurs.



Source: www.Artemis.bm Deal Directory

Figure 6 - Evolution des émissions par type de trigger

1.5. Littérature sur les modèles existants de Pricing des CAT Bond

Plusieurs études ont été menées sur le Pricing des CAT Bond depuis les premiers lancements de ces produits. Nous allons présenter deux des études effectuées dans cette section. Le premier modèle est un modèle continu par un processus de Cox et le second modèle est un modèle discret.

1.5.1. Modélisation par double processus stochastique de Poisson

Cette approche de valorisation des CAT Bond a été proposée par Baryschnikov, Mayo et Taylor [1998]. Dans cette approche, ils ont proposé une modélisation par un processus de Cox² ou processus de poisson doublement stochastique pour estimer le prix du Cat bond sous les conditions d'AOA « absence d'opportunité d'arbitrage » et de continuité du marché. Elle a été approfondie et corrigée par l'étude de Burnecki et Kukla [2003].

Les hypothèses du modèle :

- (i). Continuité du marché
- (ii). Les variables des sinistres sont Ft-adaptés et calculables
- (iii). Il existe un processus de poisson doublement stochastique N_s ayant une intensité stochastique λ_s et décrivant le nombre d'occurrence d'une catastrophe spécifique dans une région donnée.
- (iv). Les pertes $\{X_k\}$ causées par les catastrophes t_k sont indépendantes et i.i.d avec une distribution $F(x) = P(X_k < x)$
- (v). Les processus N_i et X_k sont supposés être indépendants
On pourrait donc exprimer la perte totale L_s par : $L_t = \sum_{i=1}^{N_t} X_i$
- (vi). Un processus de taux d'actualisation r continue et composé définit pour $t > s$ par :
$$e^{-R(s,t)} = e^{-\int_s^t r(u)du}$$
- (vii). Dans une optique de simplification du modèle, ils vont supposer que le principal est totalement consommé au premier dépassement de seuil. De même pour les paiements de coupons qui sont arrêtés dès le premier dépassement du seuil de déclenchement.
- (viii). Soit $\tau \in [0; T]$, la période à partir de laquelle la perte totale excède le seuil D . On a : $\tau = \inf(t : L_t \geq D)$.

Baryschnikov, Mayo et Taylor [1998] définissent un processus $M_t = 1_{L_t \geq D}$.

Ils montrent que M_t suit un processus de Poisson doublement stochastique d'intensité stochastique Λ_t :

$$\Lambda_t = \lambda_t \{1 - F(D - L_t)\} 1_{L_t < D}$$

² Appendix : Processus Cox

Sous les hypothèses du modèle, les prix des CAT Bond pourraient être définis par les formules suivantes données dans par Burnecki et Kukla [2003] :

- **CAT bond zero coupon à $t \in [0 ; T]$:**

$$V_t^{zc} = E [P e^{-R(t,T)} (1 - M_T) | \mathcal{F}_t]$$

En utilisant les propriétés des processus de Cox dans la relation précédente, ils obtiennent l'expression finale suivante :

$$V_t^{zc} = E \left[P e^{-R(t,T)} \left(1 - \int_t^T \Lambda_u du \right) | \mathcal{F}_t \right]$$

avec $\Lambda_t = \lambda_t \{1 - F(D - L_t)\} 1_{L_t < D}$

- **CAT bond payant des coupon C_t jusqu'à τ et payant P à T :**

$$V_t^C = E \left[P e^{-R(t,T)} (1 - M_T) + \int_t^T e^{-R(t,s)} C_s (1 - M_s) ds \middle| \mathcal{F}_t \right]$$

De même que dans le cas du zéro Coupon, on obtient :

$$V_t^C = E \left[P e^{-R(t,T)} \left(1 - \int_t^T \Lambda_u du \right) + \int_t^T e^{-R(t,u)} C_u \left(1 - \int_t^u \Lambda_s ds \right) du \middle| \mathcal{F}_t \right]$$

Par définition de $R(t, T)$, on a $R(t, T) = R(t, s) + R(s, T)$.

En utilisant la relation précédente, par simplification on obtient :

$$V_t^C = E \left[P e^{-R(t,T)} + \int_t^T e^{-R(t,u)} \left\{ C_u \left(1 - \int_t^u \Lambda_s ds \right) - P e^{-R(u,T)} \Lambda_u \right\} du \middle| \mathcal{F}_t \right]$$

Härdle et Cabrera [2010] ont utilisé cette approche pour calibrer le Cat Bond paramétrique émis par le gouvernement mexicain.

Cette approche de modélisation comporte certaines limites. L'hypothèse de consommation globale du principal au premier dépassement de seuil limite les flux de coupons pour les investisseurs. Dans la réalité, le principal n'est pas forcément perdu au premier déclenchement du CAT Bond. Seule une fraction du principal et une partie des coupons sont perdues par l'investisseur en cas de dépassement de seuil.

C'est le cas des CAT Bond indemnitaires où les premières indemnités sont généralement inférieures au principal. C'est aussi le cas des CAT Bond paramétriques qui sont généralement définis avec des fractions de paiements en fonction des seuils atteints par l'indice utilisé pour le déclenchement des CAT Bond.

L'approche qui sera abordée dans la section suivante est une modélisation du CAT Bond dans un environnement discret qui tient compte des différents flux au cours de la durée du contrat.

1.5.2. Approche théorie d'équilibre des prix dans un marché incomplet

Cette méthode présentée par Cox et Pedersen [2000] est basée sur une modélisation dans un marché incomplet par une approche d'agent représentatif³. La modélisation dans un marché incomplet est motivée par le fait que les CAT Bond ne peuvent être couverts avec des actifs du marché à cause du fait que les facteurs liés aux événements catastrophes sont décorrélés des facteurs financiers du marché.

Ils proposent dans leur étude un modèle de valorisation à temps discret qui se base sur le modèle de l'agent représentatif issu de la théorie d'équilibre des prix.

Les hypothèses du modèle :

- (i). On considère $(\Omega, \mathcal{F}, \mathbb{P})$ l'espace probabilisé du modèle global.
- (ii). On suppose que les prix des CAT Bond ne dépendent que des facteurs financiers du marché et/ou des facteurs liés aux événements catastrophes.
- (iii). Il existe un marché avec absence d'opportunité d'arbitrage.
- (iv). Les risques financiers ne dépendent que des informations du marché financier.
- (v). La consommation globale ne dépend que des facteurs financiers.
- (vi). On se place dans le cadre d'un marché incomplet. Ce qui implique que la probabilité risque-neutre n'est pas unique.
- (vii). On supposera que nous sommes dans un modèle discret.
- (viii). Les agents utilisent leurs fonctions d'utilité pour faire des choix d'investissement.
- (ix). $c(t)$ est un processus stochastique \mathcal{F}_t - adapté qui représente la consommation d'un agent économique à l'instant $t \in [0; T]$.
- (x). $C(t)$ est un processus stochastique \mathcal{F}_t - adapté qui représente la consommation globale disponible à l'instant $t \in [0; T]$
- (xi). On va supposer que la fonction d'utilité de l'agent représentatif est additivement séparable⁴ et dérivable.

On définit :

- \mathcal{F}^f la filtration des événements dépendant uniquement des facteurs financiers
- \mathcal{F}^c la filtration des événements dépendant uniquement des facteurs catastrophes
- u_k représente la fonction d'utilité d'un agent économique à la période k

³ Agent représentatif : En macro-économie, le terme agent représentatif représente un agent économique représentant tous les autres agents.

⁴ Une fonction d'utilité additivement séparable implique qu'il existe des fonctions d'utilités $\{u_k | k = 0, \dots, T\}$ telle que l'espérance de la fonction d'utilité d'un agent représentatif pour une consommation $\{c(k) | k = 0, \dots, T\}$ est égale à l'espérance de la somme des fonctions d'utilité : $E \left[\sum_{k=0}^T u_k(c(k)) \right]$

Etant dans un marché incomplet, la probabilité risque neutre n'est donc pas unique. Il faut donc chercher une mesure de probabilité qui permettent de ramener le calcul du prix dans des processus d'actualisation classiques. Pour cela, les auteurs ont choisi les taux d'intérêts $\{r(k)|k = 0, \dots, T - 1\}$ tels que :

$$\frac{1}{1 + r_t} = \frac{1}{u'_t(C(t))} E^P[u'_{t+1}(C(t+1)) | \mathcal{F}_t]$$

Selon la théorie de l'agent représentatif, le prix $V_{CAT}(t)$ d'une série de flux $F_{CAT} = \{F_{CAT}(k)|k = 1, \dots, T - 1\}$ d'un agent est donné par la relation :

$$V_{CAT}(0) = E^P \left[\sum_{k=1}^T \frac{u'_k(C(k))}{u'_0(C(0))} * F_{CAT}(k) \right]$$

Soit la dérivée de Radon-Nikodym de Q par rapport à P, une variable aléatoire positive telle que :

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_T} \stackrel{\text{def}}{=} \prod_{k=0}^{T-1} [1 + r_{j-1}] \frac{u'_T(C(T))}{u'_0(C(0))}$$

On peut démontrer que $E^P \left[\left. \frac{dQ}{dP} \right|_{\mathcal{F}_T} \right] = 1$ (démonstration disponible en Annexe de cette étude et en Annexe B de l'étude de Cox et Pedersen) et que Q est une mesure de probabilité équivalente à P⁵.

En effectuant un changement de probabilité de la probabilité historique P à la nouvelle mesure de probabilité, on a :

$$V_{CAT}(0) = E^Q \left[\sum_{k=1}^T \frac{u'_T(C(k))}{u'_0(C(0))} * F_{CAT}(k) * \left. \frac{dQ}{dP} \right|_{\mathcal{F}_T} \right]$$

On peut en déduire l'expression du prix du CAT Bond sous Q⁶ (démonstration disponible en Annexe).

$$V_{CAT}(0) = E^Q \left[\sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1})} * F_{CAT}(k) \right]$$

Cette expression pourrait être simplifiée avec les informations de la probabilité historique des risques catastrophes en considérant l'hypothèse que les flux des obligations catastrophes dépendent uniquement des variables liées aux risques catastrophes.

Dans ce modèle, l'aversion au risque de l'agent représentatif est prise en compte dans la formule par l'intermédiaire de sa fonction d'utilité qui est incluse dans la définition de la mesure de probabilité Q et des taux r_j . Comme stipulé dans l'hypothèse (viii), les choix d'investissements des agents sont réalisés selon leur fonction d'utilité. Il est donc important de conserver cette hypothèse dans la suite de la lecture de cette étude.

⁵ Se référer à Annexe III-1 pour la démonstration.

⁶ Se référer à Annexe III-2 pour le changement de probabilité

Zimbidis, Frangos, and Pantelous [2007] ont proposé dans leur étude une approche de modélisation des prix des CAT Bond liés aux tremblements de terre sur les côtes grecques en partant de cette approche de valorisation de CAT Bond.

Après cette revue des CAT Bond et leur mode fonctionnement, la partie suivante sera dédiée à la description des principales approches de Deep Learning.

Partie 2. Généralités sur le Deep Learning et les réseaux de neurones

L'objectif de cette section est de présenter l'approche « Deep Learning » qui sera utilisée dans cette étude. L'approche Deep Learning qui sera abordée dans cette étude est principalement le réseau de neurones. Cette approche qui tient son raisonnement du fonctionnement neuronal du cerveau qui à partir des expériences observées par le passé et de nouvelles observations, active plusieurs neurones pour l'interprétation des informations nouvelles. Dans le cadre de l'apprentissage profond, l'approche réseau de neurones artificiel essaie de reproduire ce même mécanisme.

Dans cette section, il sera d'abord question de décrire le fonctionnement d'un réseau de neurone. Ensuite, on présentera les principaux types des réseaux de neurones. Enfin, un focus sera fait sur l'approche séquentielle des réseaux de neurones.

2.1. Description de l'approche réseau de neurones

2.1.1. Structure d'un réseau de neurones

Un modèle de réseau de neurone peut être subdivisé en 3 grandes parties appelées couches :

- La première couche est la couche d'entrée qui représente les caractéristiques des variables d'observations utilisées pour l'apprentissage du modèle.
- La seconde couche appelée couches cachées est un regroupement de couches qui servent à établir des liens entre les variables explicatives et des variables expliquées. Chaque couche est composée de neurones et des liens d'activations des neurones des couches suivants (figure 7).
- La dernière couche représente la sortie du modèle c'est-à-dire la prédiction du modèle

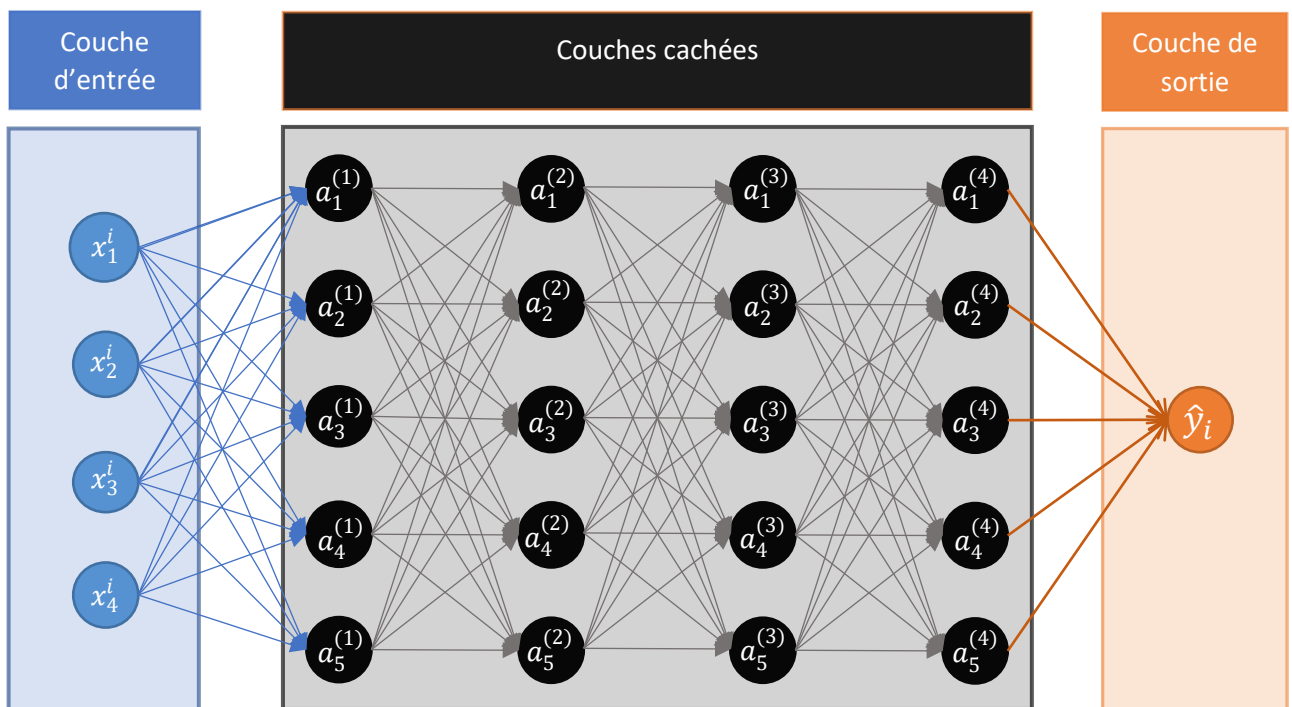


Figure 7 - Exemple de structure d'un réseau de neurone

Entre chaque couche, les informations des couches précédentes se propagent par l'intermédiaire d'une fonction d'activation. Le schéma ci-dessus résume le transfert d'information entre deux couches.

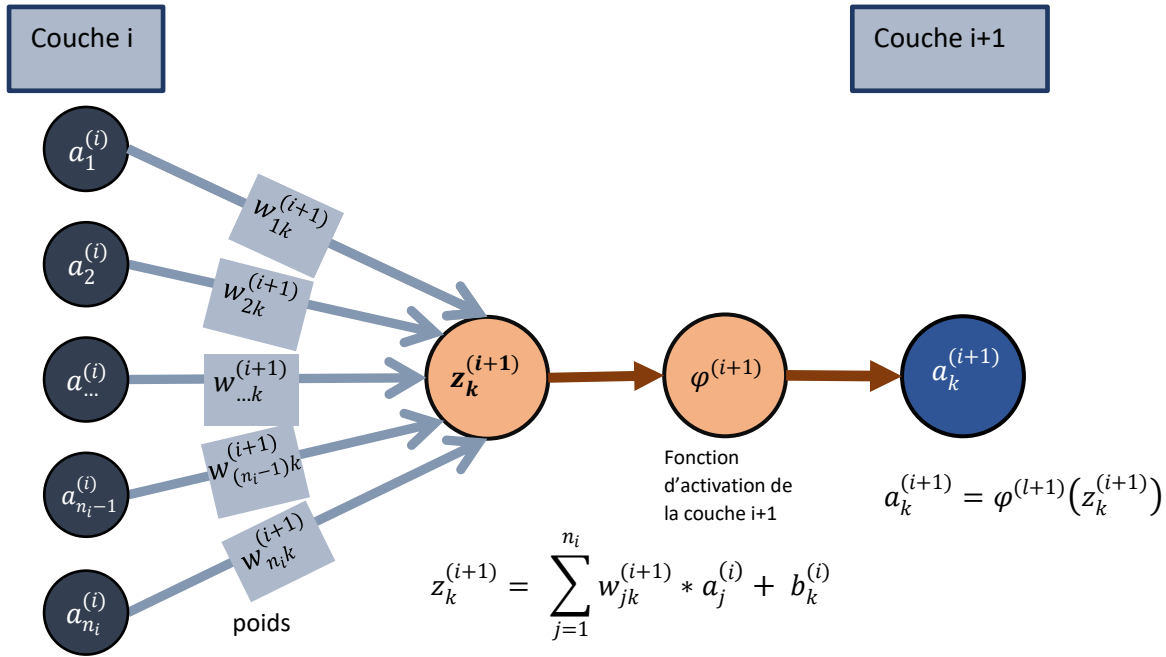


Figure 8 - Transfert d'informations entre des neurones de deux couches (i) et (i+1)

On utilisera les notations matricielles suivantes dans la suite de cette étude :

- \hat{Y} la matrice des valeurs obtenues par le modèle
- m la taille des observations utilisées pour modéliser le réseau de neurone
- n le nombre d'attribut des données d'entrées.
- La matrice des données de la couche d'entrée ou de modélisation

$$X = \begin{bmatrix} x_1^1 & \dots & x_n^1 \\ \vdots & \ddots & \vdots \\ x_1^m & \dots & x_n^m \end{bmatrix} \text{ avec le } i\text{-ème donnée } X^i = [x_1^i \quad \dots \quad x_n^i]$$

- La matrice des valeurs réelles à prédire : $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$

- La matrice des poids entre les couches $k - 1$ et k : $W^{[k]} = \begin{bmatrix} w_{11}^{[k]} & \dots & w_{n_{k-1}1}^{[k]} \\ \vdots & \ddots & \vdots \\ w_{1n_k}^{[k]} & \dots & w_{n_{k-1}n_k}^{[k]} \end{bmatrix}$

- La matrice des neurones de la couche k : $A^{[k]} = \begin{bmatrix} a_1^{[k]} \\ \vdots \\ a_{n_k}^{[k]} \end{bmatrix}$ avec $A^{[0]} = X$

- Le terme de biais de la couche k : $b^{[k]} = \begin{bmatrix} b_1^{[k]} \\ \vdots \\ b_{n_k}^{[k]} \end{bmatrix}$

2.1.2. Fonctions d'activation ou fonction de transfert

Les fonctions d'activations dans les réseaux de neurones ont pour objectif d'introduire une non-linéarité dans la modélisation des réseaux de neurones. Ils permettent ainsi de désactiver les neurones issus de combinaisons des données d'entrée qui n'ont pas d'impact significatif sur l'explication des données de sorties : on parle aussi de fonction de seuillage. On distingue plusieurs types de fonctions d'activations. On présentera ci-dessous les principales fonctions d'activation.

- **ReLU ou fonction linéaire modifiée :**

La fonction ReLU est la fonction la plus utilisée dans les modèles de régression. Sa principale fonctionnalité est de casser la linéarité du modèle. Elle permet donc de désactiver certains nœuds qui ont une influence négligeable sur la précision du modèle. Elle a aussi l'avantage d'avoir une dérivée facile à utiliser.

$$\varphi(x) = \max(x; 0)$$

Elle a pour principal désavantage, la nullité de sa dérivée pour les valeurs négatives. Ce qui entraîne la désactivation des nœuds à partir du moment où la valeur est négative. Cela empêche la propagation des informations dans les couches.

- **Fonction d'activation Swish:**

Cette fonction a été développée pour pallier les problèmes de propagation de la fonction ReLU pour les valeurs négatives. Dans cette approche, plutôt que de désactiver un nœud par la fonction d'activation, le poids des valeurs négatives est atténué avec la fonction sigmoïde.

$$\varphi(x) = x * \frac{1}{1 + e^{-x}}$$

- **Fonction d'activation sigmoïde :**

Cette fonction est principalement utilisée dans les modèles de classification dans lesquels les sorties sont de deux classes. Elle est définie comme suit :

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- **Fonction d'activation softmax :**

La fonction d'activation softmax est utilisée dans les modèles de classification dans lesquels les sorties sont de plusieurs classes c'est-à-dire de plus de 2 classes à prédire. On peut l'utiliser par exemple dans un modèle de reconnaissance des chiffres. Elle est définie comme suit :

Soit x_i la classe i ,

$$\sigma(x_i) = \frac{e^{x_i}}{\sum_{k=1}^n e^{x_k}}$$

2.1.3. Fonctions d'objectif ou fonctions de coûts

La fonction d'objectif ou fonction de coûts peut être définie comme la fonction du coût global des erreurs d'estimations entre les valeurs observées et celles prédites par le modèle. On pourrait résumer le processus de recherche du modèle optimal d'un réseau de neurones comme une procédure de recherche des paramètres qui vont permettre de minimiser cette fonction d'objectif. Plus la fonction est proche de zéro, plus le modèle arrive à prédire les données observées

On distingue plusieurs types de fonctions d'objectifs que nous allons aborder ci-dessous. On notera pour la suite J la fonction de coût qu'on cherchera à minimiser,

Les principales fonctions de coûts utilisées pour résoudre des problèmes de régression sont généralement :

- La moyenne des erreurs au carrés (MSE) :
Cette fonction a pour avantage de pénaliser les estimations éloignées des observations du fait de la présence du carré dans le terme d'erreur.

$$J(Y, \hat{Y}) = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

- La moyenne des erreurs absolues (MAE).
Cette fonction a pour objectif de donner une information sur l'amplitude des erreurs d'estimation.

$$J(Y, \hat{Y}) = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|$$

En ce concerne les problèmes de classifications, les principales fonctions utilisées sont :

- La fonction d'entropie croisée : Cette fonction est utilisée dans le cas des modèles de classification binaire c'est-à-dire deux classes à prédire.

$$J(Y, \hat{Y}) = -\frac{1}{m} \left(\sum_{i=1}^m y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i) \right)$$

- L'entropie croisée multinomiale : Cette fonction est utilisée dans le cas des modèles de classification ayant un nombre de classes à prédire supérieure à 2.

$$J(Y, \hat{Y}) = -\frac{1}{m} \left(\sum_{i=1}^m \sum_{j=1}^K y_i^{[j]} \log \hat{y}_i^{[j]} + (1 - y_i^{[j]}) \log (1 - \hat{y}_i^{[j]}) \right)$$

Avec $y_i^{[j]}$ la valeur observée de la classe j pour la donnée d'apprentissage i

2.1.4. Hyperparamètres d'un modèle de réseau de neurones

Les hyperparamètres sont des variables utilisées dans la recherche du modèle optimal durant la phase d'apprentissage et qui ne dépendent pas des données. Ces variables peuvent être définies comme des paramètres d'aide à l'optimisation des algorithmes d'apprentissage. Ils peuvent contribuer à l'accélération de la convergence des données vers un optimum local, à la généralisation du modèle, etc.

Ils existent plusieurs hyperparamètres. Ci-dessous seront présentés les principaux hyperparamètres utilisés dans les réseaux de neurones.

