

Mémoire présenté devant l'ENSAE Paris
pour l'obtention du diplôme de la filière Actuariat
et l'admission à l'Institut des Actuaire
le 15/03/2023

Par : **Comlan Rodrigue TOSSOU**

Titre : **Risk Weighted Assets : Estimation de l'effet des paramètres**

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

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

Membres présents du jury de la filière

Entreprise : Société Générale

Nom : Caroline HILLAIRET

Signature :

*Membres présents du jury de l'Institut
des Actuaire*

Directeur du mémoire en entreprise :

Nom : Céline PIQUET

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)**

Secrétariat :

Signature du responsable entreprise

Bibliothèque :

Signature du candidat



RISK WEIGHTED ASSETS : ESTIMATION DE L'EFFET DES PARAMÈTRES

RAPPORT CONFIDENTIEL

Rédigé par :

Comlan Rodrigue TOSSOU

Tutrice entreprise :

Céline PIQUET

Tutrice académique :

Caroline Hillairet

Mars 2023

RÉSUMÉ

La globalisation a entraîné ces dernières années une interdépendance très accrue entre les pays, rendant ainsi la propagation de toute défaillance financière très aisée. Pour prévenir les crises futures, le comité de Bâle a donc émis une formule de mesure des exigences en fonds propres nécessaires à faire face dans 99,9% des cas aux pertes extrêmes potentielles. Il s'agit de la formule réglementaire de calcul du risk weighted assets (RWA). Cette formule, issue du modèle monofactoriel de Gordy-Vasicek, repose sur un ensemble d'hypothèses en faible concordance avec la réalité économique. Une alternative serait de faire appel à une généralisation multifactorielle de la formule réglementaire. Cette solution, bien que pertinente, induit une complexité plus grande dans l'évaluation du risque de crédit.

Ainsi, nous proposons une estimation pertinente des effets des paramètres de la formule réglementaire qui aura l'intérêt de fournir une interprétation et un monitoring plus aisés du risque de crédit. Cette dernière information permettrait, d'un point de vue bancaire, d'avoir une idée très précise des modèles nécessitant une attention plus particulière. Pour répondre à cette problématique, ce mémoire est subdivisé en deux parties. La première fait une revue de l'évaluation du risque de crédit dans un contexte bâlois. La deuxième partie s'articule autour de deux axes. Le premier sert à l'approximation de la formule réglementaire de manière analytique et par modélisation. Le second fournit les outils nécessaires à l'estimation de l'effet des paramètres de la formule réglementaire sur les variations du risque.

Mots clés : *Risk Weighted Assets (RWA), risque de crédit, Bâle, effet marginal, machine learning.*

ABSTRACT

Globalization has in recent years led to a very increased interdependence between countries, making the spread of any financial failure very easy. In order to prevent future crises, the Basel Committee therefore issued a formula for measuring the capital requirements needed to deal with potential extreme losses in 99.9% of cases. This is the regulatory formula for calculating risk weighted assets (RWA). This formula derived from the Gordy-Vasicek monofactorial model is based on a set of hypotheses that are not in line with economic reality. An alternative would be to use a multi-factor generalization of the regulatory formula. This, while relevant, solution leads to greater complexity in the assessment of credit risk.

Thus, we propose a relevant estimation of the effects of the parameters of the regulatory formula that will have the advantage of providing easier interpretation and monitoring of credit risk. From a banking point of view, this latter information would give a very precise idea of the models requiring more attention. To address this issue, this brief is divided into two parts. The first part reviews the assessment of credit risk in a Basel context. The second part focuses on two areas. The first is used to approximate the regulatory formula analytically and by modelling. The second axis provides the tools for estimating the effect of the regulatory formula parameters on risk variations.

Key words : *Risk Weighted Assets (RWA), credit risk, Basel, marginal effect, machine learning.*

REMERCIEMENTS

Je souhaite tout d'abord témoigner toute ma gratitude à ma tutrice Céline PIQUET, Superviseur Projets Calculs, Modèles et Octroi, pour non seulement m'avoir permis de réaliser mon alternance au sein de la Société Générale, mais aussi pour tous les efforts consentis pour me fournir un cadre professionnel idéal. Mes remerciements vont également à l'ensemble de l'équipe Calculateurs et particulièrement à Samir BEDREDDINE pour son encadrement et son implication dans l'ensemble de mes travaux.

J'adresse un immense merci à Caroline HILLAIRET, ma tutrice académique, pour la rigueur scientifique dont j'ai bénéficié tout le long de la rédaction de ce mémoire. En acceptant de me superviser, elle n'a point cessé de m'accorder sa disponibilité, ses bons conseils et remarques malgré son emploi du temps chargé.

Je ne saurais remercier à juste titre toutes les personnes ayant contribué à la réalisation de ce travail sans faire une mention spéciale à tous mes collègues de Detralytics. Votre implication dans la préparation et la finalisation de ce mémoire m'a particulièrement touchée. Ce travail s'est vu énormément bonifié grâce à vous.

Enfin, j'aimerais adresser de profonds remerciements à ma famille et à mes amis les plus chers pour le soutien sans faille qu'ils n'ont de cesse de m'apporter. Sans vous, ce travail n'aurait pas vu le jour.

Merci.

NOTE DE SYNTHÈSE

L'environnement économique d'un pays est intrinsèquement lié à sa santé financière et à celle du monde qui l'entoure. Cette situation crée une interdépendance entre la vie économique d'une part et la vie financière d'autre part. Les différentes crises économiques passées ont très souvent eu une origine financière. En effet, la globalisation a entraîné ces dernières années une interdépendance très accrue entre les pays ; rendant ainsi la propagation de toute défaillance très aisée. Pour prévenir les crises futures, il a fallu édicter des règles de prudence que chaque acteur se doit de respecter afin de préserver l'environnement économique. C'est dans cet état d'esprit que le comité de Bâle a vu le jour. Son rôle est d'émettre les règles de gestion des risques financiers. Une formule de mesure des exigences en fonds propres nécessaires pour faire face dans 99,9% des cas aux pertes extrêmes potentielles a donc été décrétée. Il s'agit de la formule réglementaire de calcul du *risk weighted assets (RWA)*.

Genèse et critique de la formule réglementaire

Le RWA, ou risque pondéré, reflète le niveau de risque encouru par la banque sur chacun de ces contrats. Il se base sur le montant de l'exposition à défaut (EAD), la probabilité de défaut (PD) du tiers, le taux de perte en cas de défaut (LGD) et la maturité (M) du contrat pour fournir une mesure de risque cohérente.

$$RWA = EAD \cdot LGD \left[\Phi \left(\frac{\Phi^{-1}(PD) + \sqrt{R}\Phi^{-1}(0,999)}{\sqrt{1-R}} \right) - PD \right] \frac{1 + (M - 2,5)b}{1 - 1,5b} \cdot 12,5 \cdot 1,06$$

Avec $R = 0,12 \cdot \left(\frac{1 - e^{-50 \cdot PD}}{1 - e^{-50}} \right) + 0,24 \cdot \left(1 - \frac{1 - e^{-50 \cdot PD}}{1 - e^{-50}} \right)$ et $b = (0,11852 - 0,05478 \cdot \ln(PD))^2$. Cette formule, issue du modèle monofactoriel de **Gordy-Vasicek**, repose sur un ensemble d'hypothèses en faible concordance avec la réalité économique. Trois principales critiques sont formulées :

- l'hypothèse de portefeuille parfaitement granulaire¹ n'est que très rarement respectée. Les banques font très souvent face à une concentration de leur risque vis-à-vis d'une même contrepartie. Même si l'avènement de Bâle III apporte des outils supplémentaires pour

1. Existence d'un très grand nombre de contreparties avec des expositions négligeables vis-à-vis de l'ensemble du portefeuille.



l'évaluation et la gestion du risque de concentration, il n'en demeure pas moins que la formule actuelle de calcul du risque de crédit ne tient pas compte des effets négatifs du risque de concentration.

- l'approche monofactorielle du risque commun découlant de la seule conjoncture macroéconomique est une vision limitée de l'environnement économique. Il serait plus vraisemblable de déterminer pour chaque secteur les facteurs de risque systématique impactants.
- d'un point de vue théorique, une confusion entre probabilités de défaut conditionnelle et inconditionnelle semble parfois commise dans les estimations. Cela engendre une incohérence avec l'inconditionnalité dans ce modèle.

Ces limites nous poussent à questionner la forme de la fonction réglementaire de calcul du risque de crédit. Une alternative serait de faire appel à une généralisation multifactorielle de la formule réglementaire. Cette solution, bien que pertinente, induit une complexité plus grande dans l'évaluation du risque de crédit. Ainsi, nous proposons de construire une fonction alternative à l'actuelle qui aura l'intérêt de fournir une interprétation et un monitoring plus aisés du risque de crédit.

Aussi, les différents modèles internes utilisés par la banque pour l'estimation de la probabilité de défaut et de la perte en cas de défaut sont soumis à validation par le régulateur, et de ce fait, sujets à pénalités en cas de mauvaise qualité. On parle d'ajout d'**Add-on** sur les drivers. Ce processus revient à majorer le paramètre préalablement obtenu par les modèles internes d'un pourcentage supplémentaire pour un tant soit peu rectifier la potentielle sous-estimation du risque.

De plus, la première difficulté à l'utilisation de la formule réglementaire est l'anticipation sur le coût d'une éventuelle majoration de ses paramètres. La seconde difficulté survient dans la quantification de la contribution de chaque driver dans le niveau de risque obtenu. Cette dernière information permettrait, d'un point de vue bancaire, d'avoir une idée très précise des modèles nécessitant une attention plus particulière.

Pour répondre à cette problématique, deux principales approches sont possibles. La première, **analytique**, permet grâce à la formule de Taylor, de déterminer les effets marginaux de chacun des paramètres pour des variations infinitésimales. La seconde, quant à elle, consiste à approcher la formule réglementaire par un **modèle** qui servira à l'estimation des effets des paramètres.

Estimation des effets des paramètres par approche analytique

La formule de Taylor est un outil mathématique qui permet, à l'aide de certaines hypothèses, d'approcher une fonction régulière par une combinaison linéaire de ses dérivées partielles. Ainsi, si on considère f une fonction régulière définie sur un ouvert U de \mathbb{R}^n à valeurs dans \mathbb{R}^p , x_0 un



point de U , alors d'après **Taylor-Young**, pour tout $h \in \mathbb{R}^n$ tel que $x_0 + h \in U$

$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + \dots + \frac{1}{k!} D^k f(x_0)(h)^k + o(\|h\|^k)$$

où $D^k f(x_0)(h)^k = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x_0) h_{i_1} \dots h_{i_k}$ et $o(\|h\|^k)$ est une fonction qui tend vers 0 quand h tend vers 0.

En notant que $RWA = EAD \times RW$ avec RW le risque pondéré, une application directe à la formule réglementaire permet d'obtenir les relations suivantes aux deux premiers ordres :

- A l'ordre 1 :

$$\Delta RW_1 \approx \frac{\partial RW}{\partial PD} \Delta PD + \frac{\partial RW}{\partial LGD} \Delta LGD + \frac{\partial RW}{\partial M} \Delta M$$

- A l'ordre 2 :

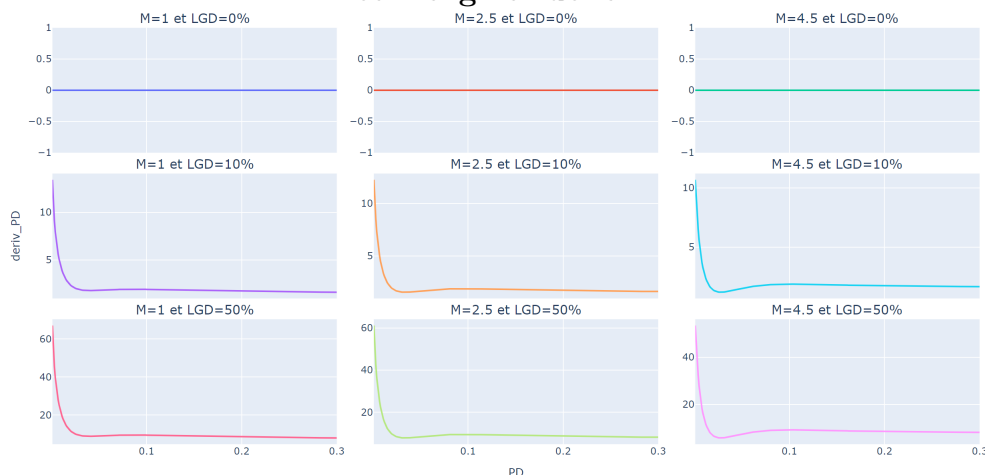
$$\begin{aligned} \Delta RW_2 \approx & \frac{\partial RW}{\partial PD} \Delta PD + \frac{\partial RW}{\partial LGD} \Delta LGD + \frac{\partial RW}{\partial M} \Delta M + \frac{1}{2} \cdot \frac{\partial^2 RW}{\partial PD^2} \Delta PD^2 \\ & + \frac{\partial^2 RW}{\partial PD \partial LGD} \Delta PD \cdot \Delta LGD + \frac{\partial^2 RW}{\partial PD \partial M} \Delta PD \cdot \Delta M + \frac{\partial^2 RW}{\partial LGD \partial M} \Delta LGD \cdot \Delta M \end{aligned}$$

avec $\frac{\partial^2 RW}{\partial LGD^2} = \frac{\partial^2 RW}{\partial M^2} = 0$

Cette approche est limitée dans son utilisation par l'hypothèse de variation infime. Toutefois, si nous la supposons vérifiée, les résultats suivants sont robustes. Les différentes dérivées partielles représentent les effets marginaux et sont fonctions des valeurs initiales des paramètres (PD, LGD et M).

De manière générale, on remarque que l'effet marginal de la PD $\left(\frac{\partial RW}{\partial PD}\right)$ décroît rapidement en fonction de la PD pour se stabiliser à un niveau assez faible. L'amplitude de variation est accentuée par des valeurs plus grandes de la maturité et de la perte en cas de défaut. L'effet d'une faible variation de la PD est d'autant plus fort que la valeur de PD originelle est faible. Ainsi, une dégradation de la qualité de crédit d'un emprunteur très bien noté présente un risque plus accru que la dégradation d'un emprunteur déjà mal noté.