- Nombre de couches et nombre de neurones par couche

Lors de la mise en place d'un modèle réseau de neurones, il faut spécifier le nombre de couches que va comporter le modèle et définir le nombre de neurones dans chaque couche du réseau de neurones. Ce choix s'avère important dans l'implémentation car un nombre très élevé de nœuds dans une couche ou de couches peut entraîner un surapprentissage du modèle c'est-à-dire une bonne réplique des données d'apprentissage mais un modèle qui a des performances mauvaises sur des données nouvelles. De même, un nombre insuffisant de couches ou de neurones pourrait limiter la capacité du modèle à apprendre des données d'observations.

Il est donc primordial, dans la mise en place d'un modèle réseau de neurones, de bien calibrer le nombre de couches et aussi le nombre de nœuds dans chaque couche.

- Taux d'apprentissage

L'hyperparamètre du taux d'apprentissage détermine la vitesse à laquelle le processus d'apprentissage va converger vers un modèle optimal.

Le choix de ce paramètre a un impact significatif sur la modélisation dans l'implémentation d'un modèle de réseau de neurone. On notera α le taux d'apprentissage.

- Régularisation

Le paramètre de régularisation est utilisé dans l'optimisation de la fonction de coût afin de limiter les problèmes de surapprentissage (« overfitting ») du modèle sur les données d'apprentissage. Il permet donc d'aider à trouver un modèle qui pourra être généralisé sur des données nouvelles. On notera λ le paramètre de régularisation. Plusieurs méthodes de régularisation sont utilisées dans les réseaux de neurones. Les techniques de régularisation couramment utilisées sont les suivantes :

o La régularisation L2

La régularisation L2 consiste à ajouter à la fonction de perte un terme additionnel qui est la norme euclidienne de la matrice des poids du modèle. Ce terme est multiplié par le paramètre de régularisation λ . On obtient une nouvelle fonction de perte :

$$J_{L2}(Y, \hat{Y}) = J(Y, \hat{Y}) + \lambda \sum_{i=1}^m \sum_{j=1}^c \sum_{k=1}^{n_j} w_{ij}^{[k]2}$$

- La régularisation L1

La régularisation L1 consiste à ajouter à la fonction de perte un terme additionnel qui est la somme des valeurs absolues des poids du modèle. Le terme de régularisation est aussi appliqué à ce terme.

$$J_{L1}(Y, \hat{Y}) = J(Y, \hat{Y}) + \lambda \sum_{i=1}^m \sum_{j=1}^C \sum_{k=1}^{n_j} |w_{ij}^{[k]}|$$

Où m représente le nombre de données utilisées pour l'apprentissage, C le nombre de couches du modèle et n_j le nombre de neurones dans la couche j .

- **Nombre d' « epoch » ou nombre de parcours de l'ensemble données d'apprentissage**

L'hyperparamètre communément appelé « epoch » dans les modèles Deep Learning, représente une itération sur toutes les données servant à apprendre les paramètres du modèle. Ainsi, le nombre d'epoch définit le nombre d'itérations de l'ensemble des données d'apprentissage. Ce paramètre définit généralement l'arrêt des itérations d'apprentissage c'est-à-dire qu'une fois le nombre spécifié dans cet hyperparamètres est atteint le modèle arrêtera son apprentissage sur les données fournis pour apprendre d'où l'importance de la calibration de ce paramètre dans un modèle réseau de neurone.

2.1.5. Optimisation du modèle d'estimation

Dans la modélisation des réseaux de neurones, l'approche principale d'optimisation est la méthode de rétropropagation ou « Back propagation » en utilisant par le mécanisme de descente de gradient pour la mise à jour des poids.

L'optimisation du modèle consistera à minimiser la fonction de coût par itérations successives en modifiant, à chaque itération, les paramètres de poids jusqu'à convergence de la fonction de coût. On pourrait résumer l'approche d'optimisation par une recherche des poids qui minimise la fonction de coût. Cette recherche peut être exprimée sous la

forme : $\min_W J(\hat{Y}, Y)$

Cette recherche d'optimisation du modèle se déroule en quatre étapes :

- Initialisation des poids du modèle
- Propagation des poids dans les différentes couches jusqu'à obtention d'une valeur prédite
- Estimation de la fonction de perte du modèle
- Rétropropagation des erreurs du modèle dans les différentes couches.

Les trois dernières étapes vont être répétées jusqu'à la convergence de la fonction de perte vers un optimum local.

- **Initialisation des poids :**

Dans un premier temps, les différents poids entre les couches sont initialisés. Cette étape d'initialisation des poids est une étape importante dans la mise en place d'un modèle réseau de neurones. On choisira généralement, de façon aléatoire, des valeurs strictement positives entre 0 et 1 comme valeur d'initialisation.

Cette étape n'intervient qu'une seule fois dans le processus. C'est l'étape initiale de la recherche d'un modèle optimal dans un réseau de neurones.

- **Propagation des poids dans le modèle et estimation de la fonction d'objectif :**

La seconde étape, généralement appelée « forward propagation » dans le langage commun de la communauté intelligence artificielle, consiste à propager ces poids jusqu'à la couche de sortie afin d'estimer la fonction de perte. Elle sert à estimer l'erreur du modèle à partir des observations et des nouveaux poids obtenus.

Cette étape sera répétée à chaque itération de retro-propagation jusqu'à convergence de la fonction d'objectif. Cette étape consiste donc à propager les informations d'entrée vers les couches suivantes avec les nouveaux poids.

- **Rétropropagation ou actualisation des poids du modèle :**

La dernière et principale étape de cette méthode d'optimisation est l'étape de rétropropagation des erreurs d'estimation dans les différentes couches. Cette étape est généralement identifiée comme « back propagation » dans les littératures sur le Deep Learning.

Elle consiste à actualiser les poids en fonction de la contribution de chaque poids dans les erreurs estimées. Cette mise à jour des poids va dans le sens opposé à de propagation des poids c'est-à-dire on part de la couche sortie à la couche d'entrée.

Elle se base sur le principe de descente de gradient.

$$w_i^{new} = w_i^{old} - \alpha \frac{\partial}{\partial w_i} J(\hat{Y}, Y)$$

2.2. Les principaux types de réseaux de neurones

Il existe plusieurs types de réseaux de neurones. Dans le cadre de notre étude, nous allons nous focaliser uniquement sur 4 types de réseaux de neurones.

- Réseau de neurone perceptron (MLP)
- Réseaux de neurones séquentiels
 - Réseau de neurone récurrents (RNN)
 - GRU – Gated Recurrent Unit
 - LSTM – “Long Short-Term Memory”

2.2.1. Le réseau de neurones perceptron (MLP)

Le modèle MLP est le modèle général qui a été décrit dans la section 2.1 qui décrit le réseau de neurones. C'est un modèle non séquentiel qui considère toutes les observations comme des données indépendantes. Dans un modèle MLP, les informations apprises des différentes observations sont extraites parallèlement. Ce point constitue son principal désavantage dans la résolution de problématiques séquentielles ou temporelles.

Dans le cadre de notre étude, nous n'aborderons pas cette approche. Nous l'avons utilisé comme modèle pour décrire le principe de fonctionnement d'un réseau de neurone dans un cadre général.

2.2.2. Les modèles de réseau de neurones récurrents ou séquentiels

L'approche traditionnelle (MLP) des réseaux de neurones considère les différentes observations de la couche d'entrée comme des données indépendantes. Dans un modèle séquentiel, les valeurs futures dépendent des états précédents. Si l'approche MLP était utilisée dans le cas d'une modélisation de température par exemple, les températures des jours précédents seraient considérées comme des cas indépendants de la situation de la période active comme si les températures des périodes précédentes n'avaient aucun impact sur la valeur future à prédire. Dans une optique de recherche d'une solution pour résoudre ce problème lié aux données séquentielles, il a été mis en place plusieurs approches dont le réseau de neurones récurrents ou RNN qui a pour spécificité de tenir compte des précédents états d'une variable dans la mise en place d'un modèle d'estimation des valeurs futures. Cette approche RNN est beaucoup utilisée dans les problématiques séquentielles telles que la modélisation des outils de traduction ou de prédiction des mots suivant, etc.

- Modélisation du modèle RNN

Dans le modèle RNN, à la période t , le modèle reçoit les variables d'observations x_t et les hypothèses h_{t-1} qui représentent les informations apprises des périodes précédentes. Ces deux informations vont permettre au modèle d'apprendre de nouvelles informations de la période t et prédire les sorties de la période t .

L'estimation de t est obtenue à partir des nouvelles hypothèses obtenues en appliquant une fonction d'activation φ aux informations apprises sur x_t et h_{t-1} .

Cette hypothèse h_t basée sur des données historiques sera utilisée comme hypothèse d'entrée pour apprendre de nouvelles informations des données des périodes suivantes et ainsi de suite.

$$h_t = \varphi(W_{xt}x_t + W_{ht}h_{t-1} + b_{ht})$$

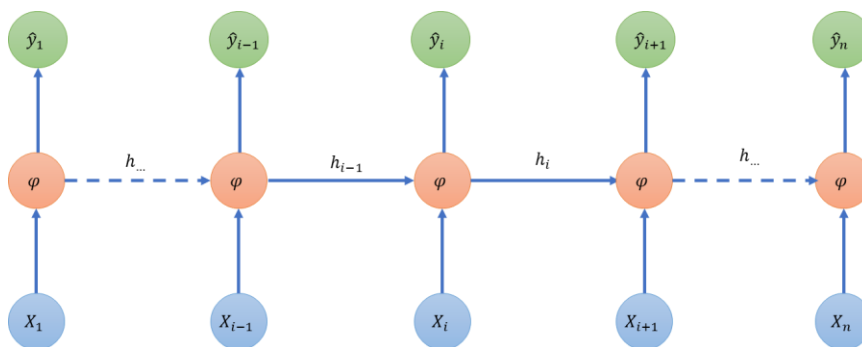


Figure 9 - Structure d'un RNN

Bien que développer pour la gestion des modèles séquentiels, le RNN connaît un désavantage majeur dans sa capacité à résoudre des problématiques séquentielles. En effet, l'un des principaux défauts du RNN est sa capacité à mémoriser de longues séquences de dépendance entre les différentes périodes d'observations. Théoriquement, le RNN devrait être capable de tenir compte des dépendances long termes c'est-à-dire être capable de prendre en compte l'impact des toutes premières séquences sur les valeurs futures. Dans la pratique, on constate qu'il a du mal à conserver les dépendances long termes. Ce problème vient principalement de l'approche d'optimisation du modèle RNN qui est la

descente de gradient à travers le temps. Cette approche d'optimisation démultiplie le processus de mise à jour entre les différentes couches.

La méthodologie de rétropropagation des erreurs va entraîner, dans l'approche RNN, un phénomène de perte de gradient avec la convergence rapide vers 0 des poids des différentes couches ou une explosion du gradient qui entraîne une divergence du gradient. Ce qui par conséquent empêche la mise à jour des poids dans les différentes couches antérieures avec les nouvelles informations provenant de la fonction de perte.

Plusieurs approches séquentielles ont été développées ces dernières années pour contourner cette contrainte du modèle RNN.

Nous allons décrire les deux principales approches séquentielles créées ces dernières années pour répondre à cette problématique.

- L'approche LSTM "Long Short-Term Memory" ou réseau de neurones récurrents à mémoire court-terme et long-terme.
- L'approche GRU "Gated Recurrent Unit" ou réseau de neurones récurrents à portes

2.2.3. Réseau de neurones récurrents à mémoire long et court-termes

Le modèle de réseau de neurones récurrents à mémoire long et court-termes ou LSTM est un approfondissement de l'approche RNN. Elle a été proposée par Hochreiter et Schmidhuber [1997]. Dans cette approche, les auteurs ont proposé une approche permettant de tenir compte des dépendances à long et court terme. L'un des avantages de l'approche LSTM est qu'elle permet de stocker les informations des périodes précédentes dans une mémoire. Le LSTM permet aussi à la mémoire d'oublier des informations contrairement au RNN classique. Le LSTM est l'approche la plus utilisée dans les projets RNN.

Comparativement à l'approche RNN classique, le modèle LSTM introduit 3 nouvelles notions qui sont la porte d'actualisation ou «Input Gate», la porte d'oubli ou «Forget Gate» et la porte de sortie ou «Output Gate». Le modèle aura comme données de sorties un état h_t et une mémoire c_t .

Nous allons conserver dans la suite de cette section les appellations usuelles des trois portes du LSTM : Input Gate, Forget Gate, Output Gate.

- « Forget Gate » ou la porte de l'oubli

Dans le modèle LSTM, le Forget Gate sert à la gestion de l'historique de la mémoire c_t . Elle a donc pour principal objectif de permettre à la mémoire du modèle d'oublier des hypothèses apprises des périodes précédentes et qui ne seront plus pertinentes à partir de la période t . Ainsi donc, à la sortie de cette porte, on a des informations sur les données apprises de l'historique qui seront effacées de la mémoire du modèle.

Il peut être utilisé pour signaler, par exemple, des fins de tendances ou de saisonnalité.

Le Forget Gate a donc pour principal objectif de permettre à la mémoire du modèle d'oublier des hypothèses obsolètes.

Le processus de mise à jour de la mémoire se déroule en deux étapes :

- A partir des observations x_t de la période t et des informations stockées h_{t-1} dans la période précédente $t-1$, la fonction sigmoïde est appliquée à la somme pondérée x_t et h_{t-1} pour identifier les informations à oublier de la mémoire c_t .

- Après application de la fonction sigmoïde à la somme pondérée de x_t et h_{t-1} , les règles décrites ci-dessous seront effectuées :
 - o plus la valeur obtenue pour une information est proche de 0, plus cette information sera considérée par le modèle comme une information historique à oublier.
 - o plus la valeur obtenue pour une information est proche de 1, plus cette information sera considérée par le modèle comme une information historique pertinente pour la période donc à conserver.

La fonction F_t de cette porte peut être formulée comme suit :

$$F_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f)$$

Où σ représente la fonction sigmoïde, W_{xf} les poids des données d'observations, W_{hf} les poids des éléments des informations apprises venant de la période précédente t-1.

Le vecteur F_t croisé avec la mémoire c_{t-1} permet d'actualiser les cellules de la mémoire pour la période t.

Etant donné que des informations pourraient être supprimées de la mémoire, la prochaine porte du modèle LSTM va proposer à la fonction mémoire les nouvelles valeurs à considérer à partir de cette période. Ce process a lieu dans l'Input Gate.

- Input Gate

Cette partie du LSTM a pour objectif la mise à jour de la mémoire c_t avec les nouvelles informations apprises et qui vont remplacer les données effacées de la mémoire à la porte Forget Gate. Le fonctionnement de cette porte est le suivant :

- A partir des observations x_t de la période t et des informations stockées h_{t-1} dans la période précédente t-1, la fonction sigmoïde est appliquée à la somme pondérée x_t et h_{t-1} afin d'identifier les cellules dont les informations ont été effacées et qui doivent être mise à jour.
- Il est ensuite appliqué la fonction tangente hyperbolique pour permettre à l'identification de nouvelles informations apprises qui seront des candidats au remplacement de celles effacées de la mémoire.

Les fonctions I_t et \tilde{c}_t de cette porte peuvent être formulées comme suit :

$$I_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i)$$

$$\tilde{c}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$$

Où

σ représente la fonction sigmoïde, \tilde{c}_t représente le vecteur des nouvelles valeurs apprises.

W_{xi} les poids des données d'observations,

W_{hi} les poids des éléments des informations apprises venant de la période précédente t-1.

Le croisement de ces deux informations va permettre d'actualiser la mémoire.

On définit l'opérateur \otimes comme le produit termes à termes de deux matrices.

A la sortie des procédures des portes Input Gate et Forget Gate, la procédure d'actualisation de la mémoire pourrait être formulée comme suit :

$$c_t = F_t \otimes c_{t-1} + I_t \otimes \tilde{c}_t$$

- **Output Gate**

L'Output Gate ou porte de sortie a pour objectif de filtrer les informations apprises de la mémoire et des données de la période t afin d'obtenir de nouvelles hypothèses pour la prédiction de la valeur de sortie. Il a donc pour objectif la mise à jour de la variable d'état h_t avec les nouvelles hypothèses.

Le fonctionnement de cet cette porte est le suivant :

- la fonction sigmoïde est appliquée à la somme pondérée x_t et h_{t-1} pour identifier les hypothèses O_t à retenir pour cette période.
- Il sera appliqué une transformation par la fonction tangente hyperbolique à la mémoire c_t afin d'obtenir des valeurs entre $[-1,1]$.

L'hypothèse de la prédiction de la période t sera le produit entre les résultats des deux opérations précédentes.

Les fonctions O_t et h_t de cette porte peuvent être formulées comme suit :

$$O_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o)$$

$$h_t = O_t \otimes \tanh(c_t)$$

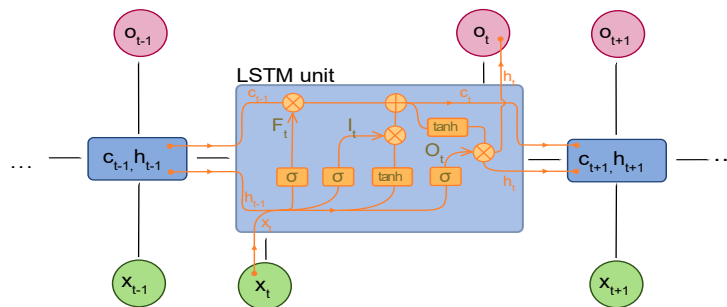


Figure 10 - Modèles LSTM (source : Wikipédia)

Le modèle LSTM est un modèle qui fait appel à plusieurs variables et paramètres. Ce qui rend complexe sa mise à place et surtout le temps de traitement des procédures d'optimisation. Un modèle similaire a été développé dans les années 2014.

2.2.4. Réseau de neurones récurrents à portes

Le réseau de neurones récurrents à portes ou communément appelé GRU a été proposé par Cho et Al [2014]. Il a un fonctionnement pratiquement similaire à celui du LSTM.

Contrairement au LSTM, le GRU n'a que 2 portes (« Update Gate » et « Reset Gate ») et n'a pas de sortie dédiée à la mémoire.

A titre de comparaison avec le modèle LSTM, la porte reset Gate du GRU pourrait être considérée comme une combinaison des deux portes Input et Forget de l'approche LSTM. Aussi, la mémoire et la variable d'état sont mergées pour une former une variable h.

Cette approche améliore le temps de recherche du modèle optimal comparativement au modèle LSTM.

- Update Gate ou porte de mise à jour des informations :

$$Z_t = \sigma(W_{xu}x_t + W_{hu}h_{t-1} + b_u)$$

- Reset Gate ou porte de réinitialisation des informations :

$$R_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r)$$

$$\tilde{h}_t = \tanh(W_{hc}x_t + W_{hc}R_t + b_h)$$

A partir de ces deux informations, on obtient l'état de sortie h_t

$$h_t = (1 - Z_t) \otimes h_{t-1} + Z_t \otimes \tilde{h}_t$$

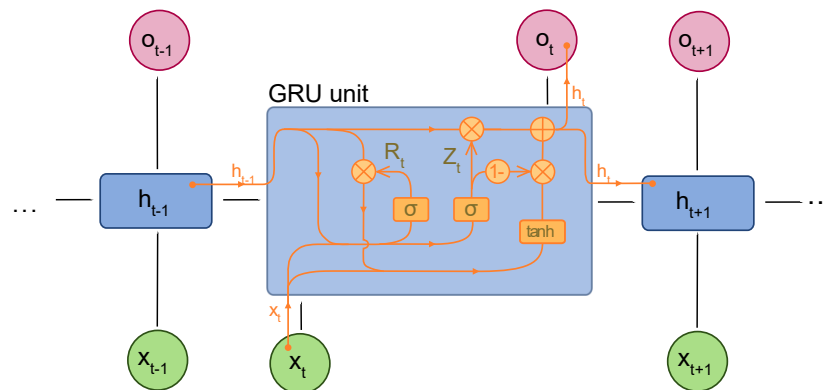


Figure 11- Modèle GRU (source : Wikipédia)

Dans les trois approches, l'objectif est d'obtenir les poids W qui minimisent la fonction d'objectif. L'approche d'optimisation utilisée est l'approche par rétropropagation qui consiste à approcher les valeurs optimales par itérations successives.

Il a été présenté dans les deux présentes parties le fonctionnement d'un CAT Bond et d'un réseau de neurones. Dans la section suivante, il sera étudié des méthodes de valorisation d'un CAT Bond.

Partie 3. Modélisation CAT Bond

Cette partie du mémoire sera consacrée à la modélisation d'un CAT Bond à déclencheur paramétrique. Dans cette optique, dans un premier temps, les principales hypothèses de modélisation seront définies.

Ensuite, deux approches de modélisations d'un CAT Bond seront étudiées. La première approche sera basée sur la modélisation des risques CAT à l'aide de lois de probabilité et la seconde approche en utilisant une approche de modélisation des risques catastrophes par des modèles d'apprentissage approfondi (Deep Learning).

3.1. Hypothèses de modélisation

3.1.1. Hypothèse sur le modèle du marché

Dans ce mémoire, le CAT bond sera modélisé dans un marché incomplet du fait de l'impossibilité de répliquer les risques catastrophes avec des produits financiers primaires. La démonstration de la pertinence de cette hypothèse est réalisée en se basant sur le même argumentaire que celui réalisé dans l'étude de Cox et Pedersen [2000].

3.1.2. Autres hypothèses de modélisations

Dans une démarche de simplification, les hypothèses suivantes seront considérées tout au long de cette étude.

- (i). Le capital assuré par le sponsor sera supposé connu dans cette étude.
- (ii). Par simplification, les coûts frictionnels liés aux coûts de transaction et de gestion du SPV seront supposés négligeables dans la modélisation. Dans la réalité, ils pourraient être pris en compte.
- (iii). L'aléa moral sera considéré comme négligeable. Dans la pratique, il pourrait exister un risque que l'assureur soit plus large dans l'estimation des sinistres lorsque la somme cumulée des pertes est proche du seuil de déclenchement. Ce qui pourrait avoir un impact sur la probabilité de déclenchement des CAT Bond.
- (iv). Dans cette étude, les risques de base du CAT Bond seront supposés négligeables. Dans la réalité, comme mentionné dans l'étude de Lee et Yu [2003], le risque de base pourrait venir de la différence entre les pertes indicelles et les pertes réelles de l'émetteur.
- (v). On supposera que le risque de défaut de l'assureur est négligeable. Dans la pratique, le risque de défaut de l'assureur pourrait exposer les investisseurs à un défaut de paiement des primes. Ce risque a un coût qui pourrait être quantifié.
- (vi). On supposera que les coupons reçus sont une fraction du principal restant après paiements des sinistres pris en charge par le CAT bond.
- (vii). On supposera que les variables des risques liés aux événements catastrophes et les variables des risques financiers sont indépendants.
- (viii). Dans un objectif de simplification du modèle, on supposera que les différents flux financiers du CAT Bond ont lieu en fin d'année de chaque période concernée par les événements catastrophes.
- (ix). On supposera que le paramètre de déclenchement est basé sur le maximum annuel d'un indice paramétrique. Ce qui implique que les flux annuels seront calculés à partir des maximums annuels des critères de déclenchement.

3.1.3. Le modèle de valorisation

Cette étude se basera sur le modèle de valorisation des CAT Bond proposé par Cox et Pedersen [2003] qui a été résumé dans la section 1.5.2. de cette étude.

Dans notre étude, nous allons supposer une indépendance entre les risques catastrophes et les risques financiers.

Dans un souci de simplification, nous supposerons que seuls les facteurs liés à la survenance des événements catastrophes influent sur les différents flux de paiements des CAT Bond.

Ce qui nous emmène à utiliser l'expression du prix du CAT Bond dans le cas où les différents flux du CAT Bond ne dépendent que des risques catastrophes. Cette expression a été spécifiée dans la section 1.5.2 concernant la littérature sur les modèles existants.

$$V_{CAT}(0) = E^Q \left[\sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1})} * F_{CAT}(k) \right]$$

Dans le cas où les différents flux du cat bond ne dépendraient que des risques catastrophes, la formule du prix pourrait être simplifiée par cette expression donnée par Cox et Pedersen [2000]⁷ dans leur étude :

$$V_{CAT}(0) = \sum_{k=1}^T E^Q \left[\frac{1}{\prod_{j=1}^k (1 + r_{j-1})} \right] E^P [F_{CAT}(k)]$$

Cette formule peut être démontrée en utilisant les hypothèses de Cox et Pedersen.

Démonstration :

Comme définit dans la section 1.5.2, on a les deux filtrations :

- \mathcal{F}^f la filtration des événements dépendant uniquement des facteurs financiers
- \mathcal{F}^c la filtration des événements dépendant uniquement des facteurs catastrophes

Sous l'hypothèse d'indépendances entre les risques financiers et des risques catastrophes On a :

$$E^Q \left[\sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1})} * F_{CAT}(k) \right] = E^Q \left[\sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1})} \right] E^Q \left[\sum_{k=1}^T F_{CAT}(k) \right]$$

$$E^Q \left[\sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1})} * F_{CAT}(k) \right] = \sum_{k=1}^T E^Q \left[\frac{1}{\prod_{j=1}^k (1 + r_{j-1})} \right] * \sum_{k=1}^T E^Q [F_{CAT}(k)]$$

On a défini dans la section 1.5.2, la probabilité Q :

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_T} = \prod_{k=0}^{T-1} [1 + r_k] \frac{u'_T(C(T))}{u'_0(C(0))}$$

En effectuant un changement de probabilité, on a :

$$E^Q [F_{CAT}(k)] = E^P \left[F_{CAT}(k) \frac{dQ}{dP} \Big|_{\mathcal{F}_T} \right]$$

⁷ Annexe III-3 : CAT Bond dans le cas où les flux du CAT Bond dépendent uniquement des risques catastrophes

Selon l'hypothèse (v) du modèle, la consommation globale C ne dépend que des facteurs financiers. Cette hypothèse implique que $\frac{dQ}{dP}\Big|_{\mathcal{F}_T}$ ne dépend que des risques financiers car nous avons aussi les variables $r(k)$ qui ne dépendent que des risques financiers (hypothèses iv du modèle). On a donc $\frac{dQ}{dP}\Big|_{\mathcal{F}_T}$ qui est \mathcal{F}^f – mesurable

Et comme F_{CAT} est \mathcal{F}^C – mesurable en plus du fait que les risques catastrophes sont indépendants des risques financiers, on en déduit que F_{CAT} et $\frac{dQ}{dP}\Big|_{\mathcal{F}_T}$ sont indépendants.

On peut ainsi écrire :

$$E^Q[F_{CAT}(k)] = E^P[F_{CAT}(k)] * E^P\left[\frac{dQ}{dP}\Big|_{\mathcal{F}_T}\right]$$

Et comme $E^P\left[\frac{dQ}{dP}\Big|_{\mathcal{F}_T}\right] = 1$

On peut conclure que :

$$E^Q[F_{CAT}(k)] = E^P[F_{CAT}(k)]$$

Par conséquent :

$$V_{CAT}(0) = \sum_{k=1}^T B(k) E^{\mathbb{P}}[F_{CAT}(k)]$$

Avec

- $B(k) = E^Q\left[\frac{1}{\prod_{j=1}^k(1+r_{j-1})}\right]$ représente le prix zéro coupon payant 1 à la maturité.
- $F_{CAT}(k)$ représente les flux de cash pour la période i du CAT Bond.

Le choix de ce modèle discret est plus adapté à une étude par une approche Deep Learning où nous travaillerons sur des probabilités historiques de survenance de risques catastrophes.

On pourrait donc résumer le calcul du prix d'un CAT bond dans cette étude à l'estimation de $B(t)$ qui est l'évaluation des risques financiers et de l'estimation de l'espérance sous la probabilité historique des flux de cash du CAT Bond $E^{\mathbb{P}}[F_{CAT}(k)]$.

Dans la pratique, on choisit des facteurs financiers dans un environnement avec absence d'opportunité d'arbitrage pour construire la structure des taux afin de déduire le prix conditionné par ces facteurs financiers. $B(t)$ pourra être estimé dans l'environnement risque neutre Q.

En ce concerne $E^{\mathbb{P}}[F_{CAT}(k)]$, deux approches de calculs seront proposées : une approche basée sur la modélisation des risques catastrophe à l'aide de lois de probabilité et une approche basée sur les techniques d'apprentissage approfondi communément appelé « Deep Learning ».

On posera : $F_{CAT}(X_k) = F_{CAT}(k)$ en partant de l'hypothèse que les flux des CAT Bond dépendent uniquement des risques catastrophes.

• **Calcul du prix sous l'approche de modélisation par une loi :**

On pose : $V_k = \frac{1}{\prod_{j=1}^k (1+r_{j-1})} F_{CAT}(X_k)$

Sous l'hypothèse d'existence de l'espérance $E^{\mathbb{Q}}[V_k]$ et avec les hypothèses des variables des risques catastrophes étant i.i.d., les variables r_k étant supposées i.i.d, les variables X_k et r_k qui sont deux à deux indépendantes.

$E^{\mathbb{Q}}[V_k] \approx \lim_{m \rightarrow +\infty} \frac{1}{m} \sum_{i=1}^m V_k^{(i)}$ où $V_k^{(i)}$ représente le i -ème simulation de V_k .

Ce qui impliquerait que :

$$V_{CAT}(0) \approx \lim_{m \rightarrow +\infty} \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1+r_{j-1}^{(i)})} F_{CAT}(X_k^{(i)})$$

où $X_k^{(i)}$ et $r_k^{(i)}$ représentent les i -ème simulation des variables X_k et r_k .

Dans cette approche de modélisation par une loi de probabilité, nous allons d'abord modéliser les risques financiers et les risques catastrophes. Ensuite nous allons simuler m trajectoires de T périodes des risques financiers et catastrophes. Enfin, nous déterminerons le prix du CAT Bond par une approche de simulation de Monte Carlo.

• **Calcul du prix sous l'approche Deep Learning:**

Sous cette approche, on utilisera l'expression du prix avec l'espérance $E^{\mathbb{P}}[F_{CAT}(k)]$. Cette sera obtenue en utilisant différentes approches.

Pour les modèles de classification, $E^{\mathbb{P}}[F_{CAT}(k)]$ sera obtenue à partir de la définition de la fonction espérance pour une variable discrète .

$$E^{\mathbb{P}}[F_{CAT}(X_k)] = \sum_{k=1}^n \mathbb{P}(A_k^{[i]}) E^{\mathbb{P}}[F_{CAT}(X_k) | A_k^{[i]}]$$

Où $A_k^{[i]}$ représente l'événement « la variable X_k appartient à la i -ème tranche à la période k » et $\mathbb{P}(A_k^{[i]})$ représente la probabilité de sortie de la i -ème classe et n le nombre de classe.

Par conséquent, on pourrait approcher le prix V_{CAT} comme une valeur actualisée des flux en sortie de modèle.

$$V_{CAT}(0) = \sum_{i=1}^n \sum_{k=1}^T B(k) \mathbb{P}(A_k^{[i]}) E^{\mathbb{P}}[F_{CAT}(X_k) | A_k^{[i]}]$$

Avec $B(k) = E^{\mathbb{Q}}\left[\frac{1}{\prod_{j=1}^k (1+r_{j-1})}\right]$

Dans le cas des modèles des régression, la sortie du modèle sera utilisée comme l'espérance $E^{\mathbb{P}}[F_{CAT}(k)]$, on notera :

$$E^{\mathbb{P}}[F_{CAT}(X_k)] = F_{CAT}(\widehat{X}_k)$$

Où \widehat{X}_k est la prédiction du modèle pour la période k

On en déduit :

$$V_{CAT}(0) = \sum_{k=1}^T B(k) F_{CAT}(\widehat{X}_k)$$

3.1.4. Les hypothèses sur la structure du contrat de CAT Bond

Dans le cadre de cette étude, les variables ci-dessous seront définies :

- Le taux d'intérêt,
- La prime payée par le sponsor,

On notera :

D_k : Il n'y a eu aucun déclenchement du CAT Bond entre k et $k + 1$

\bar{D}_k : Il y a eu un événement catastrophe entre la période k et $k + 1$

On utilisera comme hypothèses, les caractéristiques du contrat de CAT Bond suivantes :

- Les paiements S_k à la suite d'événements catastrophes entre la période k et $k+1$

$$S_k = \begin{cases} 0 & \text{si } A_k^{[0]} \\ \min[\alpha_1 P ; P_k] & \text{si } A_k^{[1]} \\ \min[\alpha_2 P ; P_k] & \text{si } A_k^{[2]} \\ \vdots & \vdots \\ \min[\alpha_n P ; P_k] & \text{si } A_k^{[n]} \end{cases}$$

Où $A_k^{[0]}$ représente l'événement aucun seuil de déclenchement a été atteint.

- On notera P_k le reste du capital restant à la fin de la période k :

$$P_k = \begin{cases} P_{k-1} - S_k 1_{D_k} & k = 1, \dots, T - 1 \\ P & k = 0 \end{cases}$$

- On notera C_k , le flux de coupons entre k et $k+1$

On supposera que les coupons dépendent uniquement :

- o du principal restant P_k ,
- o du taux de coupon fixe c ,
- o du spread δ correspondant à la prime payée par le sponsor.

$$C_k = (c + \delta) [P_{k-1} 1_{\bar{D}_k} + (P_{k-1} - S_k) 1_{D_k}]$$

Sachant que $P_k = P_{k-1} - S_k 1_{D_k}$

On pourrait en déduire une expression de C_k telle que :

$$C_k = (c + \delta) [P_{k-1} 1_{\bar{D}_k} + P_k 1_{D_k}]$$

- Les flux de cash-flow pour l'investisseur entre k et $k + 1$

$$F_{CAT}(X_k) = \begin{cases} C_k, & k = 1, \dots, T - 1 \\ C_T + P_T, & k = T \end{cases}$$

En remplaçant chaque terme de l'expression précédente par son expression détaillée, on a :

$$F_{CAT}(X_k) = \begin{cases} (c + \delta) [P_{k-1} 1_{\bar{D}_k} + P_k 1_{D_k}], & k = 1, \dots, T - 1 \\ (c + \delta) [P_{T-1} 1_{\bar{D}_T} + P_T 1_{D_T}] + P_T, & k = T \end{cases}$$

Dans les sections suivantes, nous aborderons les hypothèses de modélisation des risques catastrophes et financiers.

3.1.5. Hypothèses de modélisation des risques financiers

Dans cette étude, nous supposons que les obligations sont soumises à un seul risque financier qui sont les taux d'intérêts. Nous serons donc amenés à modéliser les taux d'intérêts dans cette étude. Dans cette partie, nous présenterons quelques modèles de modélisation de la structure des taux d'intérêts.

Le modèle de Vasicek

- Description du modèle

Le modèle de Vasicek (1977) est un modèle de taux court à un seul facteur. Ce modèle prend comme hypothèse que le taux court $r(t)$ suit un processus d'Ornstein-Uhlenbeck (voir appendice xxx).

Dans ce modèle, le taux court $r(t)$ suit une dynamique :

$$dr_t = a(b - r_t) dt + \sigma dW_t$$

avec :

a : une constante positive qui représente la force de rappel de retour à la moyenne

b : une constante positive qui définit le taux d'intérêt de long terme

σ : une constante positive qui représente la volatilité du taux d'intérêt r_t

W_t est un mouvement brownien \mathcal{F}_t - adapté

La solution de l'EDS de r_t , pour $t > s$, peut s'écrire sous la forme⁸ :

$$r_t = r_s e^{-a(t-s)} + b(1 - e^{-a(t-s)}) + \sigma \int_s^t e^{-a(t-u)} dW_u$$

r_t suit donc une loi normale dont l'espérance et la variance peuvent s'écrire sous la forme :

$$\begin{aligned} E[r_{t+\delta} | \mathcal{F}_t] &= E \left[r(t) e^{-a\delta} + b(1 - e^{-a\delta}) + \sigma \int_t^{t+\delta} e^{-a(t-u)} dW_u \middle| \mathcal{F}_t \right] \\ &= r(t) e^{-a\delta} + b(1 - e^{-a\delta}) + \sigma E \left[\int_t^{t+\delta} e^{-a(t-u)} dW_u \middle| \mathcal{F}_t \right] \end{aligned}$$

Et comme W suit un mouvement brownien, on a $E \left[\int_t^{t+\delta} e^{-a(t-u)} dW_u \middle| \mathcal{F}_t \right] = 0$

Ce qui implique que : $E[r_{t+\delta} | \mathcal{F}_t] = r(t) e^{-a\delta} + b(1 - e^{-a\delta})$

Et $\lim_{t \rightarrow \infty} E[r_{t+\delta} | \mathcal{F}_t] = b$ d'où l'appellation du paramètre b par le taux d'intérêt long terme.

$$\begin{aligned} Var[r_{t+\delta} | \mathcal{F}_t] &= Var \left[r(t) e^{-a\delta} + b(1 - e^{-a\delta}) + \sigma \int_t^{t+\delta} e^{-a(t-u)} dW_u \middle| \mathcal{F}_t \right] \\ &= \sigma^2 Var \left[\int_t^{t+\delta} e^{-a(t-u)} dW_u \middle| \mathcal{F}_t \right] \end{aligned}$$

On retrouve la forme de l'intégrale de Wiener dans le terme de la variance $\int_s^t f(u) dW_u$, on a :

$$Var \left[\int_t^{t+\delta} f(u) dW_u \middle| \mathcal{F}_s \right] = E \left[\int_t^{t+\delta} f(u)^2 du \right]$$

Ce qui implique :

$$Var[r_{t+\delta} | \mathcal{F}_t] = \sigma^2 E \left[\int_t^{t+\delta} (e^{-a(t-u)})^2 du \right]$$

On en déduit que :

⁸ Pour retrouver la solution, il suffit d'appliquer le lemme d'Ito en posant $Y_t = e^{at}(r_t - b)$

$$Var[r_{t+\delta}|\mathcal{F}_t] = \frac{\sigma^2}{2a} (1 - e^{-2a\delta})$$

On peut constater que la variance de r_t converge vers $\frac{\sigma^2}{2a}$ quant t tend vers l'infini.

$$\lim_{t \rightarrow \infty} Var[r_{t+\delta}|\mathcal{F}_t] = \frac{\sigma^2}{2a}$$

Le prix du zéro coupon obtenu est le suivant :

$$P(t, T) = A(t, T) e^{-B(t, T)r(t)}$$

Avec

$$\begin{cases} A(t, T) = \exp\left((B(t, T) - (T - t)) \left(b - \frac{\sigma^2}{2a^2} \right) - \frac{\sigma^2}{4a} B(t, T)^2 \right) \\ B(t, T) = \frac{1 - e^{-a(T-t)}}{a} \end{cases}$$

Le modèle de Vasicek n'exclut pas les valeurs négatives de taux. Ce qui dans le contexte actuel de taux négatif dans certaines économies ne pourrait être présenté comme un désavantage uniquement. On observe aussi un effet de retour à la moyenne du taux court r_t vers le taux long terme b dans ce modèle.

De plus, l'un des principaux avantages du modèle de Vasicek est qu'il est un modèle endogène c'est-à-dire que la courbe des taux peut être entièrement expliquée par les variables du modèle.

Le modèle CIR (COX, Ingersoll and Ross)

- **Description du modèle**

Ce modèle de taux développé par Cox, Ingersoll et Ross (1985) est aussi un modèle de taux court à un seul facteur.

Dans ce modèle, le taux court $r(t)$ suit une dynamique :

$$dr_t = a(b - r_t) dt + \sigma \sqrt{r_t} dW_t$$

avec :

a : une constante positive qui représente la force de rappel de retour à la moyenne

b : une constante positive qui définit le taux d'intérêt de long terme

$\sigma\sqrt{r_t}$: représente la volatilité du taux d'intérêt r_t

W_t est un mouvement brownien \mathcal{F}_t - adapté

Dans ce modèle, comme dans celui de Vasicek, on observe aussi un effet de retour à la moyenne du taux court r_t vers b .

La volatilité du taux d'intérêt dépend de la racine carrée du taux d'intérêt. Ceci entraîne le fait que ce modèle ne peut être utilisé pour modéliser les taux d'intérêts négatifs.

r_t suit une loi de χ^2 . Le prix du zéro coupon :

$$P(t, T) = A(t, T) e^{-B(t, T)r(t)}$$

Avec

$$\begin{cases} A(t, T) = \left(\frac{2\gamma e^{\frac{a+\gamma}{2}(T-t)}}{2\gamma + (a + \gamma)(e^{\gamma(T-t)} - 1)} \right)^{\frac{2ab}{\sigma^2}} \\ B(t, T) = \frac{2(e^{\gamma(T-t)} - 1)}{2\gamma + (a + \gamma)(e^{\gamma(T-t)} - 1)} \\ \gamma = \sqrt{a^2 + 2\sigma^2} \end{cases}$$

3.1.6. Hypothèses de modélisation des risques catastrophes

L'objectif de notre étude est de proposer deux approches de valorisation des CAT Bond : une approche de modélisation par une loi de probabilité et une approche Deep Learning.

Dans un premier temps, nous modéliserons les variables risques catastrophes selon différentes approches actuarielles.

Dans cette approche, nous essaierons de simuler les lois des variables des risques catastrophes. Ensuite, nous étudierons une modélisation par approche Deep Learning des risques catastrophes.

Dans notre approche, nous faisons appel aux théories actuarielles de modélisation des risques catastrophes suivantes :

- **L'approche de la théorie des valeurs extrêmes**

Les CAT Bond paramétriques ont pour déclencheur l'atteinte ou le dépassement d'un seuil de variables paramétriques tels que les indices climatiques. De plus, les événements catastrophes sont définis comme des scénarii extrêmes dont la probabilité d'occurrence est faible mais qui peuvent engendrer des pertes considérables. Nous pensons qu'il serait opportun d'étudier la modélisation des risques catastrophes en utilisant certains principes de la théorie des valeurs extrêmes.

Cette approche rejoint celle de Zimbidis, Frangos et Pantelous [2007] qui ont utilisé cette approche TVE pour modéliser la variable « Magnitude » dans la modélisation des séismes sur les côtes grecques.

Nous allons décrire dans cette section une approche de modélisation par la théorie des valeurs extrêmes.

L'un des principaux concepts de la théorie des valeurs extrêmes est le théorème de Fisher–Tippett–Gnedenko (connue aussi sous le nom du théorème de la valeur extrême ou simplement théorème de Fisher–Tippett).

Théorème de Fisher–Tippett :

Soit $\{X_i\}_{i=1,\dots,n}$ n variables aléatoires iid. Soit $M_n = \max(X_1, \dots, X_n)$

S'il existe une séquence de nombres réels (a_n, b_n) tel que :

$$\begin{cases} a_n > 0 \\ \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - a_n}{b_n} \leq x\right) = G(x) \text{ où } G \text{ est une fonction de répartition non dégénérée} \end{cases}$$

Alors G suit une loi $GEV(\mu, \sigma, \xi)$ (Generalized Extreme Value ou loi d'extremum généralisée).

On peut donc écrire G sous la forme :

$$\begin{cases} G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}, & \text{si } \xi \neq 0 \\ G(x) = \exp\left\{-\exp\left[-\left(\frac{x - \mu}{\sigma}\right)\right]\right\}, & \text{si } \xi = 0 \end{cases}$$

- un paramètre d'échelle $\sigma > 0$,
- un paramètre de forme $\xi \in \mathbb{R}$,
- un paramètre de position $\mu \in \mathbb{R}$
- La fonction G est définie sous les conditions suivantes : $1 + \xi\left(\frac{x - \mu}{\sigma}\right) > 0$

On remarquera que :

$GEV(0, 1, 0) = \exp\{-\exp[-x]\}$ est une distribution de Gumbel

$GEV\left(-1, \frac{1}{\alpha}, -\frac{1}{\alpha}\right) = \exp\{-(-x)^\alpha\}$ est une distribution de Weibull

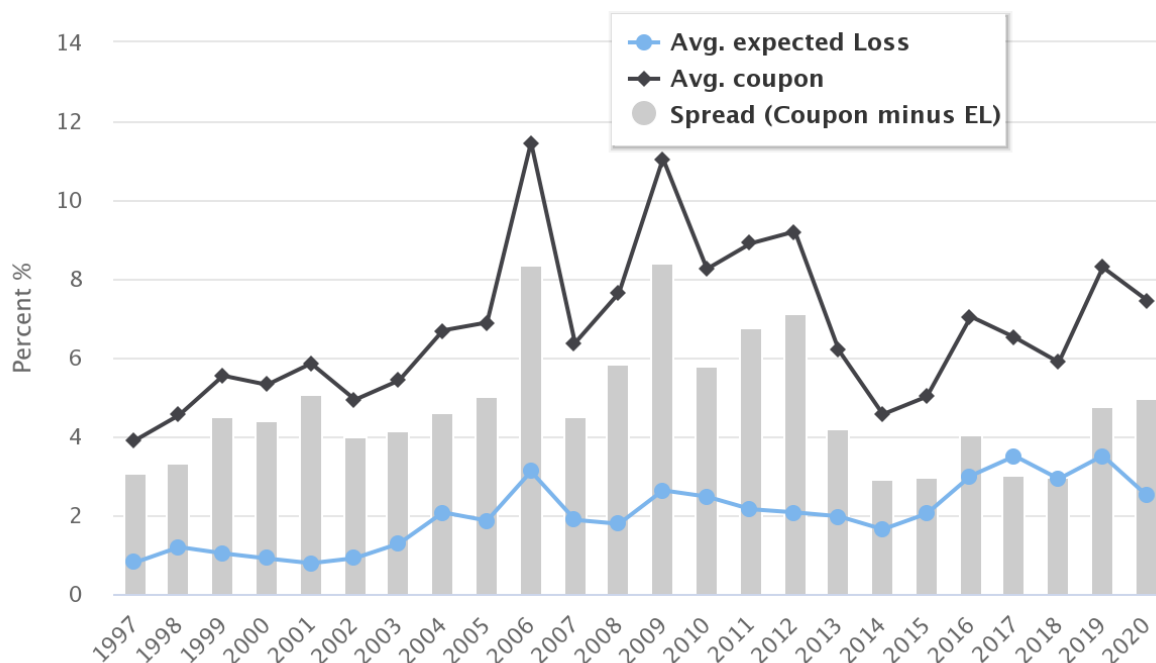
$GEV\left(1, \frac{1}{\alpha}, \frac{1}{\alpha}\right) = \exp\{-x^{-\alpha}\}$ est une distribution de Fréchet

Le principe de modélisation GEV consiste donc à estimer les 3 paramètres de la loi GEV.

3.1.7. Hypothèses sur les coupons et spread

Dans le cadre de cette étude, les données moyennes sur les coupons et les spreads seront utilisées comme respectivement le taux de coupon et le spread payé par l'assureur.

En principe, le spread est estimé à partir de la perte attendue (Expected Loss) et d'un spread additionnel qui peut être considéré comme prime de risque payée par l'assureur à l'investisseur pour accepter le risque du CAT Bond.



Source: www.Artemis.bm Deal Directory

Figure 12 - Evolution des coupons et spread

La figure 7 représente les données issues de la base de transactions d'Artemis⁹ qui montrent l'évolution des coupons et des spread payés pour les titres de catastrophes depuis 1997. Le spread payé par les assureurs avait atteint 8.42% en 2009 pour un coupon de 11% et une perte attendue de 2.63%. Entre 2013 et 2020, les spreads ont une moyenne aux alentours de 3.5%.

Dans notre étude, nous allons considérer un spread de 3.5% qui est la moyenne sur ces 7 dernières années comme hypothèse de modélisation.

De même, nous considérerons que le coupon additionnel qui est la différence entre le coupon et le spread payé par l'assureur sera de 3%. Elle représente la perte attendue ou « Expected Loss » pour une obligation CAT Bond.

⁹ <https://www.artemis.bm/dashboard/cat-bonds-ils-expected-loss-coupon/>

3.2. Modélisation des risques financiers

Comme spécifié dans la section des hypothèses, nous supposons que les obligations sont soumises au seul risque financier qui sont les taux d'intérêts. Aussi, nous allons opter pour le modèle de Vasicek pour la modélisation des taux d'intérêts. Ce choix s'inscrit dans une démarche de simplification de notre étude. Dans le modèle de Vasicek, nous pouvons avoir l'expression de r_t comme démontré dans la section 3.1.5:

$$\text{On a : } r_t = r_s e^{-a(t-s)} + b(1 - e^{-a(t-s)}) + \sigma \int_s^t e^{-a(t-u)} dW_u$$

Dans cette partie, on va d'abord décrire des approches de calibration des paramètres d'un modèle de Vasicek et ensuite proposer une approche de simulation de taux d'intérêts.