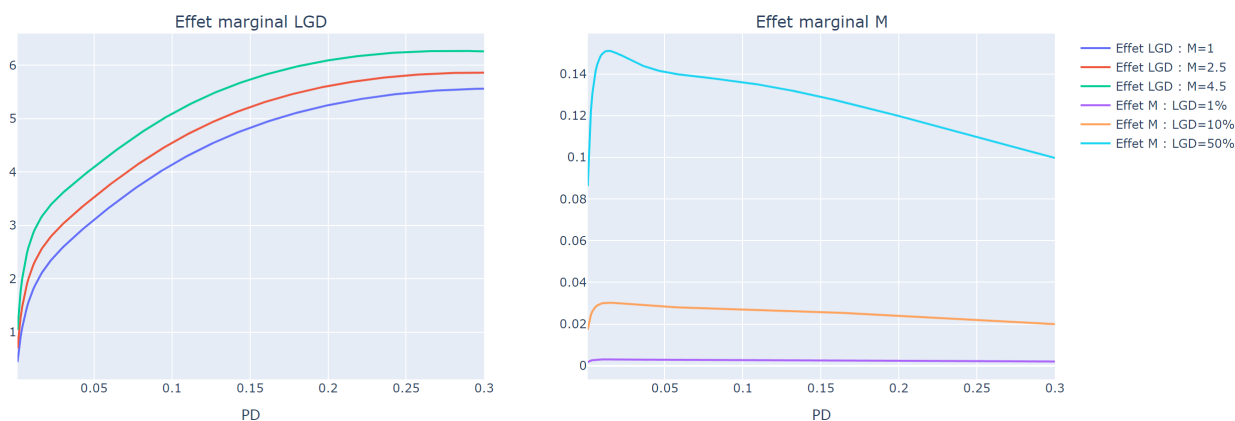
Effet marginal de la PD





L'effet marginal de la LGD $\left(\frac{\partial RW}{\partial LGD}\right)$ arbore une allure logarithmique. Une dégradation de la qualité de recouvrement d'un mauvais emprunteur² présente un risque plus accru que la dégradation de la qualité de recouvrement d'un bon emprunteur. En ce qui concerne l'effet marginal de la maturité $\left(\frac{\partial RW}{\partial M}\right)$, elle est fonction uniquement de la PD et de la LGD. On observe d'abord une croissance accélérée à mesure que la PD augmente avant d'avoir un effet inverse pour de fortes valeurs de la PD. La LGD influence l'amplitude de l'effet.

Effet marginal de la LGD et de la maturité (M)



Pour mesurer l'influence de chacun des paramètres dans les variations du niveau de risque, nous avons calculé leurs contributions. Ici, la contribution d'un paramètre correspond à la part de la variation du RW (approchée par Taylor) résultant de la variation du paramètre étudié. Le tableau suivant fournit, pour des valeurs fixées des paramètres, le niveau de risque pondéré associé ainsi que les informations suivantes pour une variation de 1% de tous les paramètres :

- l'approximation faite par la formule de Taylor aux deux premiers ordres de la variation du RW ;
- et les contributions des différentes dérivées partielles dans ces variations.

Contributions pour une variation de 1% de tous les paramètres

Valeurs initiales			RW	Variation du RW par approche analytique		Contribution des paramètres dans la variation du RW							
PD	LGD	M		delta_RW1	delta_RW2	contrib_PD	contrib_LGD	contrib_M	contrib_PD2	contrib_PD_LGD	contrib_PD_M	contrib_LGD_M	
0.011000	0.30	4.741713	0.876740	0.013876	0.013915	5.987424	63.007103	30.722828	0.000659	0.059874	-0.041654	0.307228	
0.005015	0.35	1.000000	0.430772	0.006492	0.006508	18.816831	66.189730	14.742436	0.000944	0.188168	-0.010826	0.147424	
0.046123	0.33	1.000000	0.994219	0.013592	0.013623	19.932851	72.981623	6.861454	0.009194	0.199329	-0.011142	0.068615	
0.011000	0.30	5.000000	0.900027	0.014310	0.014351	5.585545	62.715870	31.412264	0.000614	0.055855	-0.042589	0.314123	
0.011000	0.16	5.000000	0.480014	0.007632	0.007654	5.585545	62.715870	31.412264	0.000614	0.055855	-0.042589	0.314123	
0.137042	0.33	1.000000	1.544914	0.024515	0.024561	33.385286	62.901328	3.525964	0.045752	0.333853	0.000297	0.035260	
0.032596	0.33	1.000000	0.878314	0.011691	0.011712	16.651991	74.989693	8.173030	0.005428	0.166520	-0.015891	0.081730	
0.013200	0.30	1.497378	0.625997	0.008869	0.008886	14.084414	70.450627	15.274567	0.001859	0.140844	-0.023359	0.152746	
0.005015	0.30	1.398025	0.401966	0.006195	0.006212	16.510928	64.709120	18.508445	0.000828	0.165109	-0.013591	0.185084	
0.046123	0.30	1.000000	0.903835	0.012357	0.012384	19.932851	72.981623	6.861454	0.009194	0.199329	-0.011142	0.068615	

Pour une variation d'1% de tous les paramètres, on note une contribution moyenne de la LGD de 67,35%. La PD arrive ensuite avec une contribution moyenne de 15,64%. Pour la maturité, on peut noter une contribution très disparate. En effet, des valeurs très fortes (~ 5)

2. Emprunteur mal noté.



ont un effet très important tandis que des valeurs très faibles (~ 1) ont des effets très faibles. On note enfin que les effets joints ont des contributions marginales dans la variation du RW.

Approximation de la formule réglementaire par modélisation

Afin de nous affranchir de l'hypothèse de l'approche analytique (variation infime), nous avons décidé d'opter pour une approximation de la formule réglementaire par un **modèle**. Le choix s'est alors posé entre les modèles interprétables, comme la régression linéaire et polynomiale, et les modèles de machine learning. Le panel de modèles que nous avons testés se compose du modèle linéaire, du modèle polynomial (degrés 2, 3 et 4), du Random Forest, du XGBoost et des réseaux de neurones. Le modèle d'approximation retenu se doit d'avoir une qualité d'ajustement satisfaisante, de bonnes propriétés d'interprétation et une reproductibilité des principales caractéristiques d'une mesure de risque cohérente. Pour les comparer et donc sélectionner le meilleur modèle, nous nous référons aux métriques usuelles telles que le *biais*, le *Root Mean Square Error (RMSE)*, le *Mean Absolute Error (MAE)*, le *RMSE standard deviation ratio (RSR)* et le *Coefficient de détermination (R2)*. Le modèle retenu est celui possédant les meilleurs performances vis-à-vis des métriques.

Comparaison des modèles

	Modèle	Biais	RMSE	MAE	RSR	R2
<i>Modèles classiques</i>	Modèle linéaire	0.0002	0.1738	0.1270	0.343948	0.8817
	Modèle polynomial (degré 2)	0.0005	0.0845	0.0556	0.167332	0.9721
	Modèle polynomial (degré 3)	0.0004	0.0576	0.0342	0.114018	0.9870
	Modèle polynomial (degré 4)	0.0009	0.0450	0.0240	0.088882	0.9921
<i>Modèles de ML</i>	<i>Random Forest</i>	0.0004	0.0101	0.0020	0.020000	0.9996
	XGBoost	0.0006	0.0117	0.0052	0.022361	0.9995
<i>Modèle de DeepL</i>	Réseaux de neurones	-0.0113	0.0315	0.0155	0.062450	0.9980

Tout d'abord, le modèle linéaire ainsi que ses dérivées (polynomiaux) ont été exclus étant donné qu'ils ne vérifiaient pas l'une des caractéristiques principales d'une mesure de risque de crédit. En effet, ces modèles, de part leur construction (linéaire suivant les paramètres), peuvent conduire à des prédictions négatives du RW. De plus, ils ne sont pas en mesure de reproduire le caractère absorbant de la LGD³. Ces remarques nous ont poussés vers les modèles de machine learning, en particulier les modèles d'arbres. Les performances du Random Forest (qualité d'ajustement) font de lui le meilleur choix comme modèle d'approximation de la formule réglementaire.

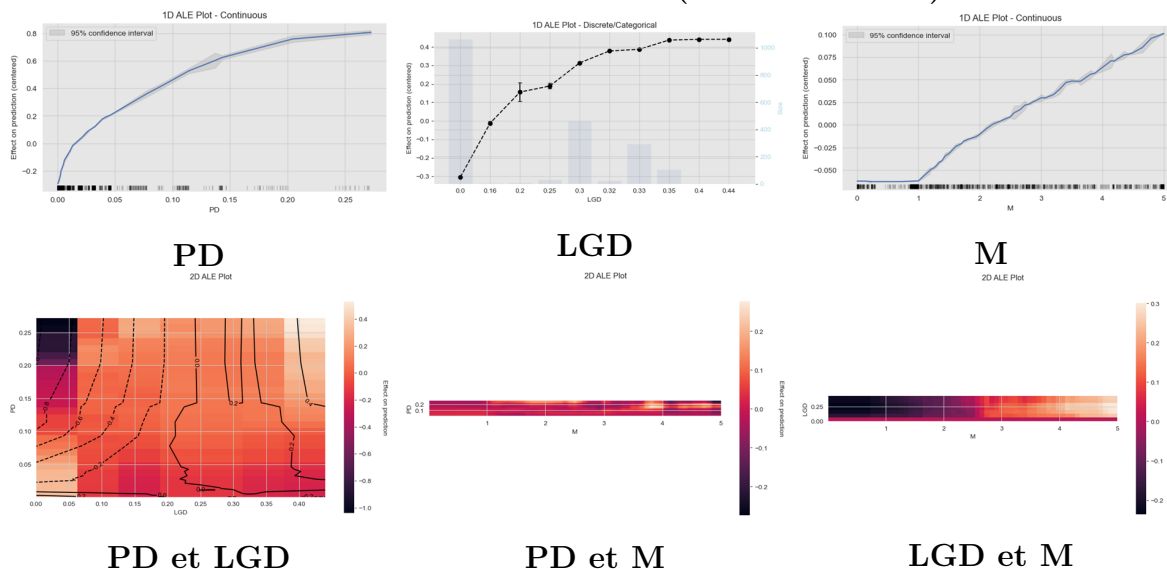
L'estimation des effets marginaux dans notre cas de figure par la Partial Dependence Plot (PDP) souffre d'une limite, l'indépendance des paramètres. Même si la corrélation entre la probabilité de défaut et la perte en cas de défaut est faible, elle existe. Ce constat, qui a tout

3. Une LGD nulle induit une valeur nulle de RW



son sens d'un point de vue financier, limite la fiabilité de la PDP dans l'évaluation des effets marginaux. Nous nous referons plutôt à la courbe des effets locaux accumulés (ALE) pour l'évaluation des effets marginaux.

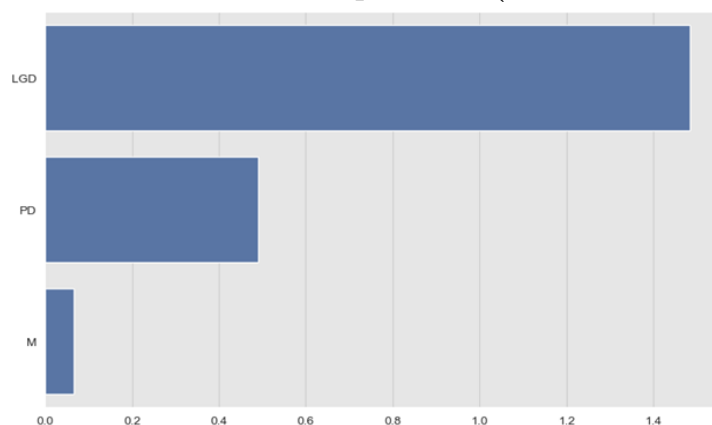
Accumulated Local Effects (Random Forest)



La probabilité de défaut possède un effet logarithmique tandis que la perte en cas de défaut et la maturité possèdent un effet plus ou moins linéaire. En effet, plus la PD est élevée, moins grand est l'impact de sa variation sur la variation du RW, tous les autres paramètres supposés constants.

Enfin, nous nous intéressons à une estimation des effets moyens de chacun des paramètres. L'importance des variables nous donne, en plus de cette information, une estimation des contributions relatives de chacune d'entre elles. On retrouve une plus grande importance de la perte en cas de défaut dans l'évaluation du risque de crédit. Cet effet est trois fois plus important que celui de la probabilité de défaut. La maturité quant à elle, a un effet assez marginal sur l'évaluation du risque, elle respecte sa propriété d'ajustement de maturité pour la prise en compte des effets de long terme.

Permutation Feature Importance(Random Forest)





Conclusion

En définitive, même si l'approximation de la formule réglementaire par machine learning ne permet pas d'obtenir une nouvelle formule simple d'évaluation du risque de crédit, nous recommandons son utilisation étant donné les résultats quantitatifs qu'elle apporte à l'estimation de l'effet des paramètres de la fonction réglementaire. En effet, il apparaît que la perte en cas de défaut est le paramètre le plus déterminant dans l'évaluation du risque vis-à-vis d'une ligne de contrat. Ainsi, l'effort de modélisation des paramètres de la fonction réglementaire se doit de prendre en compte les possibles effets négatifs que pourrait engendrer une mauvaise estimation de la perte en cas de défaut. Une attention particulière doit être portée à son estimation.

EXECUTIVE SUMMARY

The economic environment of a country is intrinsically linked to the financial health of the country and the rest of the world around it. This situation creates an interdependence between economic life on the one hand and financial life on the other. The various past economic crises have very often had a financial origin. Indeed, globalization has in recent years led to much greater interdependence between countries, making the spread of any failure very easy. To prevent future crises, it was necessary to establish rules of prudence that each actor must respect to preserve the economic environment. It was in this spirit that the Basel Committee was created. Its role is to issue financial risk management rules. A formula for measuring the capital requirements needed to meet 99.9% of the cases of potential extreme losses was therefore enacted. This is the regulatory formula for calculating *risk weighted assets (RWA)*.

Origin and criticism of the regulatory formula

The RWA or weighted risk reflects the level of risk incurred by the bank on each of its deals. It is based on the exposure at default (EAD), the third party's probability of default (PD), the loss given default (LGD) and the maturity (M) of the deal to provide a consistent risk measure.

$$RWA = EAD \cdot LGD \left[\Phi \left(\frac{\Phi^{-1}(PD) + \sqrt{R}\Phi^{-1}(0,999)}{\sqrt{1-R}} \right) - PD \right] \frac{1 + (M - 2,5)b}{1 - 1,5b} \cdot 12,5 \cdot 1,06$$

With $R = 0,12 \cdot \left(\frac{1 - e^{-50 \cdot PD}}{1 - e^{-50}} \right) + 0,24 \cdot \left(1 - \frac{1 - e^{-50 \cdot PD}}{1 - e^{-50}} \right)$ avec $b = (0,11852 - 0,05478 \cdot \ln(PD))^2$. This formula derived from the **Gordy-Vasicek** monofactorial model is based on a set of hypotheses that are not very consistent with economic reality. Three main criticisms are made :

- the perfectly granular portfolio assumption⁴ is not respected in very many cases. Banks very often end up with a concentration of their risk from the same counterparty. Although the advent of Basel III provides additional tools for assessing and managing concentration risk, the fact remains that the current credit risk calculation formula does not take into account the adverse effects of concentration risk ;

4. Existence of a very large number of counterparties with negligible exposures to the entire portfolio.



- the one-factor approach to common risk arising from the macroeconomic environment alone is a limited view of the economic environment. It would be more likely to determine for each sector the systematic risk factors impacting it.
- from a theoretical point of view, a confusion between conditional and unconditional default probabilities sometimes seems committed in the estimates. This creates an inconsistency with unconditionality in this model.

These limits lead us to question the form of the regulatory function for calculating credit risk. An alternative would be to use a multi-factor generalization of the regulatory formula. This solution, while relevant, leads to greater complexity in the assessment of credit risk. Thus, we propose to build an alternative function to the current one that will have the interest to provide a more easy interpretation and monitoring of credit risk.

Also, the various internal models used by the bank for estimating the probability of default and the loss in case of default are subject to validation by the regulator, and therefore subject to penalties in case of poor quality. We are talking about adding **Add-on** on the drivers. This process means to increase the parameter previously obtained by the internal models by an additional percentage in order to rectify the potential underestimation of risk.

In addition, the first difficulty in using the regulatory formula is anticipating the cost of a possible increase of parameters. The second difficulty occurs in quantifying the contribution of each driver in the level of risk obtained. From a banking point of view, this latter information would give a very precise idea of the models requiring more attention.