3.2.1. Méthodes de calibration des paramètres du modèle de taux d'intérêt

La première étape de la construction de la courbe des taux dans le modèle Vasicek est la calibration des trois paramètres du modèle a, b, σ . Plusieurs approches de calibration existent. Nous allons présenter quelques approches dans notre étude:

On suppose une séquence de n observations $\{r_{t_i}\}_{i=1,\dots,n}$ avec l'hypothèse que le pas entre deux périodes successives t_i et t_{i+1} est de δ .

- **L'approche MCO (estimateur des moindres carrés ordinaires)**

Dans cette approche de calibration, on va prendre comme hypothèse que les taux d'intérêt suivent un modèle linéaire.

$$r_{t_i} = \alpha r_{t_{i-1}} + \beta + \varepsilon_{t_i}$$

L'objectif de cette approche est de déterminer les coefficients α, β et σ_ε la variance des erreurs avec la régression des moindres carrés en utilisant les observations historiques du taux d'intérêt à modéliser.

Nous savons que les estimateurs MCO des paramètres $\alpha, \beta, \varepsilon_{t_{i+1}}$ sont données par les formules :

$$\hat{\alpha} = \frac{\sum_{i=1}^n r_{t_{i+1}} r_{t_i} - \frac{1}{n} \left(\sum_{i=1}^n r_{t_{i+1}} \sum_{i=1}^n r_{t_i} \right)}{\sum_{i=1}^n r_{t_i}^2 - \frac{1}{n} \left(\sum_{i=1}^n r_{t_i} \right)^2}$$

$$\hat{\beta} = \frac{1}{n} \left(\sum_{i=1}^n r_{t_{i+1}} - \hat{\alpha} \sum_{i=1}^n r_{t_i} \right)$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n-2} \sum_{i=1}^n [r_{t_i} - (\hat{\alpha} r_{t_{i-1}} + \hat{\beta})]^2$$

Cas 1 : approche par discrétisation exacte de r_t

On sait que :

$$r_t = r_s e^{-a(t-s)} + b(1 - e^{-a(t-s)}) + \sigma \int_s^t e^{-a(t-u)} dW_u$$

avec $\int_0^t e^{-a(t-u)} dW_u \sim \mathcal{N} \left(0, \frac{1-e^{-2at}}{2a} \right)$

On pourrait donc utiliser la discrétisation suivante pour approcher les valeurs de r_t :

$$r_{t+\delta} = r_t e^{-a\delta} + b(1 - e^{-a\delta}) + \varepsilon_t$$

Où $\varepsilon_t \sim \mathcal{N}\left(0, \sigma^2 \left(\frac{1 - e^{-2a\delta}}{2a}\right)\right)$

On pourrait en déduire que l'expression de r_t peut s'écrire de la forme

$$r_{t+\delta} = \alpha r_t + \beta + \varepsilon_t$$

avec

$$\alpha = e^{-a\delta}$$

$$\beta = b(1 - e^{-a\delta})$$

$$\sigma_\varepsilon^2 = \sigma^2 \left(\frac{1 - e^{-2a\delta}}{2a}\right)$$

Les estimateurs des paramètres seront déduits des relations ci-dessous :

$$\hat{a} = -\frac{\ln \hat{\alpha}}{\delta}$$

$$\hat{b} = -\frac{\hat{\beta}}{(1 - e^{-\hat{a}\delta})}$$

$$\hat{\sigma}^2 = \hat{\sigma}_\varepsilon^2 \frac{2\hat{a}}{1 - e^{-2\hat{a}\delta}}$$

Cas 2 : calibration avec le schéma de Euler de r_t

La dynamique de Vasicek est donnée par :

$$dr_t = a(b - r_t)dt + \sigma dW_t$$

On pourrait donc en déduire une approximation r_t en utilisant schéma d'Euler :

$$\Delta r_t = a(b - r_t)\Delta t + \sigma \Delta W_t$$

Ainsi pour un pas $\Delta t = \delta$, on aura :

$$r_{t+\delta} - r_t \approx a(b - r_t)\delta + \sigma (W_{t+\delta} - W_t)$$

$$\text{avec } W_{t+\delta} - W_t \sim \mathcal{N}(0, \delta)$$

Par conséquent,

$$r_{t+\delta} = (1 - a\delta)r_t + ab\delta + \sigma\sqrt{\delta} \mathcal{N}(0, 1)$$

On pourrait en déduire que l'expression de la discrétisation de r_t peut s'écrire de la forme

$$r_{t+\delta} = \alpha r_t + \beta + \varepsilon_t$$

Avec $\alpha = 1 - a\delta$; $\beta = ab\delta$; $\hat{\sigma}_\varepsilon^2 = \sigma^2 \delta$

Les estimateurs des paramètres du modèle de Vasicek sont ainsi déduits :

$$\hat{a} = \frac{1 - \hat{\alpha}}{\delta}$$

$$\hat{b} = -\frac{\hat{\beta}}{1 - \hat{a}} \text{ en remplaçant } \hat{a} \text{ par son expression}$$

$$\hat{\sigma}^2 = \frac{\hat{\sigma}_\varepsilon^2}{\delta}$$

- **L'approche EMV (estimation par maximum de vraisemblance)**

Dans un modèle de Vasicek, la variable aléatoire r_t suit une loi normale.

On a :

$$r_t | \mathcal{F}_s \sim \mathcal{N} \left(r(s) e^{-a(t-s)} + b(1 - e^{-a(t-s)}); \sigma_1^2 \right)$$

$$\text{avec } \sigma_1 = \frac{\sigma^2}{2a} (1 - e^{-2a(t-s)})$$

On pourrait en déduire la fonction de densité f de $r_t | \mathcal{F}_s$:

$$f(r_t | \mathcal{F}_s) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left\{ -\frac{[r_t - r(s) e^{-a(t-s)} - b(1 - e^{-a(t-s)})]^2}{2\sigma_1^2} \right\}$$

On suppose une séquence de n observations $\{r_{t_i}\}_{i=1, \dots, n}$ d'une suite de variables aléatoires $\{r_i\}_{i=1, \dots, n}$ suivant une distribution normale de fonction de densité f .

On a :

$$f(r_{t_i} | \mathcal{F}_{t_{i-1}}) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left\{ -\frac{[r_{t_i} - r_{t_{i-1}} e^{-a\delta} - b(1 - e^{-a\delta})]^2}{2\sigma_1^2} \right\}$$

On peut donc exprimer la vraisemblance : $L(r_{t_1}, \dots, r_{t_n}; a, b, \sigma_1) = \prod_{i=1}^n f(r_{t_i} | \mathcal{F}_{t_{i-1}})$

On en déduit la log-vraisemblance :

$$\ln L(r_{t_1}, \dots, r_{t_n}; a, b, \sigma_1) = \sum_{i=1}^n \ln f(r_{t_i} | \mathcal{F}_{t_{i-1}})$$

On trouve :

$$\ln L(r_{t_1}, \dots, r_{t_n}; a, b, \sigma_1) = \sum_{i=1}^n -\ln(\sigma_1 \sqrt{2\pi}) - \sum_{i=1}^n \frac{[r_{t_i} - r_{t_{i-1}} e^{-a\delta} - b(1 - e^{-a\delta})]^2}{2\sigma_1^2}$$

En simplifiant cette expression, on a :

$$\ln L(r_{t_1}, \dots, r_{t_n}; a, b, \sigma_1) = -n \ln(\sigma_1 \sqrt{2\pi}) - \sum_{i=1}^n \frac{[r_{t_i} - r_{t_{i-1}} e^{-a\delta} - b(1 - e^{-a\delta})]^2}{2\sigma_1^2}$$

L'estimateur du maximum de vraisemblance est l'estimateur qui maximise la fonction de vraisemblance sur les observations :

$$\ln L(r_{t_1}, \dots, r_{t_n}; \hat{a}, \hat{b}, \hat{\sigma}_1) = \max_{(\mu, \sigma_1, \kappa)} \ln L(r_{t_1}, \dots, r_{t_n}; a, b, \sigma_1)$$

On obtient les estimateurs suivants des paramètres a, b, σ_1 :

$$\hat{b} = \frac{1}{n} \sum_{i=1}^n \frac{[r_{t_i} - r_{t_{i-1}} e^{-a\delta}]}{1 - e^{-a\delta}}$$

$$\hat{a} = -\frac{1}{\delta} \ln \left[\frac{\sum_{i=1}^n (r_{t_i} - b)(r_{t_{i-1}} - b)}{\sum_{i=1}^n (r_{t_{i-1}} - b)^2} \right]$$

$$\hat{\sigma}_1 = \frac{1}{n} \sum_{i=1}^n [r_{t_i} - r_{t_{i-1}} e^{-a\delta} - b(1 - e^{-a\delta})]^2$$

On déduit de $\hat{\sigma}_1$ l'expression de

$$\hat{\sigma} = \frac{1}{n} \frac{2\hat{a}}{1 - e^{-2\hat{a}\delta}} \sum_{i=1}^n [r_{t_i} - r_{t_{i-1}} e^{-\hat{a}\delta} - \hat{b}(1 - e^{-\hat{a}\delta})]^2$$

En substituant a par son estimateur \hat{a} dans l'expression de l'estimateur de \hat{b} , on obtient :

$$\hat{b} = \frac{\sum_{i=1}^n r_{t_i} \sum_{i=1}^n r_{t_{i-1}}^2 - \sum_{i=1}^n r_{t_{i-1}} \sum_{i=1}^n r_{t_{i-1}} r_{t_i}}{n \left(\sum_{i=1}^n r_{t_{i-1}}^2 - \sum_{i=1}^n r_{t_{i-1}} r_{t_i} \right) + \sum_{i=1}^n r_{t_{i-1}} \left(\sum_{i=1}^n r_{t_i} - \sum_{i=1}^n r_{t_{i-1}} \right)}$$

3.2.2. Simulations de trajectoires de taux d'intérêt

Dans le cas de la discrétisation exacte, nous allons utiliser l'expression de $r(t)$ obtenue de l'EDS. On sait que

$$r_t = r_s e^{-a(t-s)} + b(1 - e^{-a(t-s)}) + \sigma \int_s^t e^{-a(t-u)} dW_u$$

Avec $\int_0^t e^{-a(t-u)} dW_u \sim \mathcal{N} \left(0, \frac{1 - e^{-2at}}{2a} \right)$

$$\Rightarrow \sqrt{\frac{2a}{1 - e^{-2at}}} \int_0^t e^{-a(t-u)} dW_u \sim \mathcal{N}(0,1)$$

On pourrait utiliser la discrétisation suivante pour simuler les valeurs de r_t :

$$r_{t+\delta} = r_t e^{-a\delta} + b(1 - e^{-a\delta}) + \sigma \sqrt{\frac{1 - e^{-2a\delta}}{2a}} \varepsilon_t \text{ avec } \varepsilon_t \sim \mathcal{N}(0,1)$$

Dans le cas du schéma d'Euler, nous allons utiliser l'expression approchée de $r(t)$:

$$r_{t+\delta} = r_t + a(b - r_t)\delta + \sigma\sqrt{\delta} \mathcal{N}(0,1)$$

La simulation de r_t consistera donc à simuler m trajectoires de loi normale centrée réduite $\{\varepsilon_i\}_{i=1, \dots, n}$.

3.3. Modélisation CAT Bond : approche de modélisation des risques catastrophes par une loi

3.3.1. Méthodes de calibration des paramètres d'une loi GEV

Nous allons utiliser l'approche d'estimation par maximum de vraisemblance.

On suppose une séquence de n observations $\{m_i\}_{i=1,\dots,n}$ d'une suite de variables aléatoires $\{M_i\}_{i=1,\dots,n}$ suivant une distribution $GEV(\mu, \sigma, \xi)$ de fonction de densité g qui est la dérivée de la fonction de répartition G.

On pourra donc exprimer la vraisemblance L dans le cas $\xi \neq 0$ comme suit :

$$L(m_1, \dots, m_n; \mu, \sigma, \xi) = \prod_{i=1}^n g(m_i)$$

$$\text{avec } g(m_i) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right]^{-\left(\frac{1}{\xi} + 1\right)} \exp \left\{ - \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}$$

La condition d'existence de la fonction de répartition de G s'applique aussi à g pour toute valeur m_i :

$$1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) > 0 \quad i = 1, \dots, n$$

Cette condition d'existence de la fonction GEV va permettre d'avoir une contrainte forte entre les trois paramètres à estimer.

En remplaçant g par son expression, on obtient :

$$L(m_1, \dots, m_n; \mu, \sigma, \xi) = \frac{1}{\sigma^n} \exp \left\{ - \sum_{i=1}^n \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \prod_{i=1}^n \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right]^{-\left(\frac{1}{\xi} + 1\right)}$$

Pour simplifier cette expression, on pourrait utiliser la log-vraisemblance :

$$\ln L(m_1, \dots, m_n; \mu, \sigma, \xi) = n \ln \sigma - \sum_{i=1}^n \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} - \left(\frac{1}{\xi} + 1 \right) \sum_{i=1}^n \ln \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right]$$

L'estimation par maximum de vraisemblance consistera à trouver des valeurs de μ, σ, ξ qui maximise la vraisemblance (ou log-vraisemblance)

$$\ln L(m_1, \dots, m_n; \hat{\mu}, \hat{\sigma}, \hat{\xi}) = \max_{(\mu, \sigma, \xi)} \ln L(m_1, \dots, m_n; \mu, \sigma, \xi)$$

En annulant les dérivées partielles de $\ln L(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ par rapport à chaque paramètre, on peut remarquer qu'il n'y a pas une expression explicite des estimateurs dans cette approche.

Nous allons donc utiliser des approches d'approximations de ces valeurs en utilisant la bibliothèque R «exTremes» .

3.3.2. Méthode de simulation de la GEV

Une loi GEV est définie par la fonction de répartition suivante

$$\begin{cases} G(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{-1/\xi} \right\}, & \text{si } \xi \neq 0 \\ G(x) = \exp \left\{ - \exp \left[- \left(\frac{x-\mu}{\sigma} \right) \right] \right\} & , \text{ si } \xi = 0 \end{cases}$$

avec $1 + \xi \left(\frac{x-\mu}{\sigma} \right) > 0$

L'inverse de la fonction de répartition de G est obtenu de façon explicite.

$$x = \begin{cases} \mu + \frac{\sigma}{\xi} \left[(-\ln(1 - G(x)))^{-\xi} - 1 \right], & \text{si } \xi \neq 0 \\ \mu - \sigma \ln(-\ln(1 - G(x))), & \text{si } \xi = 0 \end{cases}$$

il pourrait donc être proposé une approche de simulation suivante :

Soit $z \sim \mathcal{U}(0, 1)$ une loi uniforme [0,1]

Soit $\{M_i\}_{i=1,\dots,n}$ n simulation de variables aléatoires suivant une loi GEV avec :

$$m_i = \begin{cases} \mu + \frac{\sigma}{\xi} \left[(-\ln(1 - z))^{-\xi} - 1 \right], & \text{si } \xi \neq 0 \\ \mu - \sigma \ln(-\ln(1 - z)), & \text{si } \xi = 0 \end{cases}$$

3.3.3. Valorisation CAT Bond

Le prix du CAT Bond est obtenu par une simulation de Monte Carlo.

A cette étape, à partir des méthodes de simulations des risques financiers et catastrophe définies dans les sections précédentes, il va être généré k trajectoires indépendantes de T années du modèle de Vasicek pour le risque financier et de la loi GEV pour les risques catastrophes.

Le prix du CAT Bond sera défini comme la moyenne des simulations :

$$V_{CAT}(0) \approx \frac{1}{m} \sum_{i=1}^m \sum_{k=1}^T \frac{1}{\prod_{j=1}^k (1 + r_{j-1}^{(i)})} F_{CAT}(X_k^{(i)})$$

3.5. Modélisation CAT Bond : Approche de modélisation des variables des risques catastrophes par une approche Deep Learning

Cette section a pour principal objectif de proposer une approche de modélisation des risques catastrophes par une approche Deep Learning. Nous nous baserons sur des modèles séquentiels de Deep Learning pour modéliser les trajectoires des variables des risques catastrophes.

Cette partie sera subdivisée en 5 grandes parties. D'abord, nous allons décrire les modèles. Ensuite, le processus de construction de notre base de données sera présenté. Aussi, nous parlerons des modèles RNN utilisés pour notre modèle et des métriques d'évaluation du modèle. Enfin nous décrivons le processus de valorisation du CAT Bond.

L'approche Deep Learning de notre étude suivra les étapes suivantes :

- Etape 1 : Modélisation des risques financiers.
- Etape 2 : Modélisation des risques catastrophes par des approches Deep Learning
- Etape 3 : Simulation de m trajectoires de taux pour définir une courbe de taux.
- Etape 4 : Prédiction de la trajectoire de T période de la variable risque catastrophe à partir du modèle Deep Learning.
- Etape 5 : Evaluation du prix du CAT Bond.

NB : On aurait pu modéliser les risques financiers avec une approche Deep Learning. Nous avons opté pour une approche simulation des risques financiers afin de simplifier la comparaison entre les modèles de loi de probabilité et les modèle Deep Learning.

3.5.1. Description du modèle

L'objectif de l'approche Deep Learning est de déterminer les probabilités de se retrouver dans les différentes tranches du CAT Bond en partant des données historiques qui seront présentées sous formes de séries temporelles.

La problématique de notre modèle est de pouvoir donner une estimation du maximum annuel d'un paramètre climatique ou indice météorologique sur la durée T du CAT Bond.

Une première étape dans la résolution de cette problématique consistera donc à préparer notre base de données d'observations. Pour cela, nous allons définir les paramètres suivants :

- La matrice des données d'apprentissage de m périodes avec n attributs $X =$

$$\begin{bmatrix} x_1^1 & \cdots & x_n^1 \\ \vdots & \ddots & \vdots \\ x_1^m & \cdots & x_n^m \end{bmatrix}$$

avec l'observation de la i -ème période donnée $X^i = [x_1^i \quad \cdots \quad x_n^i]$

- La matrice des valeurs réelles des paramètres à estimer :

$$Y = \begin{bmatrix} y_1^1 & \cdots & y_T^1 \\ \vdots & \ddots & \vdots \\ y_1^m & \cdots & y_T^m \end{bmatrix}$$

avec y_t^i comme la valeur à estimer de la période i pour l'année t du contrat de CAT Bond.

Dans notre étude, les paramètres X^i représenteront les historiques des observations.

La variable x_j^i représente l'observation du paramètre j à la date i et la variable y_t^i représente le maximum annuel observé du paramètre à estimer de l'année t du contrat pour la période

i. Nous allons prendre comme hypothèse que la première colonne de la matrice $X \begin{bmatrix} x_1^1 \\ \vdots \\ x_1^m \end{bmatrix}$ est la variable dont les valeurs maximales constituent les valeurs à prédire.

3.5.2. Construction de la base de données d'apprentissage

La base de données de notre étude sera principalement constituée d'observations historiques de paramètres climatiques liés à des risques catastrophes comme, par exemple, la vitesse du vent, la pression, les précipitations, les relevés de températures, etc.

La base des valeurs à prédire sera construite à partir des observations historiques :

$$y_t^i = \max\left(\{x_1^{i+1+365(t-1)}, \dots ; x_1^{i+365t}\}\right)$$

Deux approches seront testées dans cette étude :

- Une première approche consistant à estimer les valeurs maximales. Une approche Deep Learning de régression sera utilisée pour modéliser cette approche.
- Une seconde approche consistant à prédire les probabilités pour les valeurs maximales à tomber dans les différentes tranches du CAT Bond. Cette approche est une problématique de classification

La convention d'une année de 365 jours sera utilisée comme hypothèse pour la construction des valeurs à prédire.

Cette construction se base sur le fait que notre étude ne s'appuiera pas sur des années calendaires pour définir les maximums mais sur des périodes de 365 jours à partir de la date d'observation des informations. Dans les modèles séquentiels, on cherche à estimer des valeurs futures en partant d'un historique et d'une situation courante.

Dans cette étude, il est primordial de s'affranchir des limites des années calendaires afin de pouvoir constituer une base de données adaptée à une étude Deep Learning.

De plus, cette hypothèse de construction de nos valeurs à prédire se justifie dans le cas des modèles séquentiels car elle permettra de transformer une base de k années de données en $365*(k-1) + 1$ années de données.

3.5.3. Modèles de Deep Learning

Notre étude consistera à prédire des séquences à partir d'une séquence d'observations historiques. Nous allons, pour cela, utiliser des modèles de Deep Learning adaptés aux problématiques temporelles.

Partant du principe que la problématique est la prédiction d'une séquence de valeurs à partir d'observations historiques fournies, nous allons principalement utiliser différentes approches de modèles Deep Learning séquentiels tels que le LSTM et ses dérivés. Nous n'utiliserons pas de modèle GRU car son fonctionnement ne diffère pas de l'approche classique LSTM.

Le modèle LSTM utilisé dans cette partie est le modèle « séquence à séquence ». Il consiste à considérer une séquence d'observations pour prédire une séquence de valeurs futures. Dans ce modèle, on pourrait par exemple, apprendre des deux dernières années pour prédire les 6 prochains mois.

Ce modèle LSTM est souvent appelé « Encoder-Décoder LSTM ». La première partie des séquences, composées de couches LSTM, a pour objectif de comprendre les liens entre les attributs des variables d'observations. La seconde section sert à faire le lien entre ces attributs et les séquences à prédire.

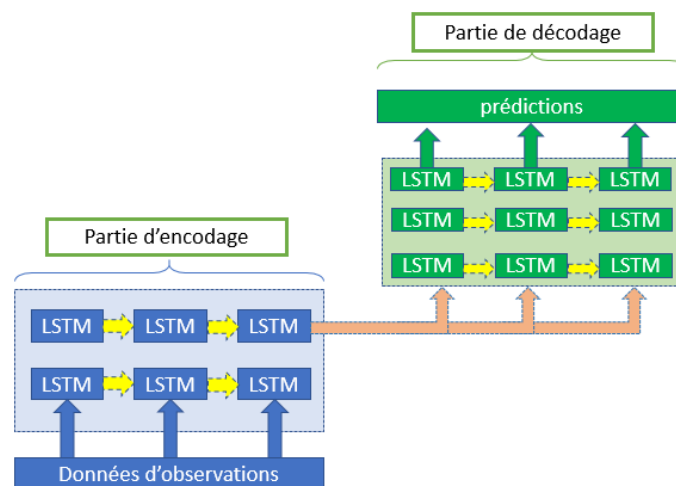


Figure 13 - Modèles LSTM Seq to Seq

Dans ce modèle, l'un des paramètres sera donc la fenêtre d'historique à considérer pour la prédiction des valeurs futures.

Une fois les modèles obtenus à partir des données, l'approche de valorisation du CAT Bond sera décrite dans la section ci-après.

3.5.4. Valorisation du CAT Bond

L'estimation du prix du CAT Bond sera obtenue en suivant le processus décrit dans les hypothèses de modélisation.

En effet, dans cette approche, nous allons estimer les probabilités de survenance des événements des différentes classes de déclenchement du CAT Bond à partir du modèle obtenu par les techniques de Deep Learning.

La formule de calcul de prix d'un CAT Bond dans le modèle réseau de neurones séquentiel sera la suivante comme expliqué dans la section 3.1.3:

- Modèle de classification :

$$V_{CAT}(0) = \sum_{i=1}^n \sum_{k=1}^T B(k) \mathbb{P} \left(A_k^{[i]} \right) E^{\mathbb{P}} \left[F_{CAT}(X_k) \mid A_k^{[i]} \right]$$

- Modèle de régression :

$$V_{CAT}(0) = \sum_{k=1}^T B(t) F_{CAT}(\hat{X}_k)$$

Pour faire suite à ce chapitre de présentation des approches de modélisation, dans la section suivante, il sera proposé une application des deux approches de modélisation sur le risque tempête dans le sud-est des Etats-Unis. Dans cette partie, on supposera que le paramètre de déclenchement du CAT Bond est la vitesse maximale annuelle des rafales de vents.

Partie 4. Application à un CAT Bond à déclencheur paramétrique

4.1. Les caractéristiques du CAT Bond étudié

Pour cette étude, nous modéliserons un CAT Bond sur des événements tempêtes avec comme déclencheur le maximum annuel de la vitesse maximale journalière des rafales de vent dans la région de notre étude. Comme spécifié dans les hypothèses, nous allons supposer que les flux financiers entre les différents acteurs, liés au CAT Bond, ont lieu en fin d'année. Il existe plusieurs indices de mesure des impacts de la vitesse de vent. Nous allons utiliser l'échelle de Saffir-Simpson, généralement utilisée pour la catégorisation des tempêtes tropicales, pour identifier les différentes tranches du CAT Bond de notre étude.

Classe de cyclone	Vents en km/h	Vents en kts
Dépression Tropicale	0 - 62	0 - 33
Tempête Tropicale	63 - 118	34 - 63
Ouragan Cat. 1	119 - 153	64 - 82
Ouragan Cat. 2	154 - 177	83 - 95
Ouragan Majeur Cat. 3	178 - 208	96 - 112
Ouragan Majeur Cat. 4	209 - 251	113 - 136
Ouragan Majeur Cat. 5	plus de 251	plus de 136

Table 2 - Echelle de Saffir-Simpson

Nous allons supposer que le CAT Bond ne se déclenchera qu'à partir de la vitesse du vent similaire à un ouragan de catégorie 1 selon l'échelle de Saffir-Simpson.

On définira donc la structure des seuils de paiements de sinistres comme décrit ci-dessous:

$$S(k) = \begin{cases} 0 & X \leq 63 \text{ kts} \\ \min(\alpha_2 P ; P_k) & 63 < X \leq 82 \text{ kts} \\ \min(\alpha_3 P ; P_k) & 82 < X \leq 95 \text{ kts} \\ \min(\alpha_4 P ; P_k) & 95 < X \leq 112 \text{ kts} \\ \min(\alpha_5 P ; P_k) & 112 < X \leq 136 \text{ kts} \\ \min(\alpha_6 P ; P_k) & X > 136 \text{ kts} \end{cases}$$

Par simplification, nous allons considérer les valeurs des différents coefficients, attribués à chaque tranche de paiement, connues. Dans la pratique, ces coefficients sont estimés à partir de l'historique de sinistralité de la zone étudiée. Au vu de la difficulté à avoir accès à des données libres fiables sur la sinistralité liée aux vitesses journalières des rafales de vents, nous allons prendre comme hypothèse d'étude des facteurs connus.

De toute évidence, ce choix n'aura d'effet que si le prix final du CAT Bond et non sur la méthodologie utilisée pour valoriser un CAT Bond.

Nous allons aussi effectuer des études de sensibilité du prix aux variations des coefficients et/ou des vitesses de déclenchement.

Application à un CAT Bond sur des risques tempêtes

Les valeurs des différentes tranches sont données dans le tableau suivant.

Caractéristiques des CAT Bond étudiés	
Déclencheur	Maximum annuelle de la vitesse journalière du vent
Tranches de sinistralité	Tranche 1 : Pas de déclenchement
	Tranche 2 : Déclenchement avec pour facteur $\alpha_2 = 20\%$
	Tranche 3 : Déclenchement avec pour facteur $\alpha_3 = 40\%$
	Tranche 4 : Déclenchement avec pour facteur $\alpha_4 = 60\%$
	Tranche 5 : Déclenchement avec pour facteur $\alpha_5 = 80\%$
	Tranche 6 : Déclenchement avec pour facteur $\alpha_6 = 100\%$
Primes d'émission	3.5% du principal restant
Capital à l'émission	100 millions
Flux de cash	Fin d'année
Coupons additionnel	3% du principal restant
Maturité	3 ans

Table 3- Caractéristiques du CAT Bond étudié

4.2. Application de la modélisation des risques financiers

4.2.1. Données d'observations des taux d'intérêts

Dans le cadre de cette étude, les observations de taux d'intérêts utilisés sont les taux d'intérêts historiques¹⁰ US 1an qui ont été observés entre 01/2010 et 11/2020.

La figure 15 représente l'évolution des taux d'intérêts US sur cette période. On constate une baisse significative entre la période 01/2020 et 04/2020 qui correspond à la période des confinements à travers le monde résultant de la crise du COVID-19.

Pour la suite de l'étude, nous ne considérons que les données entre les périodes 2017 et 2019.

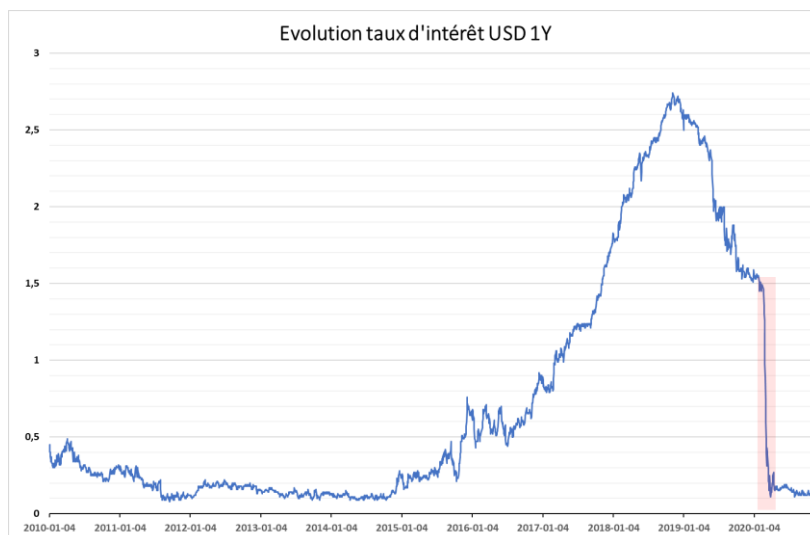


Figure 14 - Evolution journalière du taux US 1 an 2010-2020

¹⁰ Données téléchargées le 12/11/2020 sur le site : <https://www.federalreserve.gov/datadownload/Choose.aspx?rel=H15>

La figure 16 représente la relation entre deux observations successives des taux d'intérêts US 1an. On peut constater sur cette représentation que la relation entre le taux 1an observé entre deux successives pourrait être modélisé par un modèle linéaire.

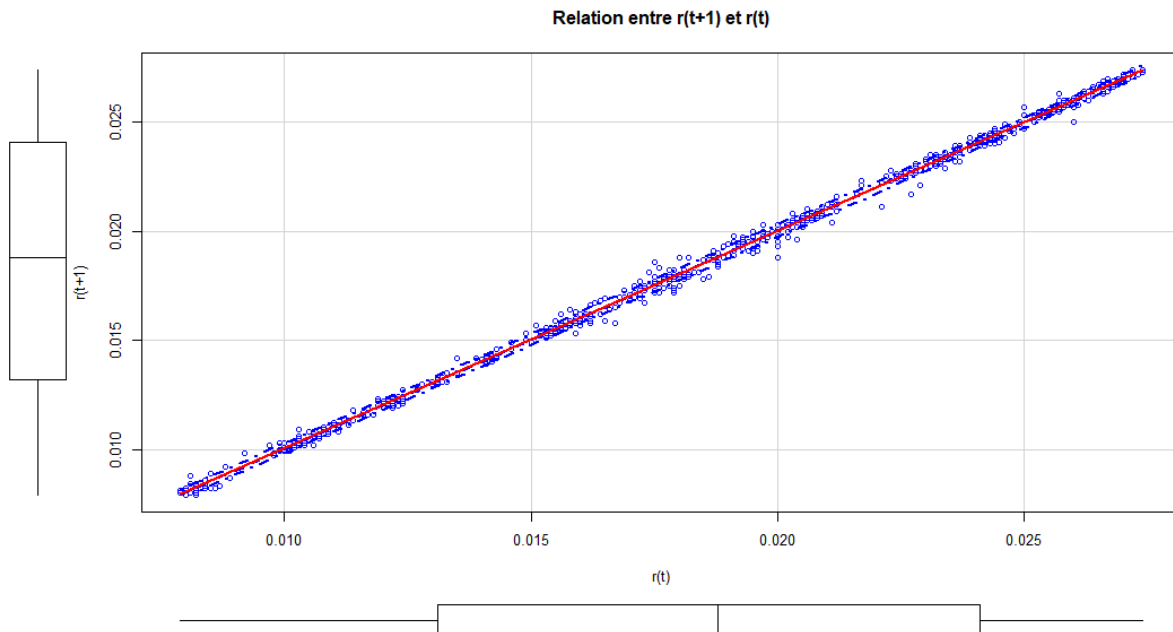


Figure 15 - Relation entre deux observations journalières successives du taux US 1an entre 2017 et 2019.

Nous allons donc estimer les paramètres de la régression linéaire associée. Ces paramètres seront utilisés dans la section de calibration du modèle de risque financier.

Les estimateurs des coefficients de la relation linéaire obtenus des données des taux d'intérêts US 1an sont les suivants :

$$\hat{a} = 9.9723 * 10^{-1}$$

$$\hat{b} = 6.1375 * 10^{-5}$$

$$\hat{\sigma}_\varepsilon = 2.343 * 10^{-4}$$

Dans le but de tester l'hypothèse de nullité de coefficient a, une analyse de la variance va être effectuée dans le tableau ci-dessous. Ce tableau sera utile à la réalisation du test de Fisher. Le tableau d'analyse de la variance obtenu des données est le suivant :

Source de variation	Somme des carrés	Degré de liberté	Carrés moyens
r_{t-1}	SCE = 0.024434	$ddl_{SCE} = 1$	$\frac{SCR}{ddl_{SCE}} = 0.0244342$
Résidu	SCR = 0.000041	$ddl_{SCR} = 747$	$\frac{SCR}{ddl_{SCR}} = 1.0 * 10^{-7}$
Total	SCT = 0.024475	$ddl_{SCT} = 748$	

La statistique de Fisher¹¹ est donnée par la relation : $F^* = \frac{SCE/ddl_{SCE}}{SCR/ddl_{SCR}}$

¹¹ La statique de Fisher est le rapport de la variabilité du modèle expliquée par les variables explicatives sur la somme des carrés des résidus (chaque terme est divisé par son degré de liberté).

On obtient la valeur de la statistique $F^* = 4.452 * 10^5 > F_{1;745}^{0,05} = 3.86$. On peut donc rejeter l'hypothèse consistant à dire que la variable $r(t)$ ne contribue pas à l'explication de $r(t+1)$.

Nous allons utiliser les paramètres obtenus pour la calibration du modèle de Vasicek.

4.2.2. Calibration du modèle de Vasicek sur les données US 1an

Cette partie de l'étude est une application des méthodes de calibration vues dans la partie précédente. Pour la méthode MCO, des tests statistiques seront faits sur la nullité du coefficient a avant toute analyse additionnelle.

Dans le cas de la discrétisation exacte, nous avons obtenu les estimateurs suivants pour les paramètres avec $\delta = 1/360$:

$$\hat{b} = 0.022901$$

$$\hat{a} = 0.966095$$

$$\hat{\sigma} = 0.004451$$

Si l'on considère l'approche par le schéma d'Euler, nous obtenons les valeurs suivantes :

$$\hat{b} = 0.022901$$

$$\hat{a} = 0.9648$$

$$\hat{\sigma} = 0.004446$$

Nous constatons dans les deux cas que les paramètres du modèle de Vasicek sont pratiquement similaires. Nous allons continuer cette étude avec les paramètres de l'approche par le schéma d'Euler. La principale raison de ce choix vient du fait que la simulation de l'approche par le schéma d'Euler nécessite une forme plus simple comparée à l'approche par discrétisation exacte.

4.2.3. Simulation de trajectoires de taux d'intérêts

Nous allons faire une simulation de 100 000 trajectoires indépendantes de taux d'intérêts avec pour point de départ le taux 1 an US au 31/12/2019.

$$r_{t+\delta} = r_t + a(b - r_t)\delta + \sigma\sqrt{\delta} \mathcal{N}(0, 1)$$

En plus des paramètres du modèle de Vasicek obtenus dans la section calibration, nous allons utiliser les paramètres additionnels suivants dans le cadre de la simulation du taux d'intérêt :

- $T = 3$ ans
- $\delta = \frac{T}{360}$
- $r_0 = r_{31/12/2019}$ avec $r_{31/12/2019} = 1.59\%$

L'algorithme de simulation sur R est décrit en annexe (section Algorithmes Python / R).

La table 4 résume les résultats de la simulation. La figure 9 représente un extrait des 100 000 trajectoires obtenues par la simulation de Vasicek. On peut remarquer la convergence de la moyenne (trait noir en pointillé) vers le taux long terme (ligne rouge en pointillé).

Application à un CAT Bond sur des risques tempêtes

	r(1)	r(2)	r(3)
nombre	100 000	100 000	100 000
moyenne	2.02%	2.19%	2.25%
Ecart type	0.30%	0.32%	0.32%
Minimum	0.74%	0.79%	0.93%
25%	1.82%	1.98%	2.04%
50%	2.02%	2.19%	2.25%
75%	2.22%	2.40%	2.47%
90%	2.40%	2.60%	2.66%
95%	2.51%	2.71%	2.78%
99%	2.71%	2.93%	3.00%
Maximum	3.31%	3.70%	3.60%

Table 4 - Résultats de la simulation de la trajectoire du taux d'intérêts

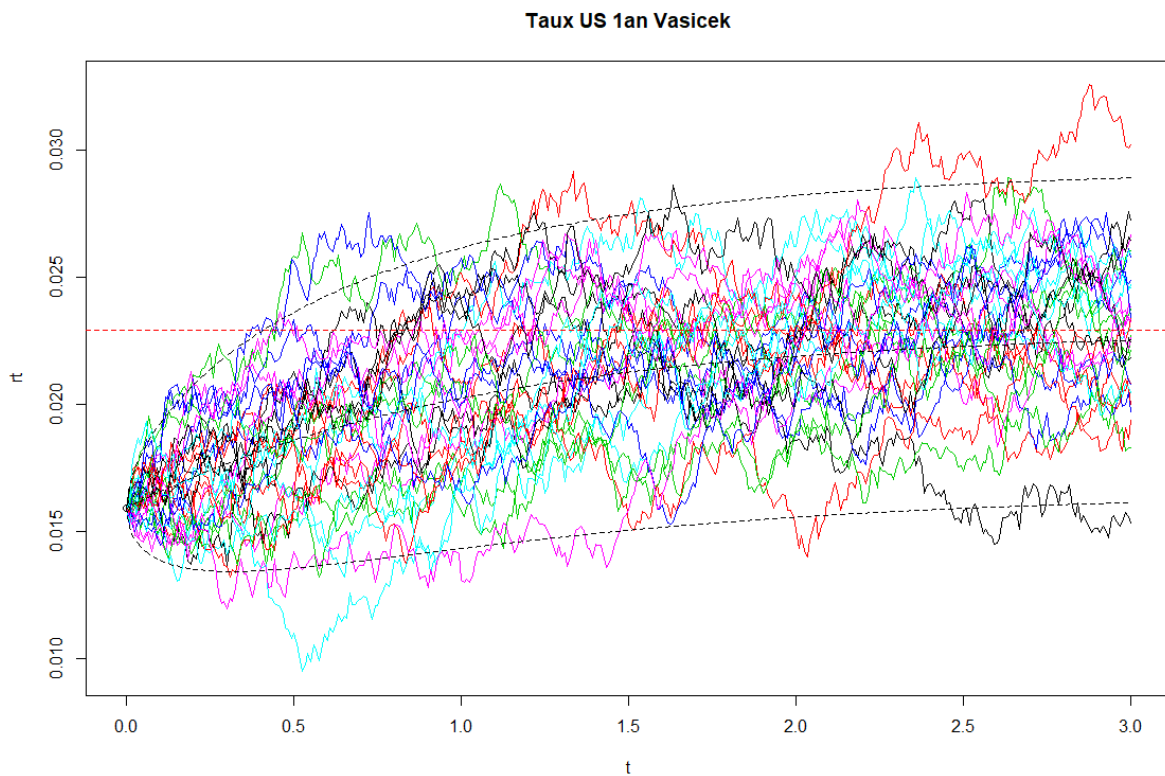


Figure 16 – Extrait de 100 trajectoires générées par la simulation de Vasicek

Dans les sections suivantes, des approches de modélisation des risques catastrophes seront étudiées.

4.3. Données d'observations des variables catastrophes

Les données¹² utilisées, pour cette étude, sont issues de la base de données du NOAA (National Oceanic and Atmospheric Administration), un organisme public américain en charge des observations des météorologiques. Les données utilisées concernent l'Etat de Floride entre 1973-2019. Elles sont issues principalement des sources fournies par les services suivants :

- GSOD – Global Surface Of Day Summary
- NCEP – National Center for Environmental Prediction
- GHCN – Global Historical Climatology Network (GHCN Daily)

4.3.1. Description des variables de la base de données d'observations

Les données utilisées sont les suivantes :

- GUST : La vitesse maximale journalières des rafales de vents (exprimée en nœuds)
- MXSPD : La vitesse maximale journalière du vent (exprimée en nœuds)
- WSPD : La vitesse moyenne journalière du vent (exprimée en nœuds)
- PRCP : La précipitation totale journalière (exprimée en pouce)
- TMAX : La température maximale journalière (exprimé en degré Fahrenheit)
- TMIN : La température minimale journalière (exprimé en degré Fahrenheit)
- SLP : Le niveau de pression journalier moyen au niveau de la mer (exprimé en millibars)
- LTFX : L'indice de soulèvement de l'air ou Lifted Index qui est l'un des indices qui aide à mesurer la stabilité de l'air (exprimé en °C)
- EVAP : L'évaporation journalière observée dans un bac d'évaporation, exprimée en décimètre.
- OMEGA : La vitesse verticale Omega, exprimée en Pascal par seconde.
- RHUM : L'humidité relative de l'air qui est une mesure de la quantité d'eau présente dans l'air, exprimée en %
- U-WIND : La vitesse zonale du vent, exprimée en m/s
- V-WIND : La vitesse méridionale du vent, exprimée en m/s
- DEWP : Le point de rosée ou la température de rosée (exprimé en degré Fahrenheit)

Ces données ont été collectées sur une base journalière dans le but de la construction de notre base de données.

Il est nécessaire d'analyser les données d'observations avant toute modélisation. Nous allons donc procéder à une analyse des données dans la section suivante.

¹² Données téléchargées en date du 21/10/2020

4.3.2. Analyse des données d'observations

L'analyse des données se fera en 3 étapes.

- Statistique descriptive des données d'observations

Le table 5 ci-dessous résume les variables d'observations.

On peut remarquer des données manquantes pour certaines variables. Les valeurs manquantes seront remplacées dans l'étude par une moyenne mobile sur une semaine.

Dans ce tableau, l'autre point qui ressort est que 99% des valeurs observées sur la variable MAX, correspondant à la vitesse maximale journalière des rafales de vents, sont inférieures à 44.1 kts. Ce qui pourrait être interprété que dans les 99% de données d'observations, la vitesse n'atteint pas le seuil de déclenchement du CAT Bond étudié.

	MAX	MXSPD	WDSP	UWND	VWND	SLP	LFTX	OMEGA	RHUM	EVAP	TMIN	TMAX	DEWP	PRCP
nombre	13 431	17 164	17 166	17 166	17 166	17 149	17 166	17 166	17 166	16 491	17 166	17 166	17 166	17 165
moyenne	23.3	13.6	6.8	- 1.9	2.4	1 017.7	5.7	- 0.0	57.2	54.7	83.9	70.2	66.6	33.1
Ecart type	6.4	4.0	2.5	2.5	3.8	4.6	8.8	0.1	20.5	32.2	6.5	8.3	8.6	104.8
Minimum	9.9	2.9	0.9	- 10.6	- 12.0	993.0	- 8.1	- 0.3	0.0	- 56.0	45.0	30.0	18.5	0.0
25%	19.0	11.1	5.0	- 3.5	- 0.0	1 014.8	- 1.1	- 0.1	43.3	38.0	80.1	66.0	62.5	0.0
50%	22.0	13.0	6.5	- 2.3	3.3	1 017.5	2.5	- 0.1	56.3	53.0	84.9	72.0	69.0	0.0
75%	26.0	15.9	8.2	- 0.5	5.3	1 020.5	11.3	0.0	72.3	71.0	89.1	75.9	73.2	5.0
90%	31.1	18.1	9.9	1.5	6.5	1 023.6	18.9	0.1	86.3	84.0	91.0	78.1	74.7	99.0
95%	35.0	21.0	11.2	2.8	7.2	1 025.6	22.8	0.1	92.0	94.0	91.9	80.1	75.4	211.0
99%	44.1	26.0	13.7	5.3	8.4	1 029.4	30.8	0.2	97.8	114.0	93.9	82.0	76.7	508.0
Maximum	99.8	49.9	35.2	12.0	11.8	1 038.4	50.5	0.4	100.0	2 809.0	98.1	84.0	79.7	2 908.0

Table 5 - Variables des données d'observations

- Corrélation entre les différentes variables des données

Cette analyse peut nous permettre de faire des hypothèses sur des relations existantes entre les variations des différentes variables. Le graphique des corrélations permet d'avoir une première information sur les variables qui pourraient avoir des impacts sur l'estimation de notre variable cible. Dans notre cas, le paramètre MAX qui représente la vitesse maximale des rafales observées sur une période est la variable à prédire dans nos modèles.

La figure 18 resume les corrélations entre les différents paramètres des données d'observations pour la zone de la Floride. On peut constater que la vitesse maximale des rafales est fortement corrélée positivement avec les variables vitesses maximales et moyennes journalières du vent et négativement avec la pression au niveau de la mer. On remarque aussi que les précipitations journalières ainsi que les différentes différentes mesures de températures sont corrélées positivement avec notre paramètre à estimer. Un autre point qui sort de ce graphique, est la corrélation significative entre les trois mesures de températures. On pourrait interpréter ce point comme si ces variables nous apportaient les mêmes informations.

Ces différentes informations nous permettent de faire une préselection des variables à retenir dans notre base d'apprentissage du modèle machine learning.

Application à un CAT Bond sur des risques tempêtes

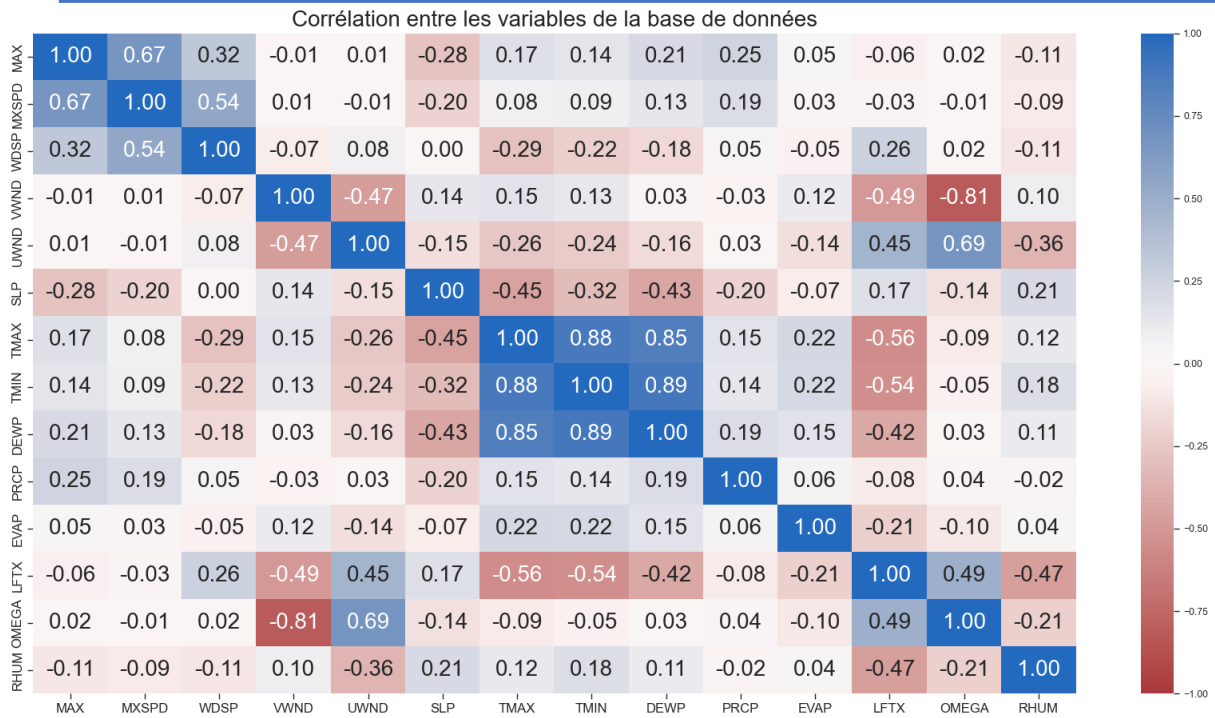


Figure 17 - Corrélation entre les variables des données d'observations

Le graphique suivant nous indique une relation presque linéaire entre les variables TMAX, TMIN et DEWP. Du graphique précédent, on avait noté une très forte corrélation (avec une corrélation minimale de 85%) entre les trois variables. Nous pourrions être amené à ne retenir qu'un seul paramètre des trois dans le cas où nous rencontrons des problèmes de performance en temps de traitements de nos modèles.