To address this issue, two main approaches are possible. The first, **analytic**, uses the Taylor formula to determine the marginal effects of each parameter for infinitesimal variations. The second is to approach the regulatory formula with a model for estimating the effects of the parameters.

Parameters's effect estimation by analytical approach

The Taylor formula is a mathematical tool that allows using certain hypotheses to approach a regular function by a linear combination of its partial derivatives. Thus, if we consider f a regular function defined on an open U of \mathbb{R}^n at values in \mathbb{R}^p , x_0 a point of U , then from **Taylor-Young**, for any $h \in \mathbb{R}^n$ such as $x_0 + h \in U$

$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + \cdots + \frac{1}{k!} D^k f(x_0)(h)^k + o(\|h\|^k)$$

where $D^k f(x_0)(h)^k = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(x_0) h_{i_1} \cdots h_{i_k}$ and $o(\|h\|^k)$ is a function that tends to 0 when h tends to 0.



Considering $RWA = EAD \times RW$ with RW the risk weighted, the direct application to the regulatory formula establishes the following links at the first two orders. :

- **Order 1 :**

$$\Delta RW_1 \approx \frac{\partial RW}{\partial PD} \Delta PD + \frac{\partial RW}{\partial LGD} \Delta LGD + \frac{\partial RW}{\partial M} \Delta M$$

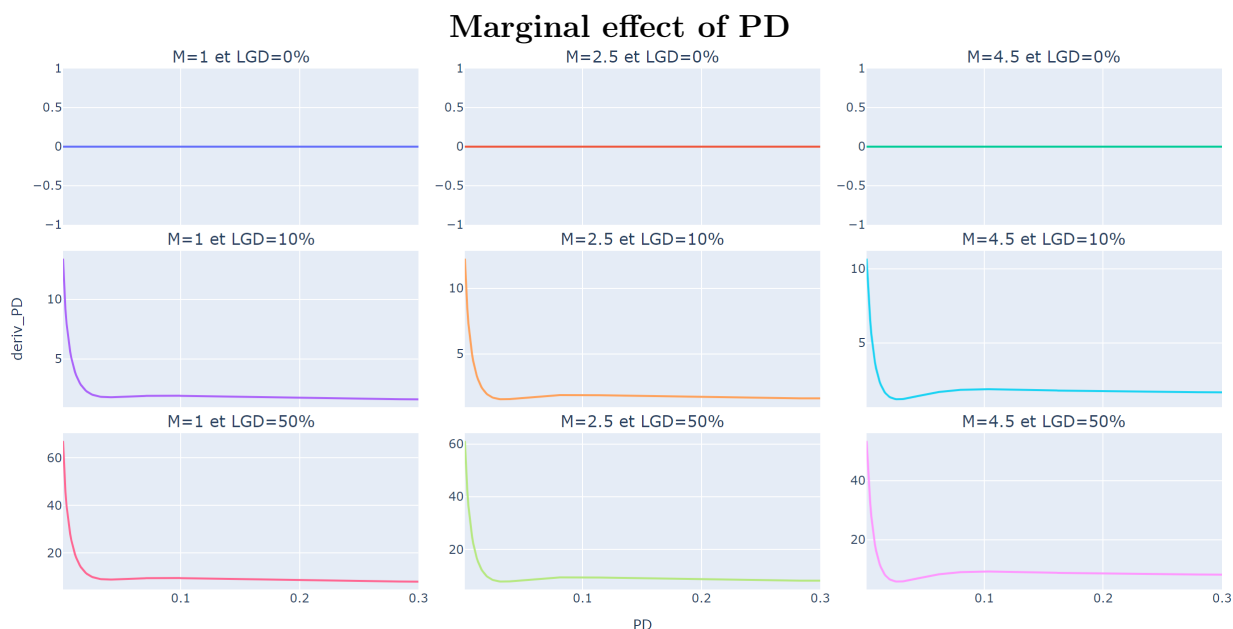
- **Order 2 :**

$$\begin{aligned} \Delta RW_2 \approx & \frac{\partial RW}{\partial PD} \Delta PD + \frac{\partial RW}{\partial LGD} \Delta LGD + \frac{\partial RW}{\partial M} \Delta M + \frac{1}{2} \cdot \frac{\partial^2 RW}{\partial PD^2} \Delta PD^2 \\ & + \frac{\partial^2 RW}{\partial PD \partial LGD} \Delta PD \cdot \Delta LGD + \frac{\partial^2 RW}{\partial PD \partial M} \Delta PD \cdot \Delta M + \frac{\partial^2 RW}{\partial LGD \partial M} \Delta LGD \cdot \Delta M \end{aligned}$$

with $\frac{\partial^2 RW}{\partial LGD^2} = \frac{\partial^2 RW}{\partial M^2} = 0$

This approach is limited in its use by the assumption of minimal variation. However, assuming verified, the following results are robust. The various partial derivatives represent marginal effects and are functions of the initial parameters values (PD, LGD and M).

In general, it is noted that the marginal effect of PD $\left(\frac{\partial RW}{\partial PD}\right)$ decreases rapidly with PD to stabilize at a fairly low level. The amplitude of variation is accentuated by higher values of maturity and loss in case of default. The lower the original PD value is, the stronger the effect of a small variation of PD will be. Thus, a deterioration in the credit quality of a highly rated borrower presents a greater risk than a deterioration of an already poorly rated borrower.



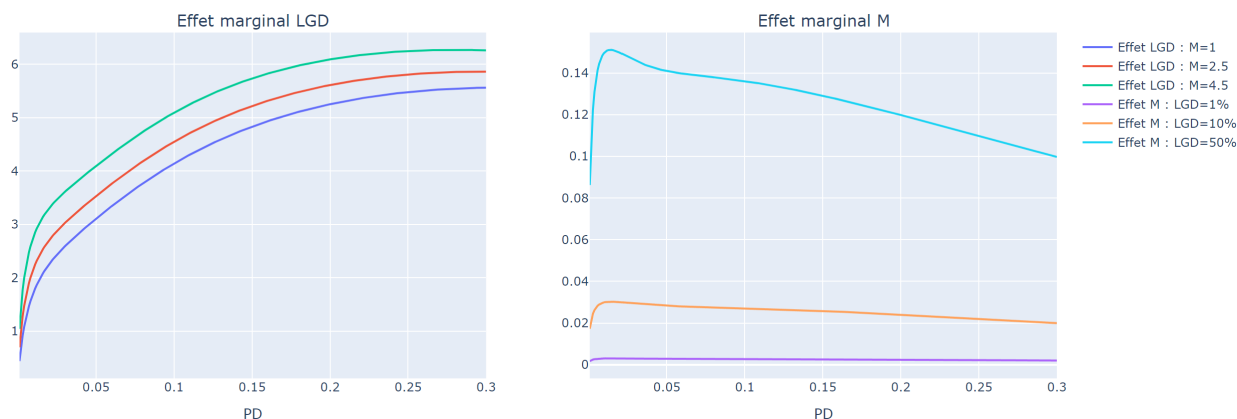
The marginal effect of LGD $\left(\frac{\partial RW}{\partial LGD}\right)$ is logarithmic in appearance. A deterioration in the quality of recovery of a bad borrower ⁵ presents a greater risk than the deterioration in the quality of a good borrower's collection. For the marginal maturity effect $\left(\frac{\partial RW}{\partial M}\right)$, it is only a

5. Borrower with bad rating.



function of the PD and LGD. We first observe an accelerated growth as the PD increases before having an inverse effect for strong PD values. LGD influences the amplitude of the effect.

Marginal effect of LGD and maturity (M)



To measure the influence of each of the parameters in the variations in the level of risk, we calculate their contributions. Here, the contribution of a parameter corresponds to the share of the RW variation (approached by Taylor) resulting from the variation of the studied parameter. The following table provides for fixed values of these parameters, the associated weighted risk level as well as the following informations for a 1% variation of all parameters :

- Taylor’s formula approximation to the first two orders of RW variation ;
- the contributions of the various partial derivatives in these variations.

Contributions for 1% variation of all parameters on actual data

Initial values			Variation of RW by analytical approach			Parameters contribution in RW variation							
PD	LGD	M	RW	delta_RW1	delta_RW2	contrib_PD	contrib_LGD	contrib_M	contrib_PD2	contrib_PD_LGD	contrib_PD_M	contrib_LGD_M	
0.011000	0.30	4.741713	0.876740	0.013876	0.013915	5.987424	63.007103	30.722828	0.000659	0.059874	-0.041654	0.307228	
0.005015	0.35	1.000000	0.430772	0.006492	0.006508	18.816831	66.189730	14.742436	0.000944	0.188168	-0.010826	0.147424	
0.046123	0.33	1.000000	0.994219	0.013592	0.013623	19.932851	72.981623	6.861454	0.009194	0.199329	-0.011142	0.068615	
0.011000	0.30	5.000000	0.900027	0.014310	0.014351	5.585545	62.715870	31.412264	0.000614	0.055855	-0.042589	0.314123	
0.011000	0.16	5.000000	0.480014	0.007632	0.007654	5.585545	62.715870	31.412264	0.000614	0.055855	-0.042589	0.314123	
0.137042	0.33	1.000000	1.544914	0.024515	0.024561	33.385286	62.901328	3.525964	0.045752	0.333853	0.000297	0.035260	
0.032596	0.33	1.000000	0.878314	0.011691	0.011712	16.651991	74.989693	8.173030	0.005428	0.166520	-0.015891	0.081730	
0.013200	0.30	1.497378	0.625997	0.008869	0.008886	14.084414	70.450627	15.274567	0.001859	0.140844	-0.023359	0.152746	
0.005015	0.30	1.398025	0.401966	0.006195	0.006212	16.510928	64.709120	18.508445	0.000828	0.165109	-0.013591	0.185084	
0.046123	0.30	1.000000	0.903835	0.012357	0.012384	19.932851	72.981623	6.861454	0.009194	0.199329	-0.011142	0.068615	

For a 1% change in all parameters, there is an average LGD contribution of 67.35%. The PD then arrives with an average contribution of 15.64%. For maturity, we can note a very disparate contribution. Indeed, very strong values (~ 5) have a very important while very low values (~ 1) have very weak effects. Finally, it is noted that the joint effects have marginal contributions in the variation of the RW.

Approximation of the regulatory formula using modeling

In order to move away from the analytical approach assumption, we decided to take a model-based approximation of the regulatory formula. The choice was then made between



interpretable models such as linear and polynomial regression and machine learning models. The panel of models we tested consists of linear model, polynomial model (degrees 2, 3 and 4), Random Forest, XGBoost and neural networks. The selected approximation model must have a satisfactory fit quality, good interpretive properties and reproducibility of the main characteristics of a consistent risk measure. To compare them and thus select the best model, we refer to common metrics such as *bias*, *Root Mean Square Error (RMSE)*, *Mean Absolute Error (MAE)*, *RMSE standard deviation ratio (RSR)* and *Coefficient of determination (R2)*. The model chosen is the one with the best performance in terms of metrics.

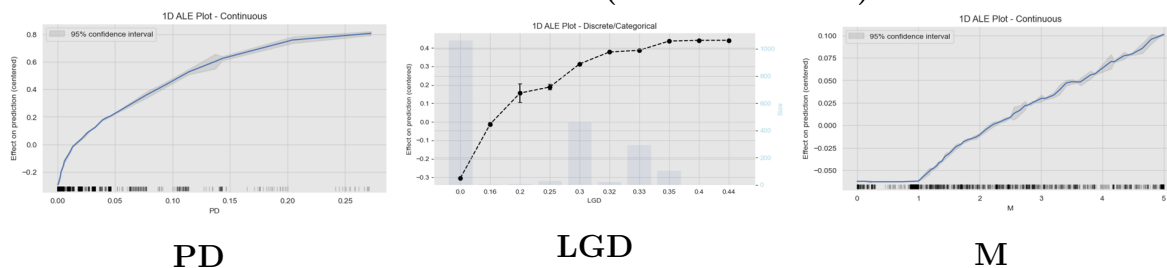
Comparison of models

	Model	Bias	RMSE	MAE	RSR	R2
<i>Conventional models</i>	Linear model	0.0002	0.1738	0.1270	0.343948	0.8817
	Polynomial model (degree 2)	0.0005	0.0845	0.0556	0.167332	0.9721
	Polynomial model (degree 3)	0.0004	0.0576	0.0342	0.114018	0.9870
	Polynomial model (degree 4)	0.0009	0.0450	0.0240	0.088882	0.9921
<i>Machine Learning</i>	Random Forest	0.0004	0.0101	0.0020	0.020000	0.9996
	XGBoost	0.0006	0.0117	0.0052	0.022361	0.9995
<i>Deep Learning</i>	Neural Network	-0.0113	0.0315	0.0155	0.062450	0.9980

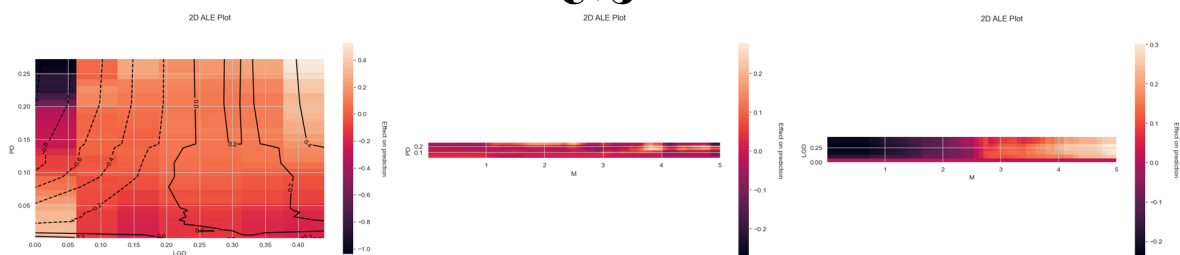
First, the linear model and its derivatives (polynomials) were excluded as they did not verify one of the main characteristics of a credit risk measure. Indeed, these models by their construction (linear according to the parameters) can lead to negative predictions of the RW. In addition, they are not able to reproduce the absorbing character of the LGD⁶. These remarks have pushed us towards machine learning models in particular tree models. The performance of the Random Forest (fit quality) makes it the best choice as an approximation model of the regulatory formula.

The estimation of marginal effects in this case by the Partial Dependence Plot (PDP) suffers from a limit, the independence of parameters. Although the correlation between the probability of default and the loss in the event of default is low, it does exist. This finding, which makes good financial sense, limits the reliability of the PDP in assessing marginal effects. Instead, we will refer to the Accumulated Local Effects (ALE) curve for assessing marginal effects.

ALE on real data (Random Forest)



6. When the LGD is null is induced a null value of RW



PD et LGD

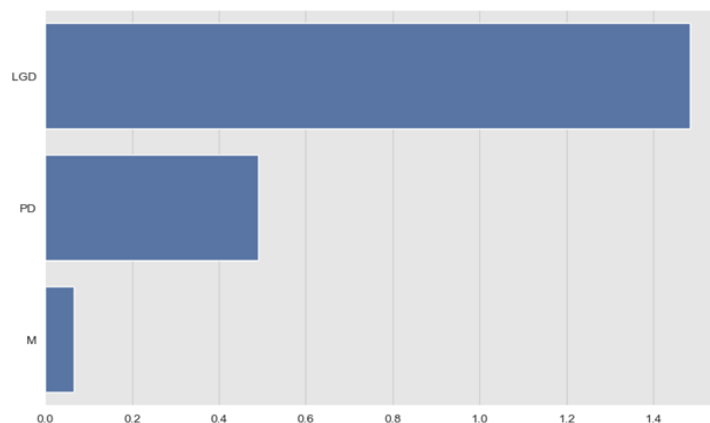
PD et M

LGD et M

The probability of default has a logarithmic effect while the loss in case of default and maturity have a more or less linear effect. Indeed, the higher the PD is, the smaller its impact on the variation of the RW is, all other parameters assumed constant.

Finally, we are interested in estimating the average effects of each parameter. The importance of the variables gives us in addition to this information, an estimate of the relative contributions of each of them. In the assessment of credit risk, there is a greater importance of loss in the event of default. This effect is three times greater than the probability of default effect. Maturity, on the other hand, has a rather marginal effect on risk assessment and respects its maturity adjustment property for taking into account long-term effects.

Permutation Feature Importance on real data (Random Forest)



Conclusion

In the end, even if the approximation of the regulatory formula by machine learning does not make it possible to obtain a new simple credit risk assessment formula, we recommend its use given the numerical results it brings to the estimation of the effect of the parameters of the regulatory function. Indeed, it appears that the loss in case of default is the most decisive parameter in the assessment of the risk from contract line. Thus, the effort to model the parameters of the regulatory function must take into account the possible negative effects that could result from a poor estimation of the loss in the event of default. Particular attention should be paid to its estimation.

TABLE DES MATIÈRES

Résumé	i
Abstract	ii
Remerciements	iii
Note de synthèse	iv
Executive summary	xi
Table des matières	xvi
Glossaire des sigles et abréviations	xx
Liste des figures	xxi
Liste des tableaux	xxiii
Introduction	1
I LE RISQUE DE CRÉDIT	3
1 L'environnement réglementaire du risque de crédit	4
1.1 La notion de risque financier	5
1.1.1 Le risque de crédit	5
1.1.2 Le risque de marché	6
1.1.3 Le risque opérationnel	7
1.2 Le cadre réglementaire	7
1.2.1 Le comité de Bâle	7
1.2.2 Evolution de la réglementation bâloise	8
1.3 Les autorités du cadre prudentiel réglementaire	12
1.3.1 Les autorités européennes	12
1.3.2 Le mécanisme de surveillance unique (SSM)	13



1.4	Le dispositif prudentiel réglementaire	14
1.5	La notion de défaut	15
2	Risque de crédit au niveau individuel	16
2.1	La probabilité de défaut	17
2.1.1	Définition et formalisme mathématique	17
2.1.2	Modélisation de la probabilité de défaut	17
2.1.3	Construction d'un modèle de PD : cas du Wholesale	18
2.2	La perte en cas de défaut (LGD)	20
2.2.1	Définition et formalisme mathématique	20
2.2.2	Modélisation de la perte en cas de défaut	20
2.3	La maturité	22
2.3.1	Définitions	22
2.3.2	Calcul de la maturité effective en IRBA	22
2.4	L'exposition en cas de défaut (EAD)	22
2.4.1	Définition et calcul de l'EAD	22
2.4.2	Méthode de calcul de l'EAD	23
2.4.3	Modélisation de l'EAD	23
2.5	Pertes anticipées	24
2.6	Pertes non anticipées	25
3	Risque de crédit au niveau du portefeuille	26
3.1	Distribution des pertes futures	27
3.2	Les risques pondérés : la formule réglementaire de Bâle 2	28
3.2.1	Construction de la formule réglementaire	29
3.3	Etude et analyse de la formule réglementaire de Bâle II	34
3.3.1	Discussion sur la formule réglementaire	34
3.3.2	Représentation graphique de la fonction réglementaire	36
II	MODÉLISATION ALTERNATIVE DU RISQUE DE CRÉDIT	38
4	Approximation de la fonction réglementaire	39
4.1	Approche analytique	40
4.1.1	Un intermède théorique : formule de Taylor	40
4.1.2	Proposition sur la formule réglementaire	41
4.1.3	Effet marginal et contribution	42
4.1.4	Application	42
4.2	Approche par modélisation classique	46
4.2.1	Régression linéaire	46
4.2.2	Régression polynomiale	49



4.2.3	Discussion des résultats et extensions	49
4.3	Approche par machine learning	50
4.3.1	Random Forest	50
4.3.2	Boosting : Gradient boosting et XGBoost	51
4.4	Approche par deep learning : Réseau de neurones	53
4.4.1	Présentation et formalisme mathématique	53
4.4.2	Application aux données simulées	54
5	Interprétabilité des modèles et applications	56
5.1	Qualité d'ajustement des modèles	57
5.1.1	Le biais	57
5.1.2	Critère RMSE (Root Mean Square Error)	57
5.1.3	Critère MAE (Mean Absolute Error)	57
5.1.4	Indicateurs normalisés	58
5.1.5	Le R^2 et le R^2 ajusté	58
5.1.6	Application aux données simulées	59
5.2	Méthodes d'interprétation des modèles de boîtes noires	59
5.2.1	Graphique de dépendance partielle (PDP)	59
5.2.2	Graphique des effets locaux accumulés (ALE)	61
5.2.3	Importance des variables	62
5.2.4	Interaction entre les variables	64
5.2.5	Modèles de substitution locaux : LIME	65
5.2.6	SHAP	66
5.3	Application à des données réelles : échantillon du portefeuille de la Société Générale	67
5.3.1	Base de travail et note sur la confidentialité	67
5.3.2	Analyse statistique succincte de la base de travail	68
5.3.3	Approche analytique	70
5.3.4	Modélisation	72
	Conclusion, limites et perspectives	78
	Références bibliographiques	III
	Annexes	IV

GLOSSAIRE DES SIGLES ET ABRÉVIATIONS

ABE	Autorité Bancaire Européenne
AIC	Akaike Information Criterion
A-IRBA	Advanced Internal Rating Based Approach
ANC	Autorités Nationales Compétentes
ALE	Accumulated Local Effects
BCE	Banque Centrale Européenne
BIC	Bayesian Information Criterion
CCF	Crédit Conversion Factor
CE	Commission Européenne
CRM	Crédit Risk Mitigation
CRR	Capital Requirement Regulation
CUE	Conseil de l'Union Européenne
CVA	Credit Valuation Ajustement
DT	Ajustements Downturn
EAD	Exposure At Default
EL	Expected Loss
F-IRBA	Foundation Internal Rating Based Approach
GE	Grandes Entreprises
IRBA	Internal Rating Based Approach
LASSO	Least Absolute Shrinkage and Selection Operator
LGD	Loss Given Default
LRA	Long-Run Average
M	Maturité
MAE	Mean Absolute Error
MCO	Moindres Carrés Ordinaires
MoC	Margin of Conservatism
MSE	Mean Squared Error
ODR	Observed Default Rates
PD	Probabilité de Défaut
PDP	Partial Dependence Plot



PE	Parlement Européen
PFI	Permutation Feature Importance
PME	Petites et Moyennes Entreprises
QIS	Quantitative Impact Studies
R^2	Coefficient de détermination
RC	Ratio de Cooke
RF	Random Forest
RMc	Ratio de McDonough
RMSE	Root Mean Square Error
RSR	RMSE standard deviation ratio
RW	Risk Weighted
RWA	Risk Weighted Assets
SSM	Single Supervisory Mechanism
SG	Société Générale
UE	Union Européenne
UL	Unexpected Loss
UTP	Unlikelihood To Pay
VaR	Value at Risk

TABLE DES FIGURES

1.1	Evolution des exigences en fonds propres	12
1.2	Institutions de régulation du système financier	14
3.1	Variation des pertes annuelles réalisées à travers le temps pour une banque donnée	27
3.2	Relation entre la VaR, l'EL et l'UL	28
3.3	Représentation de la fonction réglementaire en 2D et 3D	36
3.4	Fonction réglementaire pour les valeurs de PD sain	37
4.1	Effet marginal de la PD	43
4.2	Effet marginal de la LGD et de la maturité	43
4.3	Effet marginal de la PD^2	44
4.4	Effet marginal joint de la PD-LGD et PD-M	44
4.5	Prédiction du RW par modélisation classique	50
4.6	Random Forest sur données simulées	51
4.7	XGBoost sur données simulées	53
4.8	Architecture d'un perceptron multicouche	54
4.9	Réseau de neurones sur données simulées	55
5.1	Partial Dependance Plot sur données simulées (Random Forest)	61
5.2	ALE sur données simulées (Random Forest)	62
5.3	Permutation Feature Importance sur données simulées (Random Forest)	64
5.4	Matrice de corrélation des données réelles	69
5.5	Matrice de relation des données réelles	70
5.6	Prédiction RW par modélisation classique sur données réelles	73
5.7	Random Forest sur données réelles	74
5.8	XGBoost sur données réelles	74
5.9	Réseau de neurones sur données réelles	75
5.10	Partial Dependance Plot sur données réelles (Random Forest)	76
5.11	Permutation Feature Importance sur données réelles (Random Forest)	77
A.1	Matrice de corrélation des données simulées	XII
A.2	Matrice de relation des données simulées	XII



A.3	ALE sur données réelles (Random Forest)	XIII
A.4	Modèles de machine learning sur données simulées : Prédiction vs Données	XIII
A.5	Modèles de machine learning sur données réelles : Prédiction vs Données	XIII

LISTE DES TABLEAUX

1.1	Catégorie de rating de S&P et Moody's (KARUSISI 2007)	6
4.1	Comparaison des approximations dans l'approche analytique	45
4.2	Contribution des dérivées partielles pour une variation simultanée de 1% de tous les paramètres	45
4.3	Contribution des dérivées partielles pour une variation simultanée de 25% de tous les paramètres	45
4.4	Résultats du modèle linéaire sur données simulées	47
4.5	Résultats du modèle linéaire (fonction logarithmique) sur données simulées	48
4.6	Résultats du modèle polynomial (degré 2) sur données simulées	49
5.1	Comparaison des modèles sur données simulées	59
5.2	Contributions pour une variation simultanée de 1% de tous les paramètres sur données réelles	71
5.3	Contributions pour une variation simultanée de 25% de tous les paramètres sur données réelles	71
5.4	Comparaison des approximations dans l'approche analytique sur données réelles	72
5.5	Résultats du modèle linéaire sur données réelles	72
5.6	Résultats du modèle linéaire (fonction logarithmique) sur données réelles	73
5.7	Comparaison des modèles sur données réelles	75
A.1	Contributions pour une variation de 0,1% de tous les paramètres sur données simulées	VI
A.2	Contributions pour une variation de 10% de tous les paramètres sur données simulées	VI
A.3	Contributions pour une variation de 0,1% de tous les paramètres sur données réelles	VI
A.4	Contributions pour une variation de 10% de tous les paramètres sur données réelles	VI
A.5	Résultats du modèle polynomial (degré 3) sur données simulées	VII
A.6	Résultats du modèle polynomial (degré 4) sur données simulées	VIII
A.7	Résultats du modèle polynomial (degré 2) sur données réelles	IX
A.8	Résultats du modèle polynomial (degré 3) sur données réelles	X



A.9 Résultats du modèle polynomial (degré 4) sur données réelles XI

INTRODUCTION

L'environnement économique d'un pays est intrinsèquement lié à sa santé financière et à celle du monde qui l'entoure. Cette situation crée une interdépendance entre la vie économique d'une part, et la vie financière d'autre part. Les différentes crises économiques passées ont très souvent eu une origine financière. En effet, la globalisation a entraîné ces dernières années une interdépendance accrue entre les pays, rendant ainsi la propagation de toute défaillance très aisée. Une veille permanente des institutions de régulation est alors nécessaire pour la maîtrise des dynamiques de marché. Cette veille, qui passe par la gestion des risques, implique une évaluation pertinente et claire des risques courus par les différents acteurs du système financier.

Ainsi, l'évaluation du risque de crédit dans l'univers bancaire est une problématique centrale. Le Comité de Bâle fournit une **formule « réglementaire »** nécessaire au calcul du niveau d'exigence en fonds propres des banques pour se couvrir du risque couru. Cette formule, bien qu'ayant une construction claire, repose sur un ensemble d'hypothèses en faible concordance avec la réalité économique. Des critiques très sévères formulées à son encontre insistent sur sa crédibilité. De plus, de par sa construction, elle présente une certaine opacité sur les effets des paramètres utilisés en entrée. Ces résultats intermédiaires (effets, contributions), importants et difficilement estimables à travers sa forme actuelle, nous amènent à nous questionner sur la forme actuelle de la fonction.

De plus, les différents modèles internes utilisés par la banque pour l'estimation de la probabilité de défaut et de la perte en cas de défaut sont soumis à validation par le régulateur, et de ce fait sujets à pénalités en cas de mauvaise qualité. On parle d'ajout d'**Add-on** sur les drivers. Ce processus revient à majorer le paramètre préalablement obtenu par les modèles internes d'un pourcentage supplémentaire pour un tant soit peu rectifier la potentielle sous-estimation du risque. La première difficulté à l'utilisation de la formule réglementaire est l'anticipation sur le coût d'une éventuelle majoration de ses paramètres. La seconde difficulté survient dans la quantification de la contribution de chaque driver dans le niveau de risque obtenu. Cette dernière information permettrait, d'un point de vue bancaire, d'avoir une idée très précise des modèles nécessitant une attention plus particulière.

Ce travail vient en réponse aux difficultés énoncées précédemment. En effet, après la présentation du cadre actuel de l'évaluation du risque de crédit, nous ferons appel à plusieurs



approches pour l'estimation des effets. Ces dernières nous permettront à terme de connaître la contribution de chaque paramètre dans l'évaluation du risque de crédit.

LE RISQUE DE CRÉDIT

L'ENVIRONNEMENT RÉGLEMENTAIRE DU RISQUE DE CRÉDIT

Le contrôle prudentiel vise à garantir que les établissements de crédit disposent de fonds propres suffisants pour assumer les risques de leur activité financière. Dans ce chapitre, nous présentons l'environnement réglementaire qui régit l'activité de toutes les institutions financières en France.

1.1	La notion de risque financier	5
1.1.1	Le risque de crédit	5
1.1.1.1	Présentation	5
1.1.1.2	Mesure de risque : le rating	5
1.1.2	Le risque de marché	6
1.1.3	Le risque opérationnel	7
1.2	Le cadre réglementaire	7
1.2.1	Le comité de Bâle	7
1.2.2	Evolution de la réglementation bâloise	8
1.2.2.1	L'accord de 1988	8
1.2.2.2	De Bâle I à Bâle II	9
1.2.2.3	Bâle III : la réglementation actuelle	11
1.3	Les autorités du cadre prudentiel réglementaire	12
1.3.1	Les autorités européennes	12
1.3.1.1	La Commission européenne (CE)	12
1.3.1.2	Le Parlement européen (PE)	13
1.3.1.3	Le Conseil de l'Union Européenne (CUE)	13
1.3.1.4	L'Autorité bancaire européenne (ABE)	13
1.3.2	Le mécanisme de surveillance unique (SSM)	13
1.4	Le dispositif prudentiel réglementaire	14
1.5	La notion de défaut	15



1.1 LA NOTION DE RISQUE FINANCIER

En finance, le risque est lié à la survenance d'un évènement aléatoire impactant négativement le bilan ou le compte de résultat d'une institution financière. On peut alors distinguer l'évènement de sa cause. L'enjeu de la gestion des risques est la maîtrise de l'origine des risques financiers inhérents à toute institution financière.

Afin de mieux gérer les risques, ils ont été identifiés et catégorisés suivant leur nature. De manière générale, on distingue trois grandes catégories de risques : le risque de crédit, le risque de marché et le risque opérationnel.