On peut noter aussi une relation linéaire entre les variables vitesse maximale journalière et les vitesses maximales des rafales de vent. Ce qui était plus ou moins attendu du fait de la relation entre la vitesse maximale journalière du vent et celle des rafales de vents soutenues.

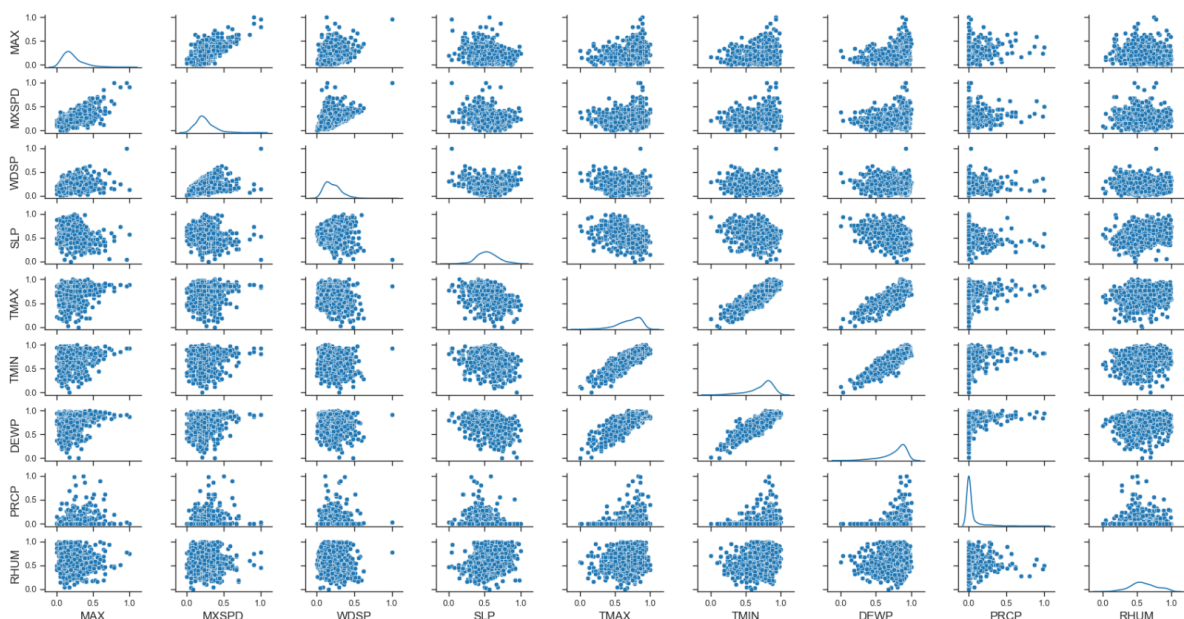


Figure 18 - Relation entre les variables

- **Analyse des séries de données d'observations**

Cette analyse aura pour but de déceler les tendances et les saisonnalités des variables. Les graphiques sur la figure 20 nous permettent de détecter une saisonnalité annuelle sur la plupart des variables. On remarque aucune tendance baissière ou haussière dans les variables. Ces informations nous sont utiles dans le cadre des modélisations des séries temporelles. Elles permettent de choisir les différentes approches de normalisation des données avant d'entraîner nos modèles.

Application à un CAT Bond sur des risques tempêtes

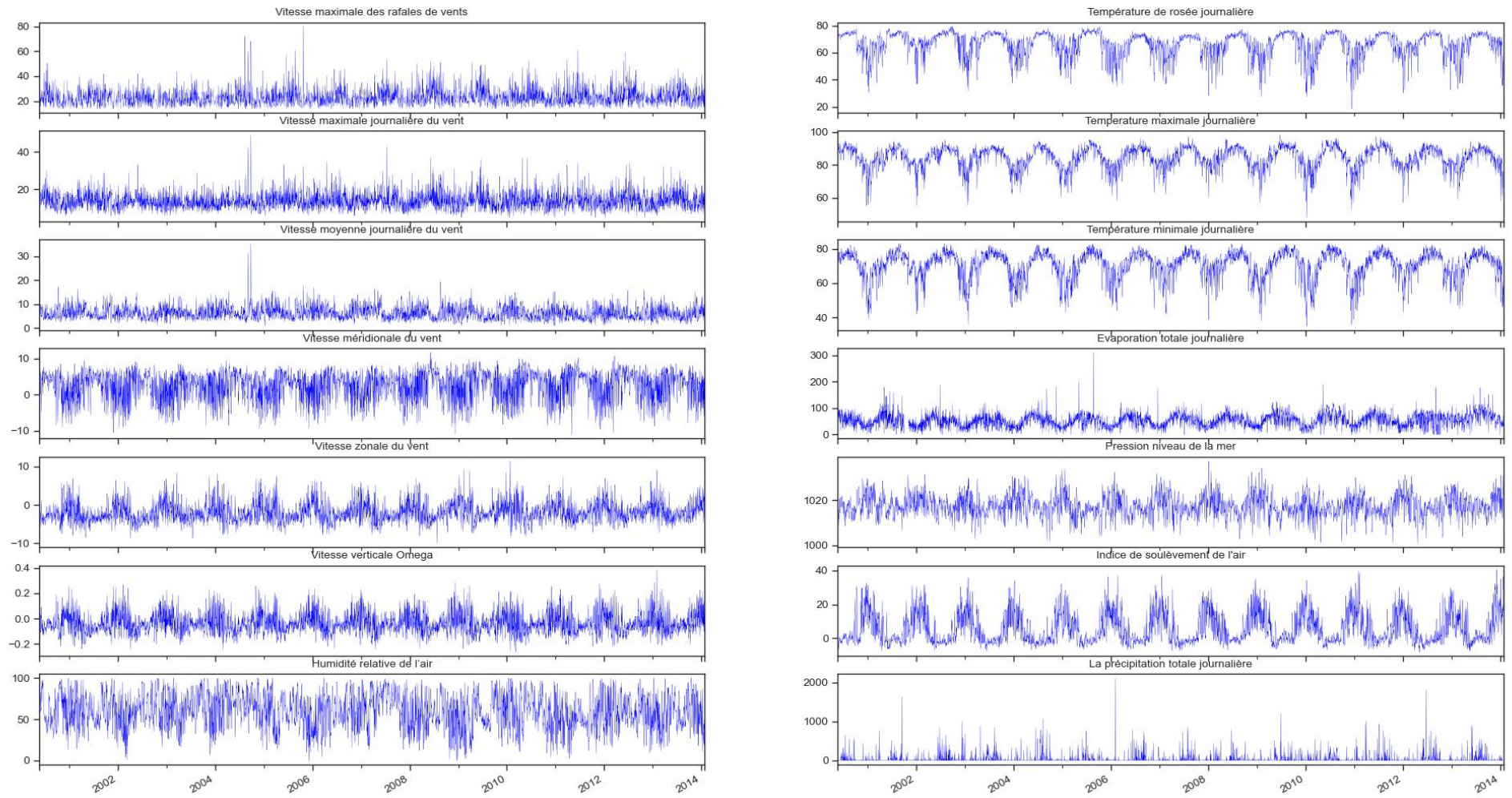


Figure 19 - Evolution des variables entre 2000 - 2014

4.4. Application de l’approche de modélisation par une loi GEV

4.4.1. Analyse des vitesses maximales annuelles

Les données utilisées pour la modélisation du GEV sont les vitesses maximales des rafales de vents entre 1973 et 2019. Les observations de la table 6 et de la figure 21 nous permettent de prendre comme une hypothèse la stationnarité des données. Un prérequis pour l’application de l’approche GEV. Nous allons étudier la stationnarité des données dans la section 4.4.2 pour confirmer cette première impression.

La vitesse maximale sur la période de notre étude a été atteinte en 1992, plus précisément en août 1992 qui correspond à la période de passage de l’ouragan Andrew sur la Floride. L’impact des pertes économiques de l’ouragan Andrew a été un fait marquant dans la création des titres financiers liés aux catastrophes. La seconde valeur maximale correspond à la période de 1979. Durant cette période, la baie de Tampa a été frappé par une série de tornade. On note aussi la période relative à l’épisode de l’ouragan Katrina en 2005.

Ces événements extrêmes apparaissent nettement sur le BoxPlot dans la figure 22 ci-dessous qui met en évidence trois valeurs que nous pouvons considérer comme extrêmes de la distribution des observations.

Années	Vitesses maximales annuelles des rafales de vents (nœuds)	Années	Vitesses maximales annuelles des rafales de vents (nœuds)	Années	Vitesses maximales annuelles des rafales de vents (nœuds)
1973	48	1989	49	2005	80
1974	50	1990	45	2006	46
1975	52	1991	59	2007	53
1976	49	1992	100	2008	52
1977	48	1993	51	2009	51
1978	51	1994	43	2010	52
1979	85	1995	51	2011	61
1980	54	1996	46	2012	59
1981	55	1997	61	2013	46
1982	53	1998	55	2014	51
1983	72	1999	61	2015	61
1984	48	2000	51	2016	59
1985	55	2001	42	2017	69
1986	50	2002	42	2018	60
1987	40	2003	44	2019	58
1988	52	2004	72		

Table 6 - Vitesse maximale annuelle des rafales de vent (source GSOD)

La statistique descriptive des données d’observations ci-dessus donne le table 4. On remarque que 75% des vitesses maximales annuelles sont inférieures à 59.10 kt.

moyenne	Ecart type	Minimum	25%	50%	75%	90%	95%	99%	Maximum
55.1	11.5	40.0	48.5	52.1	59.1	70.2	77.6	92.9	99.8

Table 7 - Analyse descriptive de la vitesse maximale annuelle

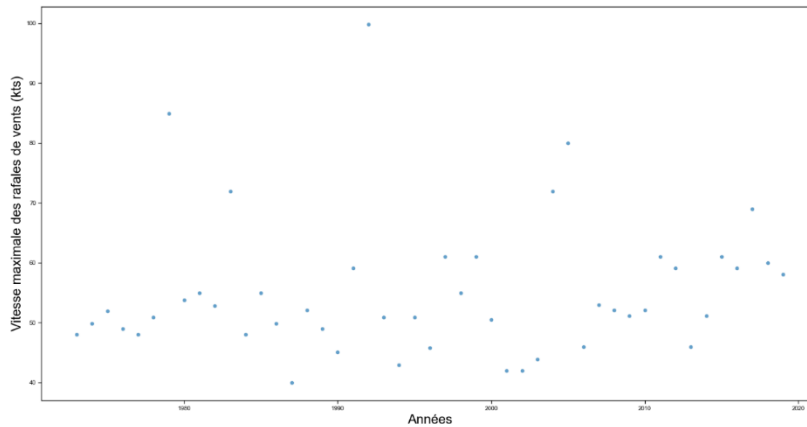


Figure 20 - Vitesses maximales entre 1973 et 2019

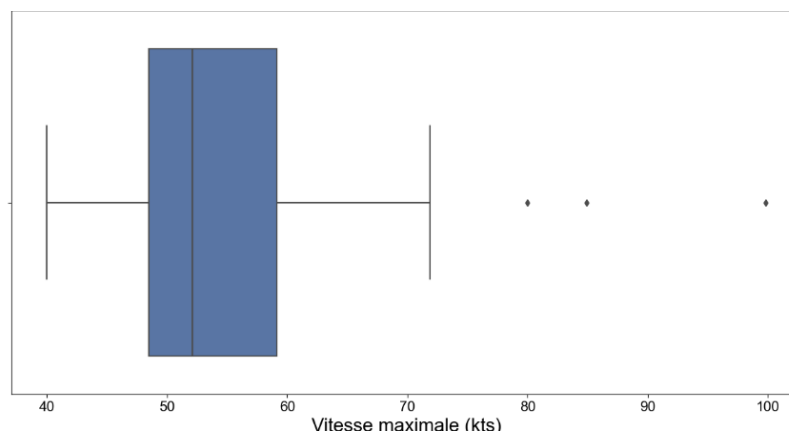


Figure 21 – BoxPlot des vitesses maximales annuelles

4.4.2. Test de stationnarité des observations

En observant la distribution des vitesses sur la figure 20, on ne décèle ni de tendance haussière ni baissière. On pourrait supposer que les données sont stationnaires. Nous allons vérifier cette hypothèse en réalisant un test statistique de stationnarité des données. Nous allons réaliser un test de Dickey et Fuller augmenté (ADF) .

Hypothèses de tests :

H0 : La série comporte une racine unitaire c'est-à-dire qu'elle n'est pas stationnaire

H1 : La série ne comporte pas de racine unitaire c'est-à-dire qu'elle est stationnaire

Statistique ADF	-4.81
p-value	0.000051
Valeurs critiques:	
1%	-3.63
5%	-2.95
10%	-2.61

Table 8 - Test de Dickey Fuller augmenté

Etant donné que la p-value obtenue dans la table ci-dessus est inférieure au niveau de signification 5%, on va donc rejeter l'hypothèse nulle H0 et donc retenir l'hypothèse alternative. De plus, la statistique ADF est inférieur à la valeur critique -3.63 à 1%. Cela suppose que l'hypothèse H0 peut être rejetée avec un niveau de signification inférieur à 1%.

Dans la suite de cette étude, nous allons supposer les données stationnaires.

4.4.3. Calibration des paramètres GEV

Dans cette section, il sera proposé une estimation des différents paramètres du modèle GEV. Nous allons utiliser la bibliothèque « exTremes » sur R avec l'approche EMV.

Les estimations obtenues sur la base de nos données sont :

Paramètres	Estimation
Position μ	49.835
Echelle σ	6.987
Forme ξ	0.154

Ces résultats sont obtenus avec une log-vraisemblance de 169.677.

La matrice de variance covariance donne les résultats suivants :

$$V = \begin{pmatrix} 1.3107 & 0.5228 & -0.0376 \\ 0.5228 & 0.7754 & -0.0074 \\ -0.0376 & -0.0074 & 0.0121 \end{pmatrix}$$

En utilisant les écarts types de la matrice V, on peut avoir une approximation des paramètres avec une intervalle de confiance à 95% :

$$\mu \in [47.59 ; 52.08]$$

$$\sigma \in [5.26 ; 8.71]$$

$$\xi \in [-0.06 ; 0.37]$$

On constate sur les graphiques de la figure 22 que le modèle issu des paramètres ci-dessus peuvent être utilisés pour notre modélisation. En effet, les graphiques des probabilités, quantiles et des niveaux de retour nous montrent presque un alignement des données modélisées et des données d'observations.

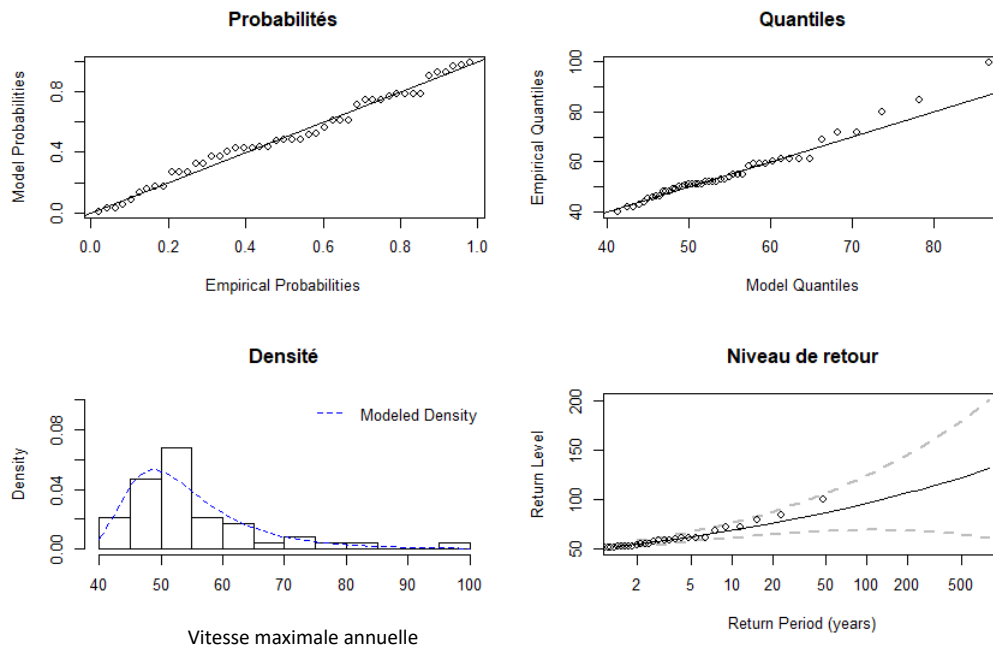


Figure 22 – Alignement des données d'observations sur le modèle calibré

De plus, en substituant les paramètres estimés dans l'expression des quantiles de la distribution GEV, il est possible d'estimer le niveau de retour de période T c'est-à-dire le niveau pour lequel le temps d'attente moyen de dépassement est de T années :

$$\hat{z}_T = \begin{cases} \hat{\mu} + \frac{\hat{\sigma}}{\xi} \left[\left(-\ln\left(1 - \frac{1}{T}\right) \right)^{-\xi} - 1 \right], & \text{si } \xi \neq 0 \\ \hat{\mu} - \hat{\sigma} \left[\ln\left(-\ln\left(1 - \frac{1}{T}\right)\right) \right], & \text{si } \xi = 0 \end{cases}$$

Nous obtenons les résultats suivants :

- 2 ans : 52.47 kts avec une IC 95% [49.86 ; 55.08]
- 20 ans : 76.14 kts avec une IC 95% [65.0 ; 87.2]
- 100 ans : 96.58 kts avec une IC 95% [69.11 ; 124.05]

On a donc une période de retour de 100 ans pour une valeur de 96.58 kts. Ce qui signifie qu'il faudrait en moyenne 100 ans pour atteindre une vitesse maximale annuelle qui est supérieure ou égale à 96.6 kts. De même, il faudrait attendre en moyenne 2 pour dépasser la valeur 52.5 kts. Ces différentes valeurs pourraient être utilisées pour l'analyse de sensibilité du prix CAT Bond aux variations des seuils de déclenchements.

4.4.4. Simulation de trajectoires de 3 années

Dans cette section, nous allons simuler 100 000 trajectoires de 3 années d’une loi GEV ayant les paramètres suivants : $(\mu, \sigma, \xi) = (49.835, 6.987, 0.154)$

Nous avons obtenu les distributions suivantes pour les 100 000 trajectoires de 3 années.

	Vitesse maximale annuelle des rafales de vents (kts)			Réel
	Simulation an 1	Simulation an 2	Simulation an 3	
nombre	100 000	100 000	100 000	47
moyenne	55.1	55.1	55.1	55.1
Ecart type	11.5	11.5	11.6	11.5
Minimum	35.0	34.6	35.0	40.0
25%	47.6	47.6	47.6	48.5
50%	52.4	52.4	52.5	52.1
75%	59.3	59.4	59.4	59.1
90%	68.5	68.7	68.6	70.2
95%	76.1	76.2	76.2	77.6
99%	97.4	96.4	96.2	92.9
Maximum	276.0	324.9	384.0	99.8

Table 9 - Distribution de la simulation GEV sur 3 années

En analysant les quantiles de la distribution des simulations, on constate que la distribution des données simulées est relativement similaire à la distribution des données réelles.

Cette distribution va être utilisée dans la section suivante pour l’évaluation du prix d’un CAT Bond.

4.4.5. Evaluation du prix du CAT Bond

Dans la méthode Monte Carlo, le prix du CAT Bond est la moyenne des différentes simulations de prix obtenues pour les 100 000 scénarios. Pour la suite, nous allons estimer le prix des CAT Bond de chaque scénario en fonctions des différentes valeurs des coefficients des tranches. Le prix obtenu avec les paramètres définis dans la section 4.1 :

Principal = 100 millions \$ et les seuils de déclenchements :

$$\begin{aligned}
 \alpha_1 &= 0\% & X \leq 63 \text{ kts} \\
 \alpha_2 &= 20\% & 63 < X \leq 82 \text{ kts} \\
 \alpha_3 &= 40\% & 82 < X \leq 95 \text{ kts} \\
 \alpha_4 &= 60\% & 95 < X \leq 112 \text{ kts} \\
 \alpha_5 &= 80\% & 112 < X \leq 136 \text{ kts} \\
 \alpha_6 &= 100\% & X > 136 \text{ kts}
 \end{aligned}$$

On obtient un prix de CAT Bond de 98.67 millions.

La table ci-dessous résume les répartitions des potentiels flux financiers sur les 100 000 simulations. On constate que l’investisseur perdrait au moins 50% de ses investissements dans 5% des cas et dans plus de 55% des cas, l’investisseur ne subirait aucune perte.

Application à un CAT Bond sur des risques tempêtes

	Montants non actualisés			
	an 1	an 2	an 3	Principal
nombre	100 000	100 000	100 000	100 000
moyenne	6.2	5.9	5.6	86.9
Ecart type	0.7	1.0	1.2	18.9
Minimum	0.0	0.0	0.0	0.0
1%	2.6	1.3	1.3	20.0
5%	5.2	3.9	2.6	40.0
10%	5.2	5.2	3.9	60.0
25%	6.5	6.5	5.2	80.0
45%	6.5	6.5	6.5	100.0
Maximum	6.5	6.5	6.5	100.0

Table 10 - Statistique descriptive des flux financiers simulés (\$ millions)

Plusieurs simulations ont été effectuées afin d'observer la distribution des prix du CAT Bond générés par le modèle. La figure ci-dessous reprend les prix obtenus à partir de 1000 simulations de prix. Les prix simulés varient entre 97.9 \$millions et 99.1 \$millions. La moyenne des prix est de 98.5 \$millions. La ligne verticale rouge représente le prix retenu dans notre première estimation.

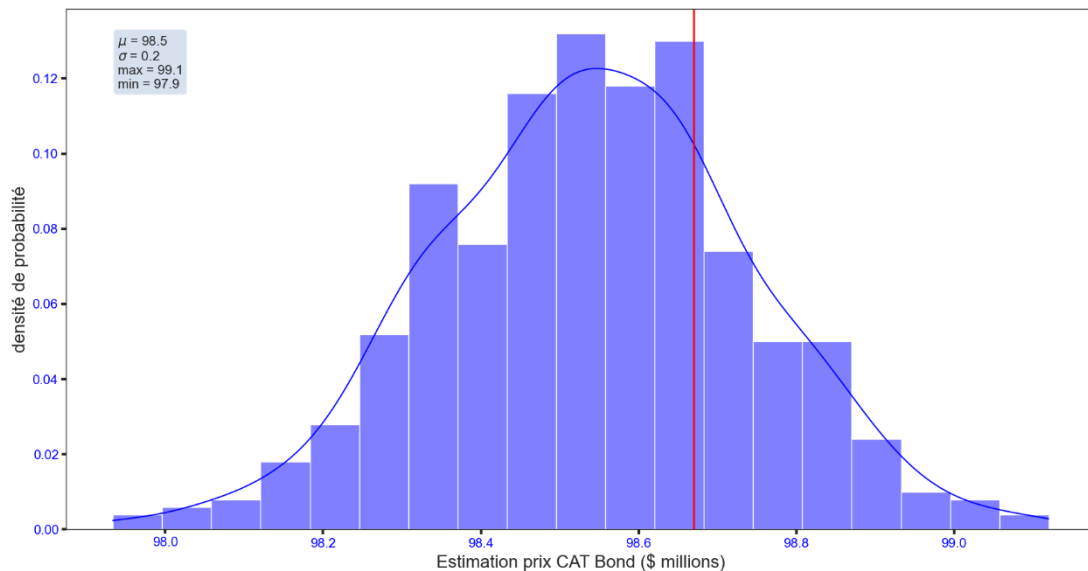


Figure 23 - Distribution des 1000 simulations de prix CAT Bond

4.4.6. Sensibilité du prix aux variations de seuil de déclenchement

Le prix d'un CAT Bond dépend de son seuil de déclenchement. Il serait intéressant d'analyser la sensibilité de la valorisation du prix d'un CAT Bond aux différents de déclenchement.