1.1.1 Le risque de crédit

1.1.1.1 Présentation

Le risque de crédit est le risque de pertes associées aux changements de qualité de signature d'un débiteur sur un engagement de remboursement de dettes qu'il a contractées (DUFFIE et SINGLETON 2003). On note trois composantes au risque de crédit :

- ◆ La *dégradation de la qualité de crédit* qui fait référence à l'évolution de la perception du tiers par le marché. Si la qualité de crédit de l'emprunteur se détériore, la prime de risque s'accroît sur les marchés financiers. Le risque de dégradation résulte alors en une possible dévalorisation de la dette au cours de sa durée de vie.
- ◆ Le *risque de défaut* qui correspond à l'incapacité à faire face à ses engagements à une date donnée.
- ◆ La *perte en cas de défaut* qui résulte de l'incertitude pesant sur le taux de recouvrement une fois le défaut survenu.

La dégradation de la qualité de crédit d'un emprunteur est fortement liée au risque de défaut dans la mesure où elle est très souvent un évènement précurseur du défaut. Toutefois, les modèles utilisés arrivent très bien à modéliser de manière totalement indépendante ces deux risques.

Pour quantifier le risque de crédit sur une contrepartie, nous avons généralement recours à sa probabilité de défaut ou plus classiquement son *rating*.

1.1.1.2 Mesure de risque : le rating

D'après KARUSISI 2007, le **rating** est une note attribuée à un débiteur qui décrit la qualité de sa solvabilité ou sa capacité de remboursement. La note d'un débiteur repose sur sa situation financière et économique, et sa sensibilité aux conditions de marché. Le rating est une variable qualitative catégorielle dont les différentes modalités sont présentées sur une échelle ordinale. Ainsi, à chaque catégorie de rating est associée une capacité de remboursement et donc une qualité de crédit.



Il existe des agences de notation internationales dont la mission est d'observer les entreprises en vue de fournir une vision neutre et indépendante de leur qualité de crédit. MOODY'S INVESTORS SERVICES et STANDARDS & POOR'S sont les principaux acteurs. Les notations vont de **AAA**, ou **Aaa**, pour la meilleure qualité de crédit à **D** pour le défaut.

TABLE 1.1 – Catégorie de rating de S&P et Moody's (KARUSISI 2007)

	S&P	Moody's	Capacité de remboursement
<i>Catégories d'investissement</i>	AAA	Aaa	Très forte capacité à rembourser
	AA+	Aa1	
	AA	Aa2	Forte capacité à rembourser
	AA-	Aa3	
	A+	A1	
	A	A2	Forte capacité à rembourser mais sensibilité à la conjoncture
	A-	A3	
	BBB+	Baa1	
	BBB	Baa2	Capacité de remboursement suffisante mais sensibilité accrue à la conjoncture
BBB-	Baa3		
<i>Catégories spéculatives</i>	BB+	Ba1	
	BB	Ba2	Doit pouvoir assurer ses engagements mais fait face à des incertitudes conjoncturelles
	BB-	Ba3	
	B+	B1	
	B	B2	
	B-	B3	
	CCC+		
	CCC	Caa	Ne pourra honorer ses engagements que si la conjoncture est favorable
	CCC-		
	CC	Ca	Créance de second rang lorsque les créances de premier rang sont CCC
C	C		
D	D	Défaut de paiement constaté ou prévisible	

1.1.2 Le risque de marché

D'après CROUHY et al. 2001, le **risque de marché** fait référence au risque de pertes sur les positions du bilan d'une institution financière à la suite de variations des prix de marché. Ainsi, il recouvre :

- ◆ le risque sur titres (*equity risk*) qui se réfère à la sensibilité de la valeur d'un portefeuille de titres ;
- ◆ le risque de taux d'intérêt (*interest rate risk*) relatif aux fluctuations de valeur des instruments de taux ;
- ◆ le risque de change (*currency risk*) relatif aux fluctuations des taux de change ;
- ◆ le risque sur produits de base (*commodity risk*) encourus pour l'ensemble de l'activité de l'entreprise.



Pour mesurer ce risque, le modèle de sélection de portefeuille de MARKOWITZ 1959 a été longtemps la méthode utilisée. Dans ce modèle, l'agent maximisait son espérance de gain pour un niveau de risque donné, qui est matérialisé par la volatilité (écart-type) des rendements de portefeuille. Cette vision basée sur l'hypothèse de normalité des rendements a laissé place à une mesure de risque plus statistique. Il s'agit de la **Value at Risk (VaR)**. La VaR fait référence à un quantile de la distribution des pertes potentielles pour un horizon donné. Cette mesure de risque est équivalente à l'écart-type dans le cas où la distribution des pertes serait gaussienne.

1.1.3 Le risque opérationnel

Le **risque opérationnel** est défini par le Comité de Bâle comme étant les risques de pertes directes ou indirectes liées à une inadéquation ou une défaillance des procédures et systèmes internes, des erreurs humaines, ou à des événements extérieurs (BANKING SUPERVISION 2001). Le risque technologique et les défaillances du système d'information sont, de ce fait, considérés comme des risques opérationnels. De par sa nature, la mesure du risque opérationnel n'est pas une tâche aisée. Le Comité de Bâle a donc proposé un ensemble de recommandations afin de prévenir au mieux ce risque.

1.2 LE CADRE RÉGLEMENTAIRE

Il existe un lien très fort entre santé économique et santé financière d'un point de vue national et international. La réglementation bancaire a donc pour mission d'établir des règles et normes régissant l'activité des institutions financières. Ces normes prudentielles comportent des exigences relatives à une dotation minimale en capital pour couvrir les risques encourus par une institution financière.

À la suite du crash boursier de 1929 et de la crise qui s'en suivit, des actions ont été menées pour régir l'activité des banques. Dans un premier temps, la réglementation s'élargissait uniquement à l'échelle nationale par un certain nombre de lois qui régissent le secteur bancaire. Les objectifs étaient : la protection des épargnants contre le risque de faillite, la sécurité des systèmes de paiement, et la prévention du risque systémique qui représente le risque qu'une défaillance bancaire se propage aux autres institutions financières. Ensuite, les états entretenant des relations entre eux, une institution internationale a vu le jour dans le but de créer un cadre général.

1.2.1 Le comité de Bâle

Les faillites des institutions financières des années 90 sont, pour la plupart, liées à des prises de risque excessives sur les marchés dérivés. Pour prévenir les crises financières et leur impact sur la solvabilité des établissements financiers, il s'est avéré indispensable de mettre en place des mécanismes de surveillance internationaux par des autorités de tutelle. Cet organe, au delà



de coordonner au niveau international le système financier, a pour principale tâche d'éviter la survenance de futures crises financières globales en limitant au mieux le risque systémique.

Le *comité de Bâle* a été institué en 1974 par les gouverneurs des banques centrales du G-10¹. A cette date, la principale tâche du superviseur était de s'assurer que les institutions financières possédaient un niveau de capital et de réserves suffisants pour supporter les risques pris dans le cadre de leurs activités.

1.2.2 Evolution de la réglementation bâloise

Les différentes crises financières qui ont secoué la vie économique des états au fil des années, ont poussé la réglementation à se réinventer pour relever les défis de son temps.

1.2.2.1 L'accord de 1988

En juillet 1988, le premier accord de Bâle, dit **Bâle I**, a été édicté. Il venait en réaction à la dégradation des fonds propres des principales banques internationales. Basé sur une approche simple et normative, l'accord de 1988 établissait un ratio de solvabilité unique que chacune des banques membres devait respecter. Il s'agit du **ratio de Cooke**. Cet accord présentait les principales caractéristiques suivantes (CHARLES 2015) :

- ✦ il était adapté au profil des grandes banques du G10 ;
- ✦ les exigences en fonds propres couvraient uniquement le risque de crédit ;
- ✦ l'échelle de pondération comportait uniquement les quatre valeurs (0%, 20%, 50%, 100%) ;
- ✦ les assujettis² devaient présenter un ratio de Cooke supérieur à 8%.

Ces mesures devaient permettre une meilleure adéquation des fonds propres par rapport aux risques et ainsi, renforcer la solidité et la stabilité du système bancaire. Au-delà de ces aspects, cet accord permettait aussi d'atténuer les inégalités concurrentielles entre les différentes banques.

Le ratio de Cooke (RC) est le rapport entre le montant des fonds propres et celui des actifs pondérés du risque.

$$RC = \frac{\text{Fonds Propres}}{\text{Actifs pondérés du risque}}$$

Les Actifs Pondérés du risque représentent la valeur des actifs risqués affectés d'un coefficient de pondération qui dépend de la nature de l'actif. Les coefficients de pondération étaient établis par le Comité de Bâle et reflétaient le niveau de risque associé à chaque actif.

1. **Groupe des dix** (Allemagne, Belgique, Canada, Etats-Unis, France, Italie, Japon, Pays-Bas, Royaume-Uni, Suede, Suisse) créé en 1962 et fondateur du Comité de Bâle.

2. Ensemble des institutions soumises à l'application des textes réglementaires.



Au début des années 2000, un premier bilan de Bâle I révèle que cet accord a porté ses fruits. Les principales améliorations étaient :

- ◆ un accroissement significatif des ratios de fonds propres ;
- ◆ une application par un large panel de pays ;
- ◆ une large réduction des inégalités de la concurrence entre les banques.

Toutefois, certaines faiblesses et critiques ont également été émises par ce bilan. Tout d'abord, le choix du niveau de 8% et des coefficients de pondération appliqués aux actifs semblait n'avoir que très peu de fondement économique. Ensuite, on pouvait noter une absence de prise en compte du profil de risque des établissements financiers ainsi que des techniques de réduction ou de transfert de risque telles que la prise de garanties ou la diversification des portefeuilles. Enfin, l'approche se limitait au risque de crédit, alors que les institutions financières étaient confrontées à une plus grande variété de risques (contrepartie, opérationnel, taux, etc.).

C'est face à ces faiblesses qu'un premier amendement en 1996 intégra un principe d'immobilisation de capital pour couvrir les pertes liées aux opérations sur les marchés financiers.

1.2.2.2 De Bâle I à Bâle II

Le 26 juin 2004, un nouvel accord voit le jour. Il instaure un dispositif prudentiel destiné à mieux appréhender et limiter les risques des établissements de crédit. De ce fait, il modifie l'assiette des risques pour intégrer le risque opérationnel et le risque de marché.

Ce nouvel accord, dit **Bâle II**, a pour but de permettre aux institutions financières, sous réserve d'autorisation par le superviseur, d'utiliser des méthodes plus avancées dans la mesure du risque de crédit et du risque opérationnel. Cette nouvelle approche interne de mesure des risques est l'un des principaux enjeux de ce nouvel accord en réponse aux critiques de celui de 1988. Ainsi, deux options sont proposées :

- ◆ une approche standard qui peut être vue comme une amélioration de Bâle I ;
- ◆ une approche par modèles internes afin d'affiner la mesure de risque aux caractéristiques de la Banque. Cette approche prend le nom d'*Internal Rating Based Approach (IRBA)*. Il s'agit d'un ensemble de techniques de mesure de risque de crédit proposé dans les règles d'adéquation des fonds propres et fondé sur les notations internes. Selon cette approche, les banques ont la possibilité d'élaborer leur propre modèle empirique afin de quantifier les fonds propres nécessaires pour faire face au risque de crédit. Il est impératif d'obtenir l'accord du régulateur pour pouvoir recourir à cette approche. Selon les inputs que les banques ont le droit d'estimer par modèle interne, on note deux différentes approches.
 - *Foundation IRBA (F-IRBA)* : seule la probabilité de défaut des tiers est estimée à l'aide des modèles internes. Les autres paramètres, comme le taux de perte en cas de défaut, sont toujours fournis par le régulateur.



- *Advanced IRBA (A-IRBA)* : avec cette approche, en plus de la probabilité de défaut, les banques peuvent avoir recours à leurs modèles internes pour modéliser le montant de l'exposition au défaut, le taux de perte en cas de défaut et la maturité effective.

Toutefois, même en approche par modèle interne, la formule de calcul du risque reste fournie par le comité de Bâle. Seuls les paramètres de cette formule peuvent être obtenus par modèle interne.

Pour améliorer la supervision des banques par le régulateur, trois piliers ont été définis :

Pilier 1 : Exigences minimales en fonds propres

Afin de couvrir l'ensemble des risques encourus par la banque, un niveau minimal en fonds propres est requis. On utilise dorénavant le **ratio de Mc Donough** (RMc).

$$RMc = \frac{\text{Fonds Propres}}{\text{Mesure Risque Crédit} + \text{Mesure Risque marché} + \text{Mesure Risque Op.}} \geq 8\%$$

Le seuil minimum de 8% sera élevé à 10,5% en 2019 pour la prise en compte d'un coussin de conservatisme dans le cadre de Bâle III. Aussi, le montant d'exigence en fonds propres est subdivisé en trois parties, appelées *tier* :

- Tier ❶ : constitué du *Core-Tier 1* (capital éligible³ et les réserves éligibles⁴) et de l'*Hybrid Tier 1* (produits mixtes⁵ ne contenant pas d'obligation de remboursement du capital).
- Tier ❷ : composé de l'*Upper Tier 2* (emprunts subordonnés perpétuels⁶ et fonds internes de sécurité) et du *Lower Tier 2* (emprunts subordonnés à long terme⁷)
- Tier ❸ : composé d'emprunts subordonnés à court terme (remboursement attendu à partir de l'échéance de deux ans) et des bénéficiaires intérimaires du trading book. Ce tier n'est plus pris en compte dans Bâle III.

Pilier 2 : Processus de surveillance prudentielle

Un processus de supervision est mis en place par les instances nationales de régulation pour s'assurer que chaque banque applique de manière rigoureuse les accords. Ce pilier vise à :

- renforcer la supervision prudentielle ;
- permettre au superviseur d'intervenir plus directement à travers des mesures correctrices en cas de besoin ;
- introduire la notion de capital économique (montant maximal de pertes admissibles par la banque pour rester bien notée) et l'évaluation de risques spécifiques comme le risque de taux, de liquidité et de concentration.

3. Capital versé par les actionnaires et les primes d'émission

4. Bénéficiaires non distribués ou non réinvestis

5. Mi-actions, mi-obligations, ...

6. Emprunts sans aucune obligation de remboursement du capital

7. Emprunts ne pouvant être remboursés avant l'échéance de 5 ans que sous certaines conditions



Pilier 3 : Discipline de marché

Ce pilier définit les exigences en matière de communication publique sur la structure du capital, l'allocation des fonds propres et l'exposition au risque.

1.2.2.3 Bâle III : la réglementation actuelle

Malgré les avancées de l'accord de Bâle II, le système financier mondial a été victime d'une crise majeure qui a mis en lumière ses limites. L'accord de **Bâle III** vise donc par la prise en compte des risques systémiques⁸ à renforcer la résilience du secteur bancaire. En l'occurrence, de nombreux intervenants pensent que la gestion du risque de contrepartie sur les opérations de marché s'est révélée être insuffisamment orientée vers un point de vue macro-prudentiel⁹, sans prendre assez en compte ces effets systémiques (SENHAJ 2019).

Principalement, Bâle III apporte les innovations suivantes :

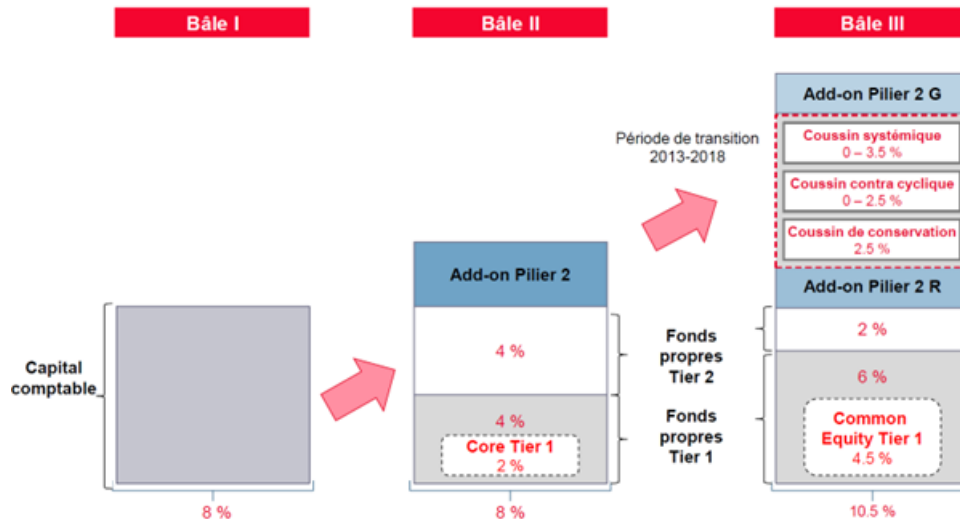
- ◆ L'amélioration de la qualité des fonds propres : de nouvelles provisions sont introduites pour augmenter le montant des fonds propres.
- ◆ La pondération des corrélations entre valeurs d'actifs (risque systémique) : au niveau de la formule de calcul du RWA (Risk Weighted Assets) sous l'approche avancée, un coefficient de **1,25** est appliqué aux grandes institutions financières pour prendre en considération le risque inhérent aux expositions à d'autres entités financières.
- ◆ La prise en compte du « *Wrong Way Risk* » : il s'agit du risque qu'une exposition à une contrepartie soit inversement corrélée à la qualité de crédit de celle-ci. Sous Bâle III, le risque de corrélation défavorable est une dimension du risque systémique qui est gérée par les établissements financiers à travers une identification des différents deals concernés par les stress tests. Une gestion fine de ces deals doit alors être réalisée par l'institution financière.
- ◆ Le calcul de la CVA (*Credit Valuation Adjustment*) pour la mesure du risque de contrepartie : il s'agit de la différence entre la valeur du portefeuille sans risque et la valeur du portefeuille risqué. La CVA correspond ainsi à la valeur de marché du risque de crédit de la contrepartie sur les opérations de marché d'un portefeuille de trading. Bâle III introduit donc des exigences supplémentaires en capitaux pour se prémunir des risques de perte sur la CVA qui surviennent généralement à la suite de la dégradation de la qualité de crédit d'une contrepartie.
- ◆ L'introduction de ratios de liquidité et de levier.

8. Le risque systémique est le risque qu'un événement particulier entraîne par réactions en chaîne des effets négatifs considérables sur l'ensemble du système pouvant occasionner une crise générale de son fonctionnement (Finance pour tous, 2019)

9. La surveillance macro-prudentielle consiste à mesurer, évaluer et limiter le risque systémique, ie. le risque d'une défaillance majeure de la fourniture de services financiers ayant des conséquences sérieuses sur l'économie réelle (Financial Stability Board – FSB, 2011)



FIGURE 1.1 – Evolution des exigences en fonds propres



Source : Documentation Société Générale

1.3 LES AUTORITÉS DU CADRE PRUDENTIEL RÉGLEMENTAIRE

On distingue deux catégories d'autorités dans le processus de veille prudentielle : les entités chargées d'édicter les directives et dispositions réglementaires ; et les organismes chargés de la supervision et de l'application du contrôle prudentiel.

Dans la zone euro, on retrouve principalement le Parlement européen, le Conseil européen, la Commission Européenne et la Banque Centrale Européenne.

1.3.1 Les autorités européennes

1.3.1.1 La Commission européenne (CE)

Il s'agit de la branche exécutive de l'Union Européenne (UE) chargée de la rédaction des propositions (directives, règlements et décisions) de nouvelles législations européennes. Elle joue donc un rôle crucial dans les propositions législatives et les études d'impact avant leur mise en application dans l'UE. Ainsi, elle :

- ✦ propose des lois au Parlement européen et au Conseil de l'UE ;
- ✦ est responsable de la planification, la préparation et la proposition de nouvelles lois européennes ;
- ✦ aide les pays de l'UE à mettre en œuvre la législation ;
- ✦ s'assure de l'application dans les délais et de manière correcte des lois ;
- ✦ propose et met en œuvre le budget européen en gérant les programmes de financements de l'UE.



1.3.1.2 Le Parlement européen (PE)

Le Parlement européen examine la proposition de la CE et décide de la modifier ou de l'approuver, auquel cas, elle est transmise au Conseil de l'UE pour être adoptée. Il s'agit d'un lieu de débat politique où les décisions au niveau de l'UE sont prises. En coopération avec les parlements nationaux des pays, le PE détermine la législation européenne.

1.3.1.3 Le Conseil de l'Union Européenne (CUE)

Les États membres de l'UE sont représentés par le Conseil de l'Union Européenne. Le CUE est le lieu de rencontre entre les ministres des gouvernements de chaque pays, le président de la CE et le président du CUE pour discuter, modifier et adopter les lois et coordonner les politiques.

Le CUE peut adopter la législation s'il est en accord avec la position du PE ou la modifier et la renvoyer au PE en seconde lecture.

1.3.1.4 L'Autorité bancaire européenne (ABE)

L'autorité bancaire européenne est une autorité indépendante de l'UE dont le but est de garantir un niveau efficace et cohérent de réglementation et de surveillance prudentielles dans l'ensemble du secteur bancaire européen. Son rôle est le maintien de la stabilité financière dans l'UE. L'ABE préserve l'intégrité, l'efficacité et le bon fonctionnement du secteur bancaire.

La mission principale de l'ABE est de contribuer à la création d'un règlement européen uniforme dans le secteur bancaire. Cela revient à fournir un ensemble de règles prudentielles harmonisées à tous les établissements financiers de l'UE. En plus de jouer le rôle de promoteur de la convergence des pratiques de surveillance au sein de la zone, elle a pour mission d'évaluer les risques et vulnérabilités du secteur bancaire européen.

1.3.2 Le mécanisme de surveillance unique (SSM)

Le Single Supervisory Mechanism (SSM) est le système de supervision bancaire européen. Il est composé de la Banque Centrale Européenne (BCE) et des autorités nationales compétentes (ANC) des pays participants. Il a pour objectifs de :

- assurer la sauvegarde et la solidité du système bancaire européen ;
- renforcer l'intégration et la stabilité financière ;
- garantir la cohérence de la supervision bancaire.

Avec le mécanisme de résolution unique, la supervision bancaire européenne est l'un des deux piliers de l'union bancaire européenne. Cela implique de :

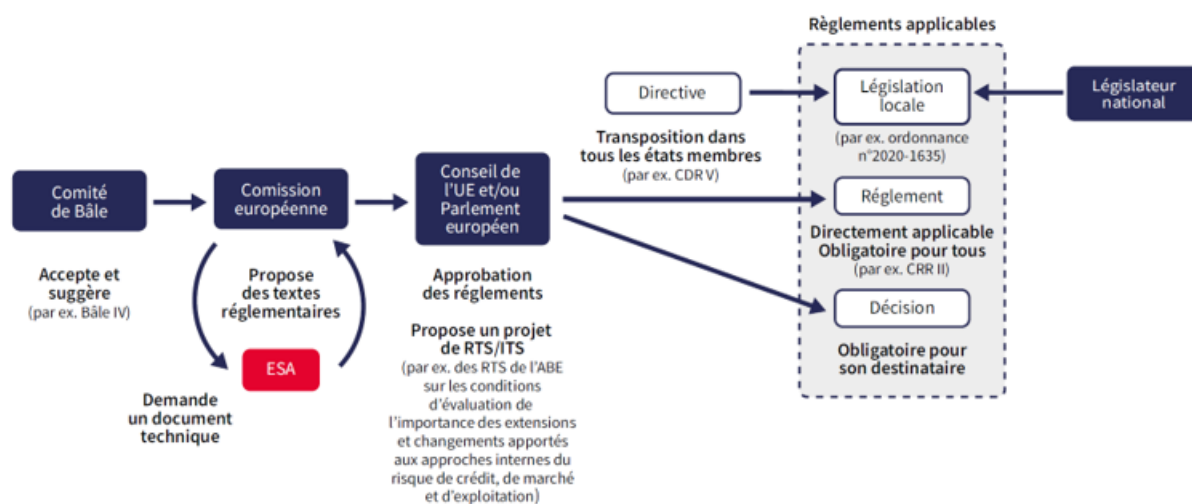
- veiller au respect des règles prudentielles de l'UE ;
- mener des revues de supervision, des inspections sur place et des enquêtes ;



- établir des exigences de fonds propres plus élevées afin de parer à tout risque financier ;
- accorder et retirer des agréments bancaires.

Dans le cadre du mécanisme de surveillance unique dans la zone euro et dans les pays hors zone UE qui ont choisi de rejoindre le SSM, la BCE est le superviseur prudentiel central des établissements financiers. Elle se charge de superviser directement les banques les plus grandes qui présentent un risque systémique conséquent. Les autres banques, de taille plus réduite, sont placées sous la supervision des superviseurs nationaux.

FIGURE 1.2 – Institutions de régulation du système financier



Source : Documentation Société Générale

1.4 LE DISPOSITIF PRUDENTIEL RÉGLEMENTAIRE

Conformément à l'article 148 du CRR, le Groupe SG doit mettre en oeuvre l'approche IRB (Internal Rating Based) pour toutes les expositions, sauf s'il a reçu l'autorisation de la BCE d'utiliser de manière permanente l'approche standard sur certaines catégories ou certains types d'expositions.

Le Groupe SG se doit donc de formaliser une stratégie relative à ses expositions qui consiste :

- soit à conserver l'approche standard conformément à l'article 150 de la CRR ;
- soit à passer à l'approche IRB conformément à l'article 143.

Le choix retenu par la SG doit être justifié et approuvé par la BCE. Ainsi, pour se conformer à la réglementation et s'assurer que les approches mises en oeuvre sont soumises à une surveillance régulière, le groupe SG a :

- élaboré un registre de la méthode utilisée par entité légale et catégorie d'actifs ;
- amélioré et clarifié la gouvernance des méthodologies ;
- mis en place un reporting et une surveillance réguliers par Business Unit, par entité et par portefeuille Bâlois.



1.5 LA NOTION DE DÉFAUT

La notion de défaut est présentée de manière différente suivant les référentiels retenus. Il s'avère donc nécessaire de réconcilier les différentes définitions pour une compréhension homogène du défaut.

Concrètement, on retiendra qu'un débiteur est en défaut quand l'un et/ou l'autre des critères suivants est rempli :

- ☛ **critère d'arriérés de paiement** : les arriérés de paiement vis-à-vis du Groupe et relatifs à des obligations de crédit ont dépassé le seuil de matérialité pendant plus de 90 jours en continu.
- ☛ **critère d'absence probable de paiement** (UTP ou *Unlikelihood to pay*) : on estime peu probable que le débiteur soit en mesure de s'acquitter de la totalité de ses obligations de crédit vis-à-vis du Groupe SG sans que soit mis en œuvre, le cas échéant, l'exercice d'un collatéral ou la réalisation d'une garantie, et ce même en l'absence d'éventuels arriérés de paiement.

En résumé

Le cadre réglementaire bancaire a fortement évolué durant ces 50 dernières années, passant d'un système micro-prudentiel à un système macro-prudentiel. Ces évolutions venues en réponse à un monde changeant ont permis la mise en place d'une structure clairement définie édictant le mode de fonctionnement du système bancaire. Cette organisation définit non seulement une hiérarchisation mais aussi pose des jalons censés prémunir les banques et leurs usagers de crises futures. Ainsi, cette organisation définit les règles d'évaluation de risque. Ces dernières feront l'objet du prochain chapitre.

RISQUE DE CRÉDIT AU NIVEAU INDIVIDUEL

Ce chapitre présente les différents éléments d'évaluation du risque de crédit au niveau individuel. Il s'attarde sur un formalisme mathématique, mais aussi pratique de la construction de ces indicateurs.

2.1	La probabilité de défaut	17
2.1.1	Définition et formalisme mathématique	17
2.1.2	Modélisation de la probabilité de défaut	17
2.1.3	Construction d'un modèle de PD : cas du Wholesale	18
2.1.3.1	Différentiation du risque via un modèle de notation	19
2.1.3.2	Quantification de la probabilité de défaut	19
2.2	La perte en cas de défaut (LGD)	20
2.2.1	Définition et formalisme mathématique	20
2.2.2	Modélisation de la perte en cas de défaut	20
2.2.2.1	Actualisation et projection des flux	20
2.2.2.2	Crédit Risk Mitigation (CRM)	21
2.2.2.3	Ajustements Downturn (DT)	21
2.2.2.4	Formule de calcul finale	21
2.3	La maturité	22
2.3.1	Définitions	22
2.3.2	Calcul de la maturité effective en IRBA	22
2.4	L'exposition en cas de défaut (EAD)	22
2.4.1	Définition et calcul de l'EAD	22
2.4.2	Méthode de calcul de l'EAD	23
2.4.3	Modélisation de l'EAD	23
2.4.3.1	Différenciation du risque	23
2.4.3.2	Quantification du risque	24
2.5	Pertes anticipées	24
2.6	Pertes non anticipées	25



Le risque de crédit est défini comme le risque de perte que la banque pourrait enregistrer lorsqu'un emprunteur se retrouve dans l'incapacité de rembourser tout ou partie de son emprunt à une échéance donnée. Selon WHANNOU 2016, cette définition suscite les questionnements suivants :

- Quelles sont les conditions qui définissent l'incapacité de remboursement ou le défaut de paiement ?
- Comment quantifier une perte potentielle, c'est-à-dire non encore réalisée ?
- Existe-t-il des caractéristiques de l'emprunteur ou du prêt qui influencent la quantification de ce risque ?

La modélisation du risque de crédit fait l'objet d'avancées majeures depuis plus de 30 ans au sein du système bancaire. En effet, étant lié à l'incertitude de pertes sur les crédits à un horizon choisi, l'objet des modèles de risque de crédit est la modélisation de la distribution des pertes. On distingue de ce fait **les modèles de défaut** des **modèles marked-to-market**. D'une part, les modèles de défaut considèrent uniquement les pertes induites par le défaut de crédit ; d'autre part, les modèles marked-to-market prennent en compte les pertes dues au changement de la qualité de crédit.

Il y a trois notions importantes sur lesquelles reposent les modèles de défaut : la *probabilité de défaut* ou *Probability of Default (PD)*, la *perte en cas de défaut* ou *Loss Given Default (LGD)* et l'*exposition au moment du défaut* ou *Exposure At Default (EAD)*.

2.1 LA PROBABILITÉ DE DÉFAUT

2.1.1 Définition et formalisme mathématique

La probabilité de défaut (**PD**) qui matérialise le *risque client* est la probabilité que le client fasse défaut à un horizon fini. Ce dernier est généralement fixé à un an. En d'autres termes, la PD est la probabilité qu'à horizon un an, un crédit migre vers la classe du défaut.

Pour chaque crédit $i = 1, \dots, m$, notons PD_i la probabilité de défaut associée. On définit la variable aléatoire de Bernoulli L_i qui matérialise le défaut ou non. Ainsi,

$$PD_i = \mathbb{P}(L_i = 1)$$

2.1.2 Modélisation de la probabilité de défaut

Suivant l'approche retenue (standard ou IRBA), la probabilité est soit déterminée, soit estimée. En effet, suivant l'approche standard, le régulateur fournit pour chaque catégorie de client et de transaction, la PD prudente associée à appliquer. Les banques peuvent définir leurs propres estimations de la PD uniquement en approche IRBA.



En méthode IRBA, plusieurs approches sont utilisées aujourd'hui pour calculer ces probabilités :

- ✦ l'approche par les **ratings** : les probabilités sont alors associées à des notations attribuées par des agences externes ou des services internes. On fait alors recours aux matrices de transition pour déterminer la qualité de crédit d'une contrepartie ;
- ✦ les approches **classiques** : ces approches font recours aux données disponibles ainsi qu'à la théorie financière et statistique pour calculer des probabilités finies.

Pour les sociétés cotées, les probabilités de défaut sont estimées en utilisant soit le « *modèle structurel*¹ » du risque de crédit de **Robert Merton** (1974), soit des « *modèles à forme réduite*² » tels que le modèle de **Jarrow-Turnbull**.

En pratique, les modèles de PD sont constitués de deux composantes : la Ranking method et les segments de calibrage. La **Ranking method** est la méthode utilisée pour classer les expositions au regard du risque de défaut. Il peut y avoir plusieurs Ranking method au sein d'un modèle de PD, à condition qu'elles renvoient vers une échelle de grade unique. Utilisée pour la différenciation du risque, la Ranking method sert à produire un score ou un rating permettant un classement ordinal des obligors au regard du risque de défaut. On retrouve au sein de la banque, deux grandes catégories de modèles :

- les modèles **Retail** : c'est la banque de détail sur les personnes physiques ou morales de très petite taille
- les modèles **Wholesales** (Non retail) : ce sont les opérations sur personnes morales de plus ou moins grande taille

2.1.3 Construction d'un modèle de PD : cas du Wholesale

Le portefeuille Wholesale est essentiellement constitué des PME (petites et moyennes entreprises), des GE (grandes entreprises), des institutions financières, des souverains, des collectivités locales et des associations. Le modèle de PD doit donc permettre la différenciation du risque afin d'attribuer la note la plus juste à chaque contrepartie, puis de quantifier ce risque à travers une valeur de probabilité de défaut associée à chaque note.

1. L'approche structurelle stipule que le défaut d'une entreprise est la résultante d'un processus qui mène une firme en difficulté à la cessation de paiement (MORAUX et RENAULT 2002). Apparu dans les années 70, ce cadre permet d'identifier tout titre de financement (créance) comme un contrat dérivé sur la valeur des actifs de la firme.

2. Les modèles à forme réduite ne fournissent aucune explication financière à la cessation de paiement d'une firme. Pour ces modèles, le défaut survient uniquement par *hasard*. De ce fait, la date d'arrivée au défaut peut être modélisée comme le premier temps de saut d'un processus stochastique.



2.1.3.1 Différentiation du risque via un modèle de notation

La différenciation du risque passe par la construction d'un modèle de notation. La modélisation se base sur un historique représentatif le plus long possible, en général un seuil minimum de trois ans est considéré. L'objectif étant d'opérer une classification des contreparties en termes de risque de crédit. Deux principales approches peuvent être envisagées : **modèle en mapping et modèle en défaut**.

L'approche de modélisation en mapping est une approche multimodale qui a pour but de prédire une note représentative du risque. Cette approche permet d'obtenir directement la note de la contrepartie.

L'approche de modélisation en défaut est une approche bimodale dont le but est la prédiction du défaut à un an. On considère ces modèles comme des modèles de scoring car générant un score qui est une mesure continue du risque. Une étape supplémentaire de discrétisation est alors nécessaire pour obtenir les notes.

Le choix de l'approche dépend des caractéristiques du périmètre modélisé. En effet, plus les taux de défaut observés seront faibles, plus on se tournera vers des modèles de mapping. Et dans l'autre sens, plus les taux de défaut seront élevés, plus on aura recours au modèle de défaut.

2.1.3.2 Quantification de la probabilité de défaut

La quantification de la probabilité de défaut selon l'approche bâloise est réalisée suivant une **approche différenciée** afin de mieux tenir compte du risque. Cette approche consiste à regrouper au sein d'un même segment de calibrage, les contreparties pour lesquelles la note a un même sens. On construit alors une échelle de PD qui représente une évaluation quantitative de la probabilité de défaut au sein de chaque segment de calibrage. L'échelle de PD se doit d'être strictement monotone en termes de risque. La PD dans le cadre wholesale est estimée en suivant les étapes ci-dessous :

- ❖ calculer les taux de défaut annuels observés (ODR - *Observed Default Rates*) à pas trimestriel par rating.
- ❖ calculer la Long-Run Average (LRA) par grade ; c'est la moyenne des ODR.
- ❖ vérifier que l'échelle de LRA est strictement monotone croissante.
- ❖ ajouter des marges de conservatismes liées à la qualité des données, à la volatilité des taux de défaut, etc.
- ❖ vérifier que l'homogénéité intra-grades et l'hétérogénéité inter-grades sont respectées sur l'échelle obtenue.
- ❖ s'assurer qu'il n'y a pas une forte concentration au sein de certains ratings et que les taux de défaut observés sont stables dans le temps.
- ❖ respecter le floor réglementaire de 0,03% qui impose qu'en dehors des contreparties souveraines, les PD estimées sur chaque grade soient supérieures ou égales au floor.



2.2 LA PERTE EN CAS DE DÉFAUT (LGD)

2.2.1 Définition et formalisme mathématique

La perte en cas de défaut, ou taux de perte, est la fraction de l'exposition qui sera perdue dans l'éventualité où le défaut se réalise. Sa mesure est assez complexe, car elle est assez spécifique à chaque institution financière. En effet, le taux de perte dépend de la capacité de l'institution à prendre des garanties et sûretés ainsi que de l'efficacité de son service de recouvrement.

Considérons la variable aléatoire $\tilde{\tau}$ qui mesure le taux de recouvrement. De manière classique, la perte en cas de défaut pour chaque crédit i , noté LGD_i , est obtenue de la manière suivante :

$$LGD_i = 1 - \mathbb{E}(\tilde{\tau}_i), \quad \tilde{\tau}_i \in [0, 1]$$

2.2.2 Modélisation de la perte en cas de défaut

Il existe deux familles de modèles de LGD :

- modèle de LGD des actifs sains : le calibrage du modèle se fait sur la base des défauts au moment de leur défaut et l'application sur la base des encours sains ;
- modèle de LGD des actifs en défaut : le calibrage se fait sur la base des défauts durant la vie du défaut et s'applique à la base des défauts.

L'estimation de la perte future par ces modèles correspond à une moyenne des pertes observées sur une population équivalente. De par cette définition, la LGD peut être mesurée au moyen de la formule suivante :

$$LGD = \frac{Pertes}{Exposition}$$

Toutefois, on arrive plus aisément à connaître l'exhaustivité des recouvrements (principe de prudence), on utilise donc à la place cette formule de calcul :

$$LGD = 1 - \frac{Recouvrements}{Exposition}$$

2.2.2.1 Actualisation et projection des flux

En cas de défaut, les recouvrements ne sont pas immédiats. Le manque à gagner par la banque durant cette période de recouvrement est ajusté à travers une actualisation des recouvrements avec le taux Euribor, plus 5% annuel. De plus, il arrive que le recouvrement ne soit pas complet pour toutes les facilités rattachées à un tiers. Un ajustement par projection des recouvrements à venir est réalisé.



2.2.2.2 Crédit Risk Mitigation (CRM)

Un Crédit Risk Mitigation (CRM) est un atténuateur de risque de perte. En l'occurrence, la banque peut avoir recours à des collatéraux (biens matériels ou non) ou à des garants (personnes physiques ou morales) pour atténuer le risque de non-récupération de l'exposition en cas de défaut. Pour que ces atténuateurs de risques soient recevables, ils doivent respecter certains critères.

Un garant ne peut être considéré que s'il possède une qualité de crédit meilleure que la contrepartie. Dans ce cas, la banque prend en compte l'atténuation via un mécanisme de substitution des paramètres bâlois (PD, LGD) de l'emprunteur par ceux du garant.

De manière équivalente, un collatéral qui sera considéré comme éligible viendra en atténuation du risque à travers une substitution du taux de perte (LGD). On applique en plus des modèles de haircut pour tenir compte de la décote futur de la valeur du bien.

2.2.2.3 Ajustements Downturn (DT)

En accord avec le principe de prudence de la réglementation bâloise, en plus de la mesure des LGD observées, on ajoute des marges de prudence (**MoC** : Margin of Conservatism) et on réalise un ajustement destiné à faire face à une éventuelle dégradation économique (Ajustement Downturn).

Les marges (MoC) sont au nombre de trois (comme pour la PD) :

- **MoC A** : pour la qualité des données ;
- **MoC B** : pour les évolutions de processus métier ;
- **MoC C** : pour l'incertitude sur les données (marge de volatilité).

Pour l'ajustement downturn, on tient compte de la dégradation de LGD au cours des crises dans le passé sur une période de 20 ans.

2.2.2.4 Formule de calcul finale

La formule de calcul finale de la LGD tient compte de l'ensemble des ajustements supplémentaires opérés sur la LGD observée. En fin de compte, la formule retenue est la suivante :

$$LGD = LGD_{observe} + Adj_{projection} + Addon_de_cots + MoC_A + MoC_B + MoC_C + Adj_{DT}$$



2.3 LA MATURITÉ

2.3.1 Définitions

Suivant les besoins, trois visions de la maturité peuvent être adoptées.

La première vision se place dans une vision statique du monde et définit la **maturité initiale** d'une exposition comme la durée initiale de l'exposition. Elle est fixe durant toute la vie du contrat et est déterminée à l'ouverture du contrat.

La deuxième vision se réfère plutôt à une vue actuelle du monde. Ainsi, à une date d'observation donnée, on considère la durée restant à courir jusqu'à la fin du contrat. C'est la **maturité résiduelle**.

Enfin, la dernière vision du monde tient compte des montants des échéances. C'est la **maturité effective**. Elle correspond à la moyenne des échéances de remboursement pondérées par les montants des flux à chaque échéance. Il s'agit de la valeur de maturité retenue dans le calcul du RWA.

2.3.2 Calcul de la maturité effective en IRBA

En approche *foundation*, la maturité effective est définie de manière forfaitaire à **2,5** ans ou à **0,5** an pour les opérations de pension et de prêt-emprunt de titres.

En approche *advanced*, la maturité effective est minorée à 1 an et capée à 5 ans.

$$M = \max \left\{ 1, \min \left\{ \frac{\sum_t t \cdot CF_t}{\sum_t CF_t}, 5 \right\} \right\}$$

CF_t indique le flux de trésorerie que le débiteur doit payer durant la période t .

2.4 L'EXPOSITION EN CAS DE DÉFAUT (EAD)

2.4.1 Définition et calcul de l'EAD

Les mesures précédentes renseignent sur la survenance et l'ampleur d'un défaut potentiel sans donner exactement le montant qui peut être perdu. Un autre déterminant du risque de crédit est le montant dû. L'exposition est définie comme la totalité des actifs qui sont liés à des transactions sur le marché, ou avec un client, et enregistrés dans le bilan et le hors bilan de la banque. L'exposition en cas de défaut peut être définie comme étant la valeur nominative des montants des prêts au cours de leur vie. Le montant de l'exposition est alors réajusté au cours du temps selon le mode de remboursement et d'amortissement.

En d'autres termes, la valeur exposée au risque (Exposure At Default) peut être définie comme l'exposition du Groupe en cas de défaut de la contrepartie. L'EAD comprend des expositions inscrites au bilan et au hors bilan. Les expositions hors bilan sont converties en



équivalent bilan à l'aide de facteurs de conversion (Crédit Conversion Factor, CCF) internes ou réglementaires.

Pour chaque crédit i , l'exposition en cas de défaut est notée EAD_i . Elle est évaluée dans le cadre du risque de crédit de manière déterministe (dans un souci de simplicité). Pour le risque de contrepartie, il est considéré comme une variable aléatoire et est calculé en référence à l'approche réglementaire retenue (méthode standard ou interne).

2.4.2 Méthode de calcul de l'EAD

Le calcul de l'EAD est issu de celui de la valeur exposée au risque, tel que défini dans les textes réglementaires. La valeur exposée au risque comporte trois dimensions :

- ☞ la valeur exposée au risque des éléments du bilan (leur valeur comptable). Ce montant est éventuellement ajusté en conformité à la réglementation et est observé ou déterminé à la date de calcul de manière déterministe.
- ☞ la valeur exposée au risque des éléments du hors bilan correspondant au montant engagé mais non tiré.
- ☞ un facteur de conversion qui est défini comme le rapport entre la partie actuellement non prélevée d'une ligne de crédit qui pourrait être prélevée et serait donc exposée en cas de défaut et la partie actuellement non prélevée de cette ligne de crédit.

$$CCF = \frac{EAD - \text{Encours de Bilan la date de calcul}}{\text{Limite} - \text{Encours de Bilan la date de calcul}}$$

Le facteur de conversion étant inconnu à la date de calcul, il fait l'objet d'un modèle et à défaut d'un dire d'expert.

2.4.3 Modélisation de l'EAD

La modélisation de l'EAD passe successivement par les étapes de différenciation du risque et de quantification du risque comme la PD et la LGD.

2.4.3.1 Différenciation du risque

Durant cette étape, les risk drivers pertinents sont spécifiés, les méthodes statistiques adéquates sont retenues et les paramètres du modèle sont estimés.

Conformément aux distinctions précédentes entre bilan et hors bilan, cette étape a pour but une segmentation experte nécessaire à l'identification des segments où le calcul du CCF n'est pas pertinent (*region of instability*), voire impossible. Pour ces segments, l'EAD est estimée directement au moyen du taux de variation d'encours. Enfin, la segmentation experte peut être affinée par des méthodes statistiques lorsque le volume de données le permet.



2.4.3.2 Quantification du risque

Cette partie du processus d'estimation des paramètres de risque conduit à une quantification appropriée du risque en s'assurant que les estimations de CCF correspondent au CCF moyen de long terme ou à l'estimation de CCF *downturn* à un niveau pertinent pour la méthode appliquée.

Le CCF moyen de long terme d'un segment est obtenu en calculant la moyenne arithmétique simple des CCF observés du segment considéré. Cette moyenne peut être très élevée à cause des valeurs extrêmes de CCF, ce qui va se traduire par des estimations d'EAD très élevées et des charges en capital plus importantes.

Au CCF moyen de long terme s'ajoutent les marges de conservatisme pour corriger des biais éventuels dans la quantification du CCF, ou une incertitude accrue qui n'est pas entièrement prise en compte par l'erreur d'estimation générale.

2.5 PERTES ANTICIPÉES

La perte anticipée est le coût moyen qui découlerait du défaut d'un crédit (KARUSISI 2007). Si on considère la variable aléatoire de Bernoulli L_i qui matérialise le défaut ou non, la perte anticipée peut être vue comme l'espérance mathématique de la réalisation de cet événement multiplié par le montant de l'exposition potentiellement perdu.