Dans cette étude de sensibilité, nous allons considérer les scenarios suivants :

- Scenario 1 : Seuil de déclenchement comme défini dans l'étude
- Scenario 2 : Seuil de déclenchement à partie de la catégorie 2 (similaire à une vitesse d'un ouragan de catégorie 2 selon l'échelle Saffir-Simpson)

$\alpha_1 =$	0%	$X \leq 63$ kts
$\alpha_2 =$	0%	$63 < X \leq 82$ kts
$\alpha_3 =$	40%	$82 < X \leq 95$ kts
$\alpha_4 =$	60%	$95 < X \leq 112$ kts
$\alpha_5 =$	80%	$112 < X \leq 136$ kts
$\alpha_6 =$	100%	$X > 136$ kts

- Scenario 3 : Seuil de déclenchement à partie de la catégorie 3 (similaire à une vitesse d'un ouragan de catégorie 3 selon l'échelle Saffir-Simpson)

$\alpha_1 =$	0%	$X \leq 63$ kts
$\alpha_2 =$	0%	$63 < X \leq 82$ kts
$\alpha_3 =$	0%	$82 < X \leq 95$ kts
$\alpha_4 =$	60%	$95 < X \leq 112$ kts
$\alpha_5 =$	80%	$112 < X \leq 136$ kts
$\alpha_6 =$	100%	$X > 136$ kts

- Scenario 4 : Seuil de déclenchement à partie de la catégorie 4 (similaire à une vitesse d'un ouragan de catégorie 4 selon l'échelle Saffir-Simpson)

$\alpha_1 =$	0%	$X \leq 63$ kts
$\alpha_2 =$	0%	$63 < X \leq 82$ kts
$\alpha_3 =$	0%	$82 < X \leq 95$ kts
$\alpha_4 =$	0%	$95 < X \leq 112$ kts
$\alpha_5 =$	80%	$112 < X \leq 136$ kts
$\alpha_6 =$	100%	$X > 136$ kts

- Scenario 5 : Seuil de déclenchement à partie de la catégorie 3 (similaire à une vitesse d'un ouragan de catégorie 3 selon l'échelle Saffir-Simpson)

$\alpha_1 =$	0%	$X \leq 63$ kts
$\alpha_2 =$	0%	$63 < X \leq 82$ kts
$\alpha_3 =$	0%	$82 < X \leq 95$ kts
$\alpha_4 =$	0%	$95 < X \leq 112$ kts
$\alpha_5 =$	0%	$112 < X \leq 136$ kts
$\alpha_6 =$	100%	$X > 136$ kts

Application à un CAT Bond sur des risques tempêtes

Le tableau ci-dessous nous donne les résultats pour les différents scénarios.

- Le prix moyen varie entre 98.6 (pour le scénario 1) et le 112.2 (scénario 5) principalement dû au fait d'un seuil plus élevé ce qui implique moins de paiements de sinistres.
- En observant le second quartile, on peut constater que pour plus de 50% des cas, le prix est presque similaire quel que soit le seuil de déclenchement.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
nombre	100 000	100 000	100 000	100 000	100 000
moyenne	98.6	107.7	110.1	111.5	112.2
Ecart type	20.1	16.5	13.3	9.5	5.6
Minimum	0.0	0.0	0.0	0.0	0.0
1%	25.0	30.0	32.7	32.7	110.8
5%	52.6	68.0	111.0	111.3	111.3
10%	70.2	111.0	111.5	111.6	111.6
25%	90.8	111.9	112.0	112.0	112.0
50%	111.7	112.4	112.5	112.5	112.5
90%	113.2	113.4	113.4	113.4	113.4
99%	114.0	114.1	114.2	114.2	114.2
max	115.6	115.6	115.6	115.6	115.6

Table 11 - Sensibilité du prix du CAT Bond par rapport au seuil de déclenchement (\$ millions)

Le seuil de déclenchement est un paramètre dont l'analyse est essentielle pour la mise en place d'une obligation catastrophe. On peut voir que pour un seuil de déclenchement de catégorie 5, l'investisseur a moins d'1% de risque de perdre une partie son investissement. A l'inverse, pour un seuil de déclenchement de catégorie 1, il se retrouverait avec un risque de plus de 25% de perdre une partie de son investissement.

Après avoir exposé une application de modélisation du prix d'un CAT Bond paramétrique avec une approche de modélisation des risques liés aux événements catastrophes par une loi de probabilité, la prochaine section sera consacrée à l'étude d'une application avec les réseaux de neurones récurrents.

4.5. Application de l'approche de modélisation Deep Learning

L'approche Deep Learning consistera dans ce mémoire à mettre en place un modèle séquentiel de réseau de neurones, qui à partir d'une séquence de données d'observations va nous permettre d'estimer des valeurs maximales annuelles des vitesses de vent.

Dans le but de mettre en place ces modèles, nous avons proposé, dans la section analyse de données, une analyse de détection de corrélation entre variables et des tendances dans les évolutions temporelles des données. Cette étape d'analyse des données est une étape importante dans la mise en place d'un modèle Machine Learning. Elle est aussi utile à identifier les probables tendances, saisonnalités et stationnarités des variables.

Ces informations sont utiles dans la sélection du modèle de réduction ou normalisation des données.

Dans cette section, nous procéderons, dans un premier temps, à la description de notre modèle. Ensuite, une analyse des résultats de l'apprentissage des modèles sera effectuée. Enfin, les deux dernières étapes concerneront l'estimation de la trajectoire de la variable à prédire ainsi que l'estimation du prix du CAT Bond.

4.5.1. Construction de la base de données des modèles

La base de données, utilisée pour l'apprentissage, a été construite avec environ une cinquantaine d'années d'historiques journalières des variables d'observations pour les différents états inclus dans notre étude. Dans cette base, il y a 50 observations de maximum annuels qui sont insuffisantes pour réaliser une étude Deep Learning. Nous allons donc procéder à l'augmentation de notre base en passant d'année calendaire en année constituée de 365 jours d'observations. Cette base de 50 années d'observations va nous permettre de construire une base de données avec environ 18 000 observations de vitesses maximales annuelles. Le processus est décrit dans la figure 24 ci-après.

Cette nouvelle base construite sera le point de départ de notre étude. Cette base sera divisée en deux groupes, le premier groupe sera utilisé pour l'apprentissage du modèle et le second pour la validation des modèles obtenus.

Dans la suite de ce document, les données d'apprentissage sera noté « Base d'apprentissage » et les données de validation « Base de Test ».

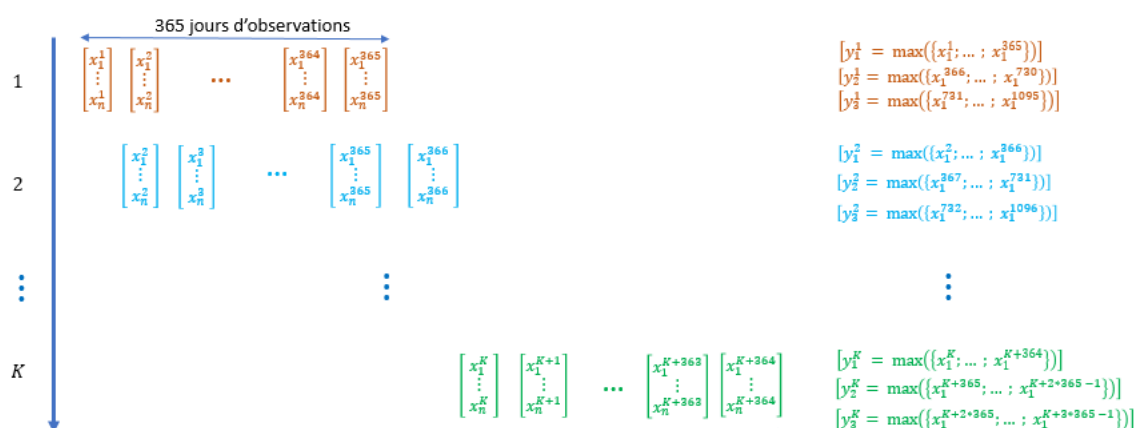


Figure 24 - Passage en base 365 jours

4.5.2. Description des modèles

Dans cette section, nous allons décrire les paramètres des modèles utilisés. Dans une optique de synthétisation, nous n'allons pas proposer tous les modèles étudiés. Nous allons limiter cette section aux modèles optimaux obtenus selon les capacités des outils à notre possession.

Le modèle LSTM utilisé, dans le cadre de cette étude, est un modèle séquence à séquence de type « Encoder Décodeur ». Dans la partie d'encodage et décodage, nous allons utiliser deux couches LSTM comportant chaque plusieurs neurones. Le nombre de neurones dans chaque couche fera partie des hyperparamètres utilisés pour l'optimisation du modèle.

Dans la mise en place d'un modèle LSTM, un paramètre indispensable est la taille de la fenêtre d'historiques c'est-à-dire le nombre d'observations historiques nécessaires pour la prédiction d'une séquence de valeurs. La figure ci-dessous décrit cette méthode.

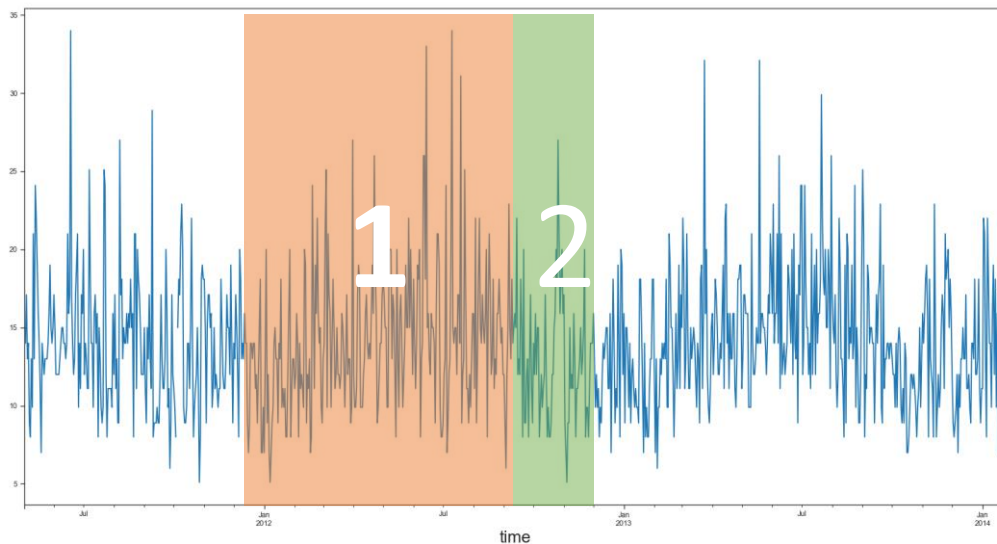


Figure 25-LSTM Fenêtre d'historiques (1) et de prédictions (2)

Dans notre étude, nous avons principalement retenu 4 tailles de fenêtres : 1 année, 2 années, 3 années et 5 années.

Les performances des modèles seront analysées afin de sélectionner le modèle optimal. Dans cette optique d'évaluation de la performance, nous allons d'abord, définir les métriques de performance qui seront utilisées dans la section suivante et ensuite nous évaluerons les modèles.

4.5.3. Evaluation des modèles

Dans un modèle de Deep Learning, le choix du modèle est fait à partir de métriques de d'évaluation de la performance des modèles. Dans cette partie dédiée à l'évaluation des modèles, nous allons d'abord définir les métriques utilisées et après faire une analyse des performances des différents modèles.

- **Métriques d'évaluation de modèles :**

Les métriques suivantes seront utilisées pour l'évaluation des modèles.

Dans le cas des modèles de régression, on utilisera :

- ✓ RMSE ou la racine carrée des erreurs quadratiques moyenne : Cette métrique est utilisée pour sa capacité à pénaliser les estimations qui sont éloignées de la valeur cible.

$$RMSE (Y, \hat{Y}) = \sqrt{\frac{1}{m} * \sum_{i=1}^m (y_i - \hat{y}_i)^2}$$

- ✓ MAE ou la moyenne des erreurs absolues : Cette métrique informe de l'amplitude moyenne des écarts entre les estimations et les valeurs cible.

$$MAE (Y, \hat{Y}) = \frac{1}{m} * \sum_{i=1}^m |y_i - \hat{y}_i|$$

Plus le RMSE et le MAE sont proche de zéro, meilleur sera le modèle.

Dans le cas des modèles classification, on utilisera :

- ✓ L'entropie croisée multinomiale : Soit K le nombre de classe du modèle, on définit la fonction d'entropie croisée multinomiale comme suit :

$$ECM (Y, \hat{Y}) = -\frac{1}{m} \left(\sum_{i=1}^m \sum_{j=1}^K y_i^{[j]} \log \hat{y}_i^{[j]} + (1 - y_i^{[j]}) \log (1 - \hat{y}_i^{[j]}) \right)$$

L'ECM du modèle optimal est proche de zéro.

- ✓ La précision moyenne d'un modèle de classification multiple : Soit K le nombre de classe du modèle, on définit la fonction « Accuracy » comme suit :

$$ACC (Y, \hat{Y}) = \frac{1}{K} \left(\sum_{i=1}^K \frac{\text{Nombre de prédiction correcte}_i}{\text{nombre totale de valeur à prédire}} \right)$$

Une ACC de 100% implique que le modèle arrive à prédire sans erreur. Et une ACC=0% implique que le modèle n'arrive à prédire aucune valeur correctement.

- **Evaluation des modèles :**

Les modèles ont été mis en place sous Python avec la bibliothèque « Tensorflow ». La structure du modèle est disponible en annexe.

Les tableaux obtenus, ci-dessous, sont le résultat de l'apprentissage des modèles sur 75% des données d'observations et de la validation du modèle sur les 25% restantes.

- **Modèle LSTM de régression :**

Les données du tableau, ci-dessous, résument les résultats obtenus des différents test effectués.

		Apprentissage		Test	
		RMSE	MAE	RMSE	MAE
LSTM - 5y	t+1	7.8	3.3	10.4	6.4
	t+2	8.0	3.2	13.4	10.6
	t+3	9.1	4.8	14.2	10.3
LSTM - 3y	t+1	5.3	3.2	7.1	5.4
	t+2	6.4	4.1	7.3	6.7
	t+3	6.8	5.0	8.4	6.5
LSTM - 2y	t+1	7.7	4.5	8.5	6.8
	t+2	6.8	4.0	9.1	7.2
	t+3	8.1	6.1	11.4	9.9
LSTM - 1y	t+1	8.1	4.7	12.8	10.9
	t+2	8.3	4.6	11.5	9.5
	t+3	7.3	3.9	13.8	12.1

Table 12-Résultats modèles LSTM de régression (kts)

On constate que pour une fenêtre d'historique d'une année d'observations, on a le phénomène de l'Overfitting soit le surapprentissage. En effet, le MAE, entre l'apprentissage et le test, est passé au double (d'environ 4.5 à 10 kts). Cela est une indication que le modèle n'arrive pas à généraliser les informations apprises sur les données de test. C'est aussi une indication qu'il faudrait augmenter la taille de la fenêtre ou des données.

En utilisant des historiques de 2ans ou 3 ans, on constate une amélioration entre les écarts des métriques d'apprentissage et de test.

Pour le modèle LSTM avec 5 années d'historiques, on constate le même phénomène que pour le modèle LSTM avec une année d'historique. Ce qui pourrait être interprété comme une période trop longue d'historiques d'observations.

Le modèle 3 années d'historiques paraît être celui qui réplique le mieux les données d'observations. Pour la suite de cette section, on utilisera les valeurs du modèle 3 années. On peut aussi le constater sur la figure 26 ci-dessous.

Application à un CAT Bond sur des risques tempêtes

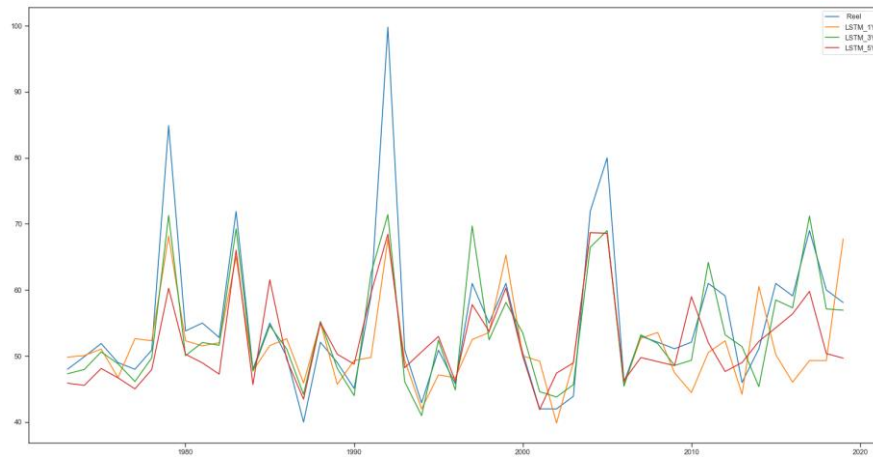


Figure 26 - Résultats des modèles LSTM

- Modèles de Classification :

Pour les besoins de la classification, une nouvelle variable va être ajoutée dans la base de données. Cette variable représentera les différentes classes de l'échelle de Saffir-Simpson.

La valeur de cette classe sera déterminée à partir des valeurs maximales annuelles comme ci-dessous.

$$y_t^{[i]} = \begin{cases} 1 & y_t^i \leq 21 \text{ kts} \\ 2 & 21 < y_t^i \leq 33 \text{ kts} \\ 3 & 33 < y_t^i \leq 63 \text{ kts} \\ 4 & 63 < y_t^i \leq 82 \text{ kts} \\ 5 & 82 < y_t^i \leq 95 \text{ kts} \\ 6 & 95 < y_t^i \leq 112 \text{ kts} \\ 7 & 112 < y_t^i \leq 136 \text{ kts} \\ 8 & y_t^i > 136 \text{ kts} \end{cases}$$

Les données de table 13 résume les performances des modèles utilisés pour l'approche classification du modèle LSTM.

		Apprentissage		Test	
		ECM	ACC	ECM	ACC
LSTM - 5y	t+1	0.239	90.0%	0.164	94.5%
	t+2	0.249	90.8%	0.048	91.1%
	t+3	0.310	90.2%	0.097	88.4%
LSTM - 3y	t+1	0.351	89.8%	0.604	85.9%
	t+2	0.127	92.3%	0.184	90.7%
	t+3	0.230	94.4%	0.451	91.4%
LSTM - 2y	t+1	0.296	87.8%	1.422	69.0%
	t+2	0.332	86.8%	1.020	63.4%
	t+3	0.106	96.1%	0.534	90.2%
LSTM - 1y	t+1	0.365	85.6%	0.848	71.4%
	t+2	0.366	85.8%	0.835	72.8%
	t+3	0.069	97.9%	0.730	83.7%

Table 13 - Résultats des modèles de classification

Application à un CAT Bond sur des risques tempêtes

Les lignes des modèles LSTM 1y et LSTM 2y montrent que ces modèles n'arrivent pas à généraliser les informations apprises lors de la phase d'apprentissage du modèle. On parle d'un surapprentissage.

Le phénomène opposé est constaté sur le modèle LSTM 5y où les performances des données de Test sont meilleures que celles des données d'apprentissage. On parle dans ce cas de sous-apprentissage. Dans les deux cas, cette situation suggère une modification des hyperparamètres du modèle.

Comme dans le modèle de régression, on constate, pour le modèle LSTM 3y, des performances presque similaires entre les phases d'apprentissage et de test du modèle. De plus, l'ECM montre que les erreurs de classes dans le modèle LSTM 3y sont moyenne très inférieures à une classe.

4.5.4. Estimation du prix du CAT Bond

A partir des modèles obtenus dans la section précédente, cette étape consistera à prédire les évolutions de la vitesse annuelle maximale des rafales de vent. Il sera proposé deux estimations, l'une à partir du modèle de classification et l'autre à partir du modèle de régression.

Pour les estimations des valeurs futures, on utilisera les données de l'année 2019 comme dernière année de données disponibles afin de permettre au modèle d'estimer la trajectoire de 3 années après cette période.

- Cas modèle de régression LSTM

Les résultats estimés par le modèle LSTM 3y sont résumés dans la table ci-dessous.

On constate que le modèle ne prédit pas de déclenchement du CAT Bond sur sa période d'activité. Etant donné que le modèle prédit à une erreur, il serait intéressant de voir les bornes des prix en tenant compte des erreurs constatées sur la base de Test.

Le prix du CAT Bond estimé varie entre 70 \$millions (cas où il y a deux déclenchements avec l'ajout du RMSE sur la sortie du modèle) et 112.51 \$ millions (cas où il n'y a aucun déclenchement du CAT Bond).

		Résultats modèles	Cas an 1 + RMSE Test	Cas an 2 + RMSE Test	Cas an 3 + RMSE Test	Cas + RMSE sur toutes les années
LSTM - 3y (kts)	t+1	57.11	64.21	57.11	57.11	64.21
	t+2	51.38	51.38	58.68	51.38	58.68
	t+3	55.94	55.94	55.94	64.34	64.34
Flux coupons CAT Bond non actualisés (\$ millions)	t+1	6.50	5.20	6.50	6.50	5.20
	t+2	6.50	5.20	6.50	6.50	5.20
	t+3	6.50	5.20	6.50	5.20	3.90
	Principal	100.00	80.00	100.00	80.00	60.00
Prix CAT Bond (\$ millions)		112.51	90.01	112.51	92.53	70.03

Table 14 - Prix CAT Bond - LSTM régression

Application à un CAT Bond sur des risques tempêtes

- Cas modèle de classification LSTM

Dans le modèle de classification, on obtient les données suivantes :

		Probabilités					
		Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6
LSTM - 3y	t+1	0.0009%	0.0032%	89.2923%	10.1607%	0.4848%	0.0581%
	t+2	0.0032%	0.0049%	87.6166%	12.0323%	0.2513%	0.0863%
	t+3	0.0026%	0.0085%	88.0841%	11.5603%	0.3220%	0.0225%

Le modèle prévoit plus de 88% qu'il n'y aura pas déclenchement du CAT Bond.

Le prix du CAT Bond est obtenu à partir de la formule spécifiée dans les sections 3.5.4.

En calculant le terme $\mathbb{P}\left(X_k^{[i]}\right) F_{CAT}\left(X_k^{[i]}\right)$ (procédure en Annexe), on trouve les flux non actualisés suivants :

	Flux an 1	Flux an 2	Flux an 3	Principal
Flux probables CAT Bond non actualisés	6.363	6.203	6.110	93.999

Le prix du CAT Bond sera la moyenne des flux actualisés avec les 100 000 trajectoires de taux simulés. On trouve $V_{CAT}(0) = 106.09$ \$ millions

4.6. Analyse des résultats

Les différents prix estimés par les modèles dans ce mémoire nous donnent des fourchettes de prix pour le CAT Bond étudié. La volatilité constatée du prix en fonction du modèle choisi et des paramètres nous amène aux conclusions suivantes :

- Les modèles Deep Learning dépendent essentiellement de la qualité des données incluses dans la base. On a constaté dans le cas du modèle de régression que le prix variait de 112 \$millions à 70 \$millions en tenant compte des erreurs d'estimation du modèle. Il est donc nécessaire de s'intéresser à la variation des erreurs dans la mise en place d'un modèle Deep Learning. Dans cette étude, il est estimé, dans certains cas, à plus de 10% de la valeur moyenne de la vitesse maximale. On pourrait donc conclure que le modèle retenu pourrait être amélioré afin de réduire cette erreur.
- Dans le modèle de classification LSTM qui est un modèle basé sur les probabilités d'occurrence de chaque classe de risque, la faible probabilité des classes déclenchant le CAT Bond entraîne un prix légèrement élevé par rapport au modèle GEV.
- Dans l'approche GEV, certaines valeurs extrêmes ont été générées dans les simulations. Ces valeurs ont certainement contribué à la réduction du principal et des coupons d'où la réduction prix. Le prix constaté varie entre 97.9 et 99.1 \$millions.
- Comme on peut le constater ci-dessous, le modèle LSTM a un prix plus élevé du fait du non déclenchement du CAT Bond sur la période du contrat. Par contre, en effectuant une analyse de sensibilité sur le prix du CAT Bond avec les erreurs du modèle, on constate un déclenchement des CAT Bond ce qui entraîne une baisse du prix vers des bases comparables avec les autres modèles. De plus 10% des prix simulés du modèle GEV sont de l'ordre du prix du scénario extrême du modèle LSTM – régression. Environ 50% des données simulées donnent un prix sensiblement égal aux modèles LSTM.