Ainsi, on définit L_i de la manière suivante :

$$L_i = \begin{cases} 1 & \text{si défaut} & \mathbb{P}(L_i = 1) = PD_i \\ 0 & \text{sinon} & \mathbb{P}(L_i = 0) = 1 - PD_i \end{cases}$$

Et, la perte potentielle est obtenue à l'aide de la variable aléatoire :

$$\tilde{L} = EAD \times LGD \times L$$

Cette forme de la perte potentielle suppose le montant d'exposition à défaut EAD et le taux de perte en cas de défaut LGD déterministes. La seule variable de l'écriture est le défaut L . De cette écriture, on déduit la perte anticipée, notée EL pour *Expected Loss*, comme :

$$EL = EAD \times LGD \times \mathbb{E}[L] = EAD \times LGD \times PD$$

La réglementation pousse les banques à constituer un montant de **provisions** au moins égal à la perte anticipée.



2.6 PERTES NON ANTICIPÉES

En raison de sa nature, le risque réside dans la partie de la perte qui ne peut être maîtrisée par l'institution financière. De ce fait, en plus de la perte anticipée, les banques doivent avoir une mesure de la variabilité de la perte autour de la valeur moyenne. Les outils statistiques fournissent l'écart-type comme mesure de cette variabilité. On a alors la perte non anticipée (*Unexpected Loss*) qui est définie comme :

$$UL = \sqrt{\text{Var}(\tilde{L})} = \sqrt{EAD^2 \times LGD^2 \times \text{Var}(L)}$$

L étant une variable de Bernoulli, on obtient :

$$UL = EAD \times LGD \times \sqrt{PD(1 - PD)}$$

Cette mesure de risque souffre d'une faiblesse majeure : elle est robuste uniquement si la distribution des pertes suit une loi gaussienne. En effet, elle suppose que la répartition des réalisations est symétrique par rapport à la valeur moyenne. On verra dans la suite que cette distribution présente un étalement sur la droite, signe de pertes supérieures à la moyenne moins fréquentes, mais plus sévères. Une autre mesure de risque plus adéquate sera alors introduite.

En résumé

Une gestion des risques pertinente passe par une mesure fiable des pertes potentielles courues par la firme. Ces dernières étant une combinaison des connaissances de l'historique de la firme ainsi que de la santé économique actuelle, elles reposent sur une estimation cohérente des défauts futurs. Ainsi, ce chapitre a su asseoir le cadre théorique dans un contexte bancaire d'estimation des métriques indispensables à la mesure des impacts du défaut. Il s'est précisément agi d'étudier la probabilité de défaut, la perte en cas de défaut, la maturité et l'exposition à défaut d'un contrat. La question qui se pose alors est la méthode de combinaison de ces métriques pour la mesure du risque de crédit. Une première approche par l'écart-type a été proposée dans ce chapitre. Cette approche, bien que pertinente, repose sur l'hypothèse de normalité de la distribution des pertes. En pratique, cette hypothèse n'étant presque jamais vérifiée, nous introduirons l'approche réglementaire d'évaluation du risque de crédit dans le chapitre suivant.

RISQUE DE CRÉDIT AU NIVEAU DU PORTEFEUILLE

Dans ce chapitre, nous présentons les indicateurs d'évaluation du risque de crédit au niveau du portefeuille de la banque. En particulier, nous relatons la construction de la formule réglementaire utilisée et nous en faisons une critique.

3.1	Distribution des pertes futures	27
3.2	Les risques pondérés : la formule réglementaire de Bâle 2	28
3.2.1	Construction de la formule réglementaire	29
3.2.1.1	Le modèle de Merton (1974)	29
3.2.1.2	Modèle monofactoriel de Gordy-Vasicek	30
3.2.1.3	Formule réglementaire	33
3.3	Etude et analyse de la formule réglementaire de Bâle II	34
3.3.1	Discussion sur la formule réglementaire	34
3.3.2	Représentation graphique de la fonction réglementaire	36
3.3.2.1	Construction du jeu de données théorique	36
3.3.2.2	Analyse graphique	36

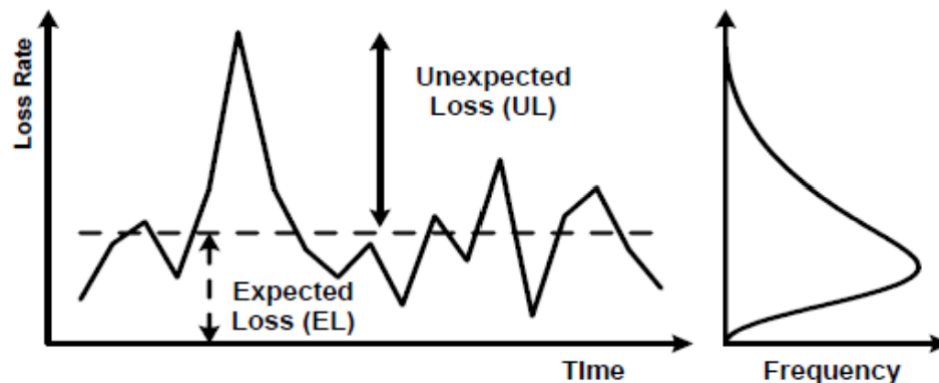


3.1 DISTRIBUTION DES PERTES FUTURES

Selon l'approche par modèles internes, les institutions bancaires font recours à leurs propres mesures d'indicateurs de risque comme paramètres d'entrée d'une formule de calcul du risque sur le portefeuille. Les mesures des paramètres de risques sont effectuées par exposition et converties en exigences en fonds propres par le biais de courbe de pondération.

Ces courbes de pondération sont obtenues en observant les pertes effectivement subies au cours d'une année tout en notant que ces réalisations diffèrent d'une année à une autre suivant le nombre et la sévérité des défauts. En pratique, la figure suivante montre comment la courbe de fréquence est construite.

FIGURE 3.1 – Variation des pertes annuelles réalisées à travers le temps pour une banque donnée



Source : CHARLES 2015

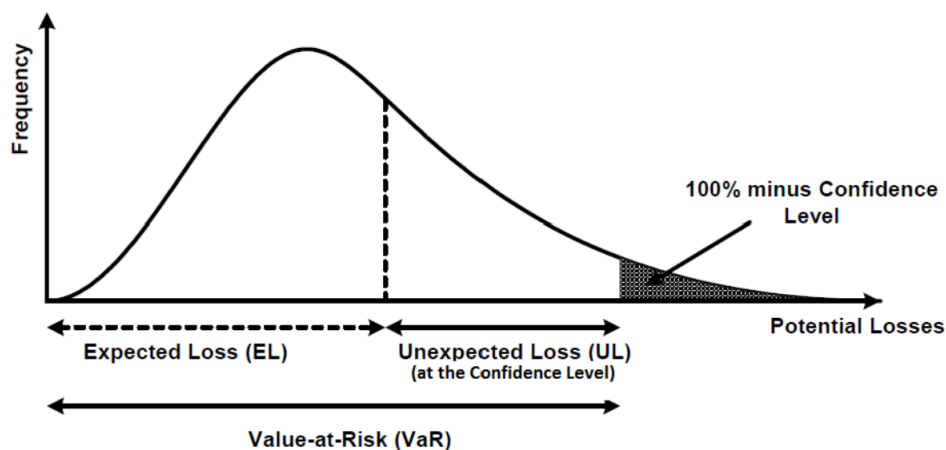
Le graphique 3.1 met en lumière le niveau des pertes attendues et de celles inattendues. Pour qu'un établissement financier puisse demeurer solvable, il doit disposer d'assez de capital pour couvrir le niveau des pertes inattendues. On pourrait alors penser qu'il suffirait de mobiliser un capital égal au niveau de l'ensemble du portefeuille détenu par la banque. Toutefois, cette approche « naïve » est économiquement inefficace. En effet, tout capital mobilisé en moins pour la couverture de risque pourrait être utilisé pour des investissements profitables. Ainsi, l'enjeu de la mesure de risque réside dans la mobilisation d'un niveau de capital pouvant couvrir les pertes futures tout en permettant à la banque de réaliser de manière optimale des investissements.

L'approche IRBA adoptée par Bâle II fournit un cadre pour la mesure du capital nécessaire. Cette approche s'appuie sur un modèle stochastique et permet l'estimation du montant des pertes. Ce dernier est estimé de sorte qu'il n'existe qu'une probabilité très faible que les pertes réalisées puissent excéder celles estimées. Enfin, la probabilité retenue est choisie pour satisfaire une fréquence de faillite des banques acceptable.

A partir de la distribution des pertes futures, il suffit alors d'avoir recours au quantile pour obtenir le niveau de pertes associé à une probabilité de réalisation donnée.



FIGURE 3.2 – Relation entre la VaR, l'EL et l'UL



Source : CHARLES 2015

Le graphique 3.2 met en exergue la *Value-at-Risk (VaR)* qui correspond au seuil à partir duquel la perte réalisée est jugée excessive, on observe alors la **faillite**. Pour un portefeuille, la $VaR_{1-\alpha}$ pour un niveau de confiance $1 - \alpha$, $\alpha \in [0, 1]$ se définit par l'équation suivante :

$$\mathbb{P}(\tilde{L} > VaR_{1-\alpha}) = \alpha$$

Le choix de la *VaR* comme mesure du risque montre que l'on s'intéresse à la sévérité des pertes. Sous Bâle II, le montant de capital requis a été fixé à 99,9%.

3.2 LES RISQUES PONDÉRÉS : LA FORMULE RÉGLEMENTAIRE DE BÂLE 2

Suivant l'approche *advanced Internal Rating Based Approach (A-IRBA)* les paramètres nécessaires au calcul du risque pondéré sont fournis par la banque elle-même :

- la probabilité de défaut à un an de chaque contrepartie ou emprunteur. Ces probabilités de défaut sont issues des échelles de notation interne ;
- le taux de perte en cas de défaut ;
- l'exposition à défaut (montant moyen de perte en cas de défaut) ;
- et la maturité effective de chaque position.

En *foundation Internal Rating Based Approach (F-IRBA)*, la banque n'estime que la probabilité de défaut pendant que les autres paramètres sont fournis par le régulateur.

Le *Risk Weighted Asset* ou risque pondéré correspond au produit de l'exposition par la pondération de risque $\psi(\cdot)$ et reflète donc le risque intrinsèque au deal.

$$RWA = EAD \times \psi(PD, LGD, M)$$



3.2.1 Construction de la formule réglementaire

3.2.1.1 Le modèle de Merton (1974)

Le modèle de risque de crédit proposé par Merton est le point de départ de plusieurs modèles dans le cadre de l'approche structurelle. Cette approche est directement liée à la valeur des actifs d'une compagnie. Ainsi, la compagnie sera considérée comme étant en défaut lorsque la valeur des actifs touche une barrière définie. Le temps avant défaut τ est alors un temps d'arrêt prévisible par rapport aux informations disponibles. Dans cette approche, on se réfère aux informations et données qui reflètent le bilan du débiteur.

Merton cherche à évaluer la probabilité de défaut d'une firme à un certain horizon à l'aide de l'historique ou une mesure de risque neutre. Pour lui, le bilan d'une entreprise peut être résumé en :

- ▣ actifs V_t d'un côté ;
- ▣ et passifs de l'autre. On y retrouve les dettes D_t et les fonds propres E_t .

Ces différentes valeurs sont aléatoires et dépendantes du temps hormis la dette qui est supposée constante au cours du temps (D , hypothèse de simplification). De plus, Merton émet les hypothèses suivantes LONGIN 2014 :

- ◆ **marché parfait** : absence de coûts de transaction, d'impôts et de taxes ;
- ◆ **marché complet** : tous les actifs sont répliquables, c'est-à-dire qu'ils peuvent s'écrire comme une combinaison des autres actifs dans tous les états du monde ;
- ◆ **absence d'opportunité d'arbitrage** : il n'y a pas de profit sans risque ;
- ◆ **validité du théorème de Modigliani-Miller** : indépendance de la valeur de l'entreprise par rapport à sa structure en capital ;
- ◆ **taux d'intérêt constant** r .

A l'aide de ces hypothèses, on est en mesure de considérer une entreprise en faillite lorsque la valeur de ses actifs se retrouve en dessous du niveau de ses dettes à maturité T . Ainsi, la valeur des fonds propres à maturité est alors :

$$E_T = (V_T - D)^+$$

En considérant la probabilité risque neutre \mathbb{Q} et un facteur d'actualisation (taux zéro-coupon) $B(0, T)$, on estime la valeur des fonds propres à la date initiale par :

$$E_0 = B(0, T) \cdot \mathbb{E}_{\mathbb{Q}} [(V_T - D)^+]$$

Les fonds propres peuvent être considérés comme un **call européen sur l'actif avec un strike D et une échéance en T** .



En supposant de plus que l'actif V_t admet la dynamique $dV_t = \mu V_t dt + \sigma V_t dW_t$ avec W_t un mouvement Brownien, on peut appliquer la formule de **Black-Scholes** et ainsi obtenir la formule caractéristique de l'approche structurelle de Merton :

$$E_0 = V_0 \Phi(d_1) - B(0, T) \cdot D \Phi(d_2) \quad \text{avec} \quad \begin{cases} d_1 := \frac{\ln(V_0/D) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \\ d_2 := d_1 - \sigma\sqrt{T} \end{cases}$$

Sous la mesure de risque neutre, $dV_t = rV_t dt + \sigma V_t dW_t$ et

$$\mathbb{P}(\tau \leq t) = \mathbb{P}(V_t \leq D) = \Phi\left(\frac{1}{\sigma\sqrt{t}} \left[\ln\left(\frac{D}{V_0}\right) - rt + \frac{\sigma^2 t}{2} \right]\right)$$

Le modèle structurel de Merton est un cadre simplifié qui présente quelques limites :

- le niveau de dette devrait être endogène : en effet, lorsque la firme est proche du défaut, le niveau de dette augmente ;
- l'utilisation d'une option à barrière¹ est plus réaliste vu que le défaut peut survenir à n'importe quelle date ;
- la valeur de l'actif est considérée continue et donc n'induit aucune perte au moment du défaut. Or, la valeur de l'actif peut suivre un processus à saut et donc engendrer des pertes extrêmes à un moment donné.

3.2.1.2 Modèle monofactoriel de Gordy-Vasicek

Le modèle de base de Gordy-Vasicek est un modèle utilisé pour la quantification du risque de crédit. Il s'applique sur un portefeuille contenant un nombre élevé de créances bien diversifiées. Il se fonde sur l'hypothèse de portefeuille infiniment granulaire (*infinitely granular portfolio assumption*).

Infinitely granular portfolio assumption

Cette hypothèse implique :

- l'existence d'un très grand nombre (potentiellement infini) de contreparties ;
- que chaque exposition est négligeable vis-à-vis de l'exposition totale du portefeuille.

Cette hypothèse permet des simplifications dans le calcul de la Value-at-Risk.

Supposons que le portefeuille comporte n emprunteurs. On note par $D_i = \mathbb{1}_{\{V_i \leq \gamma_i\}}$ la variable aléatoire matérialisant le défaut de chaque emprunteur dans le cadre du modèle de Merton.

1. Une option à barrière est un instrument financier qui s'active ou se désactive en fonction de l'évolution du sous-jacent sur laquelle elle porte. Ainsi, le passage du sous-jacent au-dessus ou en dessous d'une valeur limite (la barrière) active ou désactive l'option. On distingue donc les options à barrière activante qui permettent l'exercice de l'option, des options à barrière désactivante qui entraîne la cessation de l'option.



On considère aussi $A_i = EAD_i \times LGD_i$ la perte engendrée par l'emprunteur i , supposée indépendante de D_i . La perte totale