	Prix CAT Bond (\$ millions)
LSTM - Regression	112.51
Cas an 1 + RMSE Test	90.01
Cas an 2 + RMSE Test	112.51
Cas an 3 + RMSE Test	92.53
Cas ajout RMSE sur toutes les périodes	70.03
LSTM - Classification	106.90
Modèle GEV	98.67
MC GEV - Quantile 90%	113.16
MC GEV - Quantile 50%	111.66
MC GEV - Quantile 25%	90.78
MC GEV - Quantile 10%	70.16

Ces différentes approches peuvent permettre de définir une fourchette de prix pour une obligation catastrophe. Cette étude comporte certainement des limites que nous essaierons d'aborder dans la partie suivante.

Comme toute étude qui se basent sur des hypothèses de simplification théoriques, les méthodes de Pricing utilisées dans cette étude ont aussi des limites sur lesquelles l'attention du lecteur doit être attirée. Ces hypothèses viennent principalement de contraintes techniques et théoriques de simplification du modèle.

Ci-dessous seront listées quelques limites et contraintes de cette étude. Elles ne sont pas exhaustives, elles attirent l'attention du lecteur sur des réserves à avoir en lisant cette étude.

(i). Hypothèses d'indépendance entre les risques catastrophes et les risques financiers

Dans un objectif de simplification, il a été pris, comme hypothèse, l'indépendance entre les risques catastrophes et les risques financiers. Dans la pratique, on peut constater, à l'exemple de la pandémie du COVID, que les risques financiers peuvent être influencés par les événements catastrophes. Cette hypothèse tend à négliger la corrélation entre les deux types de risques.

(ii). Risque de défaut de l'assureur supposé négligeable

Cette hypothèse comporte un risque pour l'investisseur dans le sens où il suppose que seul les événements catastrophes auront un impact sur les flux financiers alors qu'il est exposé au risque de non-paiements des primes par l'assureur en cas de défaut de celui-ci.

(iii). Les placements effectués par le SPV sont sans risque

Durant la crise des subprimes en 2008, la faillite de Lehman Brothers a entraîné des défauts sur 4 CAT Bond. La structure des investissements du SPV peut donc dans certains cas comporter des risques. Cette hypothèse rend négligeable les risques des actifs dans lesquels le SPV a investi.

(iv). Limites de l'approche Deep Learning :

- Disponibilité des données :

L'approche Deep Learning est essentiellement basée sur la disponibilité des données. Elle dépend donc de la qualité des informations collectées pour entraîner le modèle.

Dans notre modèle, nous avons dû remplacer des données manquantes par des valeurs moyennes. Il faudrait faire attention dans l'utilisation de cette approche dans un modèle Deep Learning.

- Recherche du modèle optimal :

La méthode Deep Learning comporte aussi une contrainte non négligeable qui est la capacité des outils à faire tourner le modèle. Du fait du temps de traitement nécessaire excessivement long dans certains cas, certains paramètres sont fixés ou choisis lors de la mise en place du modèle. Ce choix prive le modèle de la recherche du modèle optimal. On se retrouve avec un modèle optimal local.

Conclusion

Les obligations catastrophes, par leur complexité, sont un réel challenge pour les investisseurs. Si l'on rajoute à cela, la complexité de mettre en place des stratégies de couverture avec des produits financiers traditionnels, l'hésitation des investisseurs sur ce marché devient compréhensible.

Le travail réalisé dans cette étude avait pour objectif de proposer deux approches d'estimation du prix d'un CAT Bond paramétrique : l'une basée sur une modélisation des variables des risques catastrophes avec une loi de distribution et la seconde approche consistait à mettre en place un modèle séquentiel Deep Learning.

La première étape de cette étude, a consisté à définir une approche de valorisation des obligations catastrophes. Le modèle de valorisation utilisé, au vu de la structure du marché incomplet, est une modélisation à temps discret mise en place à partir de l'approche théorie d'équilibre des prix. Dans cette approche, le prix du CAT Bond dépend de deux groupes de variables : les variables des risques financiers et des risques catastrophes.

La seconde étape consistait donc à modéliser les risques financiers et les risques catastrophes. Cette étape a été subdivisée en 2 deux grandes parties selon le groupe de variable à modéliser.

Tout d'abord, la variable des risques financiers a été modélisée avec un modèle stochastique de taux, qui dans notre étude est le modèle de Vasicek. Ce modèle de taux a été utilisé à des fins de simplification du modèle. En effet, connaissant les limites du modèle Vasicek dans la réplique des courbes de taux telles qu'observées, il serait tout à fait normal que certains auraient privilégié des approches différentes de modélisation des taux d'intérêts.

Ensuite, il a été proposé une modélisation des risques catastrophes avec le modèle GEV étant donné que le CAT Bond de l'étude a pour paramètre de déclenchement le maximum annuel d'une variable liée aux événements catastrophes. Dans cette approche, le prix du CAT Bond est estimé avec une méthode de Monte Carlo après simulation de plusieurs trajectoires de taux d'intérêts et de loi GEV.

Enfin, une approche Deep Learning a été mise en place en utilisant le modèle séquentiel des réseaux de neurones LSTM. Dans cette approche, la principale difficulté est la taille de l'échantillon. Cette approche a été réalisée en plusieurs étapes. Dans la première étape, il a été proposé une approche d'augmentation de la base de données. Dans la seconde étape, deux modèles séquentiels LSTM ont été appris des données d'apprentissage. Dans la dernière étape, le prix a été estimé en utilisant les définitions de l'espérance.

Même si ces deux approches ont permis de proposer des estimations de prix d'un CAT Bond paramétrique sur un risque de tempête, nous restons toutefois prudents avec les limites que comportent cette étude.

De plus, cette étude pourrait être approfondie avec l'étude des cas où ils existent une corrélation entre les risques financiers et risques catastrophes. Ces groupes de risques ont été supposés indépendants et le modèle de valorisation est obtenue principalement à partir de cette hypothèse. Il arrive souvent qu'il y ait une corrélation entre ces deux risques. La crise du COVID est un parfait exemple.

Conclusion

De nos jours, les avancées des différentes méthodes de Deep Learning pourraient contribuer à la simplification de certaines problématiques assurantielles et financières. On pourrait également citer comme piste exploitable pour la suite de cette étude, la modélisation des CAT Bond sur la base d'une historique d'images météorologiques. Cette approche pourrait faire intervenir une approche nouvelle du LSTM : ConvLSTM, qui combine les techniques de Deep Learning dédiées à la reconnaissance d'images et l'architecture du LSTM.

On pourrait aussi appliquer cette approche à un autre domaine de l'assurance qui est l'assurance paramétrique.

Annexes

Algorithme de simulation du prix du CAT Bond : Approche modélisation par loi GEV

```

# Caracteristiques du CAT Bond et facteurs des differentes tranches
a1 <- 0
a2 <- 0.2
a3 <- 0.4
a4 <- 0.6
a5 <- 0.8
a6 <- 1

principal = 100
P <- matrix(0, 100)

c <- 0.03
spread <- 0.035

## Simulation des risques financiers ##
## Paramètres du modèle
r0 <- 0.0159
theta <- 0.022901
k <- 0.9649
beta <- 0.004446

## simulation des taux
n <- 10000 # Nombre de simulations
T <- 3 # Maturité du Cat Bonds
m <- 360 # Convention annuelle
dt <- T/m # le temps delta entre deux simulations
run <- 1000 # Nombre de lancement du calcul de prix

for (z in 1:run){
  r <- matrix(0,m+1,n) # Initialisation de la matrice des taux
  r[1,] <- r0

  # Simulation des trajectoires de taux r
  for(j in 1:n){
    for(i in 2:(m+1)){
      dr <- k*(theta-r[i-1,j])*dt + beta*sqrt(dt)*rnorm(1,0,1)
      r[i,j] <- r[i-1,j] + dr
    }
  }

  ## Construction de la base annuelle des taux
  R <- matrix(1, n, T)

  for(i in 1:n){
    for(j in 1:T)
      R[i,j] <- r[(360/T)*j,i]
  }
}

# Simulation des trajectoires GEV
library(evd)
u <- 49.834835 # paramètre de position
kh <- 6.987297 # paramètre d'échelle
esp <- 0.153812 # paramètre de forme

X_gev <- matrix(0, n, T) # Initialization des données GEV
for (k in 1:3)
  { X_gev[, k] <- rgev(n, u,kh,esp) } # Génération des trajectoires GEV

# Calcul du prix du CAT Bond
Coup <- matrix(0, n, T)
sinis_gev <- matrix(0, n, T)
princ_gev <- matrix(principal, n, 1)
flux_gev <- matrix(0, n, 1)
for (j in 1:T)
  {
    if (X_gev[i, j] <= class_gev[1])
      { sinis_gev[i, j] = 0
        princ_gev[i] <- max(princ_gev[i] - sinis_gev[i, j],0)
        Coup[i,j] <- (spread + c)*princ_gev[i]
      }
    if (X_gev[i, j] > class_gev[1] && X_gev[i, j] <= class_gev[2])
      { sinis_gev[i, j] <- min(a2*principal, princ_gev[i])
        princ_gev[i] <- max(princ_gev[i] - sinis_gev[i, j],0)
        Coup[i,j] <- (spread + c)*princ_gev[i]
      }
    if (X_gev[i, j] > class_gev[2] && X_gev[i, j] <= class_gev[3])
      { sinis_gev[i, j] <- min(a3*principal, princ_gev[i])
        princ_gev[i] <- max(princ_gev[i] - sinis_gev[i, j],0)
        Coup[i,j] <- (spread + c)*princ_gev[i]
      }
    if (X_gev[i, j] > class_gev[3] && X_gev[i, j] <= class_gev[4])
      { sinis_gev[i, j] <- min(a4*principal, princ_gev[i])
        princ_gev[i] <- max(princ_gev[i] - sinis_gev[i, j],0)
        Coup[i,j] <- (spread + c)*princ_gev[i]
      }
    if (X_gev[i, j] > class_gev[4] && X_gev[i, j] <= class_gev[5])
      { sinis_gev[i, j] <- min(a5*principal, princ_gev[i])
        princ_gev[i] <- max(princ_gev[i] - sinis_gev[i, j],0)
        Coup[i,j] <- (spread + c)*princ_gev[i]
      }
    if (X_gev[i, j] > class_gev[5])
      { sinis_gev[i, j] <- min(a6*principal, princ_gev[i])
        princ_gev[i] <- max(princ_gev[i] - sinis_gev[i, j],0)
        Coup[i,j] <- (spread + c)*princ_gev[i]
      }
    if (j == T && princ_gev[i]>0)
      { Coup[i,j] <- Coup[i,j] + princ_gev[i]
      }
    flux_gev[i] = (1/(1 + R[i,1]))*Coup[i,1] +
      (1/((1 + R[i,1])*(1 + R[i,2])))*Coup[i,2] +
      (1/((1 + R[i,1])*(1 + R[i,2])*(1 + R[i,3])))*Coup[i,3]
  }
}
Price=mean(flux_gev)

```

Model LSTM :

```

def LSTM_model(input_size, output_size, num_features):
    model = Sequential()
    model.add(BatchNormalization(input_shape = (input_size, num_features)))
    model.add(LSTM(units = 50, kernel_regularizer=keras.regularizers.l2(0.01), return_sequences = True))

    model.add(Dropout(0.2))
    model.add(BatchNormalization())

    model.add(LSTM(units = 50, kernel_regularizer=keras.regularizers.l2(0.01),
                    return_sequences = False))

    model.add(Dropout(0.2))
    model.add(BatchNormalization())

    model.add(RepeatVector(output_size))

    model.add(LSTM(units = 50, kernel_regularizer=keras.regularizers.l2(0.01),
                    return_sequences = True))

    model.add(Dropout(0.2))
    model.add(BatchNormalization())

    model.add(LSTM(units = 50, kernel_regularizer=keras.regularizers.l2(0.01),
                    return_sequences = False))

    model.add(Dense(units=output_size, activation = 'relu',
                    kernel_regularizer=keras.regularizers.l1_l2(0.001, 0.001)))

    model.compile(loss = "mse", optimizer = RMSprop(),
                  metrics=[tf.keras.metrics.RootMeanSquaredError(name="rmse"),
                           tf.keras.metrics.MeanAbsoluteError(name="mae")])

    model.summary()
    return model

```

Algorithme – Prix CAT Bond : Modèle LSTM

```

1  # Calcul du principal restant des différentes trajectoires du CAT Bond
2  Principal_Calcul <- function()
3  {
4    Pr_1 <- matrix(principal, 3, 6)
5    Pr_2 <- matrix(principal, 3, 6)
6    Pr_3 <- matrix(principal, 3, 6)
7    for (m in 1:6)
8    {
9      for (z in 1:6)
10     {
11       for (k in 1:6)
12       {
13         Pr_1[2,k] = max(Pr_1[1, k] - coef_a[k]*Pr_1[1, k],0)
14         Pr_2[1,k] = max(Pr_1[1, k] - coef_a[k]*Pr_1[1, k],0)
15         Pr_3[1,k] = max(Pr_1[1, k] - coef_a[k]*Pr_1[1, k],0)
16       }
17       Pr_2[2,z] = max(Pr_2[1, z] - coef_a[z]*Pr_2[1, z],0)
18       Pr_3[2,z] = max(Pr_2[1, z] - coef_a[z]*Pr_2[1, z],0)
19     }
20     Pr_3[3,m] = max(Pr_3[2, m] - coef_a[m]*Pr_3[2, m],0)
21   }
22 }
23 # Calcul des coupons de chaque trajectoire du CAT Bond
24 Coupon_Calcul <- function()
25 {
26   fCAT_1 <- 0
27   fCAT_2 <- 0
28   fCAT_3 <- 0
29   fCAT_Pr <- 0
30
31   for (k in 1:6)
32   {
33     fCAT_1 = fCAT_1 + CAT_Results[1, k] * (spread + c) * max(Pr_1[1, k] -
34     coef_a[k]*Pr_1[1, k],0)
35     for (z in 1:6)
36     {
37       fCAT_2 = fCAT_2 + CAT_Results[1, k]*CAT_Results[2, z] * (spread + c) *
38       max(Pr_2[1, z] - coef_a[z]*Pr_2[1, z],0)
39       for (m in 1:6)
40       {
41         fCAT_3 = fCAT_3 +
42         CAT_Results[1, k]*CAT_Results[2, z]*CAT_Results[3, m] * (spread + c)
43         * max(Pr_3[2, m] - coef_a[m]*Pr_3[2, m],0)
44         fCAT_Pr = fCAT_Pr +
45         CAT_Results[1, k]*CAT_Results[2, z]*CAT_Results[3, m]
46         * max(Pr_3[2, m] - coef_a[m]*Pr_3[2, m],0)
47       }
48     }
49   }
50 # Calcul du prix du CAT Bond
51 Price_Compute <- function()
52 {
53   for (i in 1:n)
54   {
55     flux_gev[i] = (1/(1 + IR[i,1]))*fCAT_1 +
56     (1/((1 + IR[i,1])*(1 + IR[i,2])))*fCAT_2 +
57     (1/((1 + IR[i,1])*(1 + IR[i,2])*(1 + IR[i,3])))*(fCAT_3 + fCAT_Pr)
58   }
59 }
60 Principal_Compute()
61 Coupon_compute()
62 Price_Compute()
63 P=mean(flux_gev)

```


Annexe I : Rappels Finance & statistiques

Annexe I-1. Estimation par maximum de vraisemblance

Soit (X_n) une suite de variables indépendantes, on définit la fonction de vraisemblance :

- Cas des v.a. discrètes :

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n P(x = x_i | \theta)$$
- Cas des v.a. continues de densité $f(x, \theta)$

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

On appelle estimateur du maximum de vraisemblance (EMV) toute fonction de l'estimateur $\hat{\theta}$ qui vérifie : $L(x_1, \dots, x_n; \hat{\theta}) = \max L(x_1, \dots, x_n; \theta)$

Annexe I-2. Processus de poisson doublement stochastique ou processus de Cox

Soit $\{N(t), t \geq 0\}$ un processus de poisson de paramètre $\lambda(t)$.

On définit un processus de Cox ou processus de poisson doublement stochastique comme étant un processus poisson dans lequel $\{\lambda(t), t \geq 0\}$ est lui-même un processus stochastique.

Annexe I-3. Loi forte des grands nombres

Soit $\{X_i\}_{i=1, \dots, n}$ une suite de variables indépendantes et de même loi, et identiquement distribuées admettant une espérance notée $E(X_i) = m$ avec $i=1 \dots n$.

On note :

$$\bar{X}_n = \frac{1}{n} * \sum_{i=1}^n X_i$$

Alors la suite \bar{X}_n tend vers m presque sûrement quand $n \rightarrow \infty$.

$$\bar{X}_n \xrightarrow[n \rightarrow \infty]{ps} m$$

Annexe I-4. Théorème de l'espérance totale

Soit X, Y deux variables aléatoires.

On a : $E(X) = E(E(X|Y))$

- Soient $\{A_i\}_{i=1, \dots, n}$ des événements deux à deux disjoints et leur union forme l'univers. On peut écrire :

$$E(X) = \sum_{i=1}^n E(X|A_i) * P(A_i)$$

Annexe II : Modélisation avec une approche par agent représentatif

Annexe II-1 : Démonstrations de la dérivée de Radon-Nikodym

On a par définition dans la section 1.5.2:

$$\frac{1}{1+r_t} = \frac{1}{u'_t(C(t))} E^P[u'_{t+1}(C(t+1)) | \mathcal{F}_t]$$

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_T} \stackrel{\text{def}}{=} \prod_{k=0}^{T-1} [1+r_k] \frac{u'_T(C(T))}{u'_0(C(0))}$$

Soit :

$$\xi(k) \stackrel{\text{def}}{=} E \left[\frac{dQ}{dP} \Big|_{\mathcal{F}_T} \Big| \mathcal{F}_k \right]$$

Soit un processus stochastique définit par :

$$\zeta(k) = \begin{cases} \prod_{j=0}^{k-1} (1+r_j) \frac{u'_k(C(k))}{u'_0(C(0))} & \text{pour } k = 1, \dots, T \\ 1 & \text{pour } k = 0 \end{cases}$$

On peut remarquer que la relation suivante:

$$\zeta(T) = \frac{dQ}{dP} \Big|_{\mathcal{F}_T} = \xi(T)$$

On peut aussi établir une relation entre $\zeta(k)$ et $\zeta(k+1)$, On a :

$$\begin{aligned} \zeta(k+1) &= \prod_{j=0}^k (1+r_j) \frac{u'_{k+1}(C(k+1))}{u'_0(C(0))} \\ &= (1+r_k) \prod_{j=0}^{k-1} (1+r_j) \frac{u'_{k+1}(C(k+1))}{u'_0(C(0))} \frac{u'_k(C(k))}{u'_k(C(k))} \\ &= (1+r_k) \zeta(k) \frac{u'_{k+1}(C(k+1))}{u'_k(C(k))} \end{aligned}$$

On peut en déduire que $\zeta(k)$ est une martingale, On a :

$$\begin{aligned} E[\zeta(k+1) | \mathcal{F}_k] &= E \left[(1+r_k) \zeta(k) \frac{u'_{k+1}(C(k+1))}{u'_k(C(k))} \Big| \mathcal{F}_k \right] \\ &= \zeta(k) (1+r_k) E \left[\frac{u'_{k+1}(C(k+1))}{u'_k(C(k))} \Big| \mathcal{F}_k \right] \end{aligned}$$

Avec la définition de $r(t)$, on trouve :

$$E[\zeta(k+1) | \mathcal{F}_k] = \zeta(k) (1+r_k) * \frac{1}{1+r_k} = \zeta(k)$$

On peut en déduire que : $E[\xi(T)] = E[\zeta(T)] = E[\zeta(0)] = 1$.

Ce qui implique

$$E \left[\frac{dQ}{dP} \right] = 1$$

Par conséquent Q est une mesure de probabilité.

Annexe II-2 : Changement de probabilité de l'expression du prix

On a :

$$V_{CAT}(0) = E^P \left[\sum_{k=1}^T \frac{u'_k(C(k))}{u'_0(C(0))} F_{CAT}(k) \right]$$

En isolant le terme à l'intérieur de la somme, on a :

$$E^P \left[\frac{u'_k(C(k))}{u'_0(C(0))} F_{CAT}(k) \right] = E^P \left[\frac{1}{\prod_{j=0}^{k-1} (1+r(j))} \zeta(k) F_{CAT}(k) \right]$$

On pose : $B(k) = \prod_{j=0}^{k-1} (1+r_j)$

On remplace $\zeta(k)$ par $\xi(k)$. Cela est une conséquence du fait $\zeta(k)$ est une martingale
Avec

$$\xi(k) \stackrel{\text{def}}{=} E \left[\frac{dQ}{dP} \Big|_{\mathcal{F}_T} \Big| \mathcal{F}_k \right]$$

$$\begin{aligned} E^P \left[\frac{u'_k(C(k))}{u'_0(C(0))} F_{CAT}(k) \right] &= E^P \left[\frac{\xi(k)}{B(k)} F_{CAT}(k) \right] \\ &= E^P \left[E \left[\frac{dQ}{dP} \Big|_{\mathcal{F}_T} \Big| \mathcal{F}_k \right] \frac{1}{B(k)} F_{CAT}(k) \right] \\ &= E^P \left[E \left[\frac{dQ}{dP} \Big|_{\mathcal{F}_T} \frac{1}{B(k)} F_{CAT}(k) \Big| \mathcal{F}_k \right] \right] \end{aligned}$$

par déduction en utilisant l'espérance totale, on a:

$$= E^P \left[\frac{dQ}{dP} \Big|_{\mathcal{F}_T} \frac{1}{B(k)} F_{CAT}(k) \right]$$

En effectuant le changement de probabilité, on retrouve ce que nous cherchons à démontrer

- [1]. Yu. Baryshnikov, A. Mayo and D. R. Taylor [1998], Pricing of CAT Bond
- [2]. Cox, S. H. and Pedersen, H. W. [2000]. Catastrophe risk bonds. *North American Actuarial Journal*, 4(4): p. 56–82.
- [3]. Cummins, J. David and Lalonde, David and Phillips, Richard D. [2000], The Basis Risk of Catastrophic-Loss Index Securities
- [4]. K. Burnecki, G. Kukla (2003), Pricing of zero-coupon and coupon CAT Bond, *Applicationes Mathematicae* 30, p. 315–324.
- [5]. Zimbidis, A. A., Frangos, N. E., and Pantelous, A. A. (2007). Modeling earthquake risk via extreme value theory and pricing the respective catastrophe bonds. *ASTIN bulletin*, 37(1): p. 163–184.
- [6]. David Cummins (2008), CAT Bond and Other Risk-linked Securities: State of the Market and Recent Developments, *Risk Management and Insurance Review*, volume 11, p. 23 – 47
- [7]. Härdle, Wolfgang K. and Cabrera, Brenda López [2010], Calibrating CAT Bond for Mexican Earthquakes. *Journal of Risk and Insurance*, Vol. 77, Issue 3, pp. 625-650
- [8]. Shi, Xingjian, Chen, Zhouong, Wang, Hao, Yeung, Dit-Yan, Wong, Wai-Kin, and Woo, Wang-chun [2015]. Convolutional LSTM network: A machine learning approach for precipitation nowcasting, *Advances in Neural Information Processing Systems*, volume 28, page 802-810
- [9]. S. Hochreiter and J. Schmidhuber. [1997]. Long short-term memory. *Neural Computation*, 9(8):1735–1780.
- [10]. KyungHyun Cho and Bart van Merriënboer and Dzmitry Bahdanau and Yoshua Bengio [2014], On the Properties of Neural Machine Translation: Encoder-Decoder Approaches
- [11]. Scott Lundberg and Su-In Lee [2017], A Unified Approach to Interpreting Model Predictions, *Advances in Neural Information Processing Systems*, Pages 4765–4774
- [12]. Andy Polacek [2018], Catastrophe bonds: A primer and retrospective, *Chicago Fed Letter*, 2018 NUMBER 405
- [13]. Swiss Re Institute [2020], Natural catastrophes in times of economic accumulation and climate change, *Sigma*, n°2/2020

BIBLIOGRAPHIE

Sites internet:

1. <https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn>
2. <https://www.artemis.bm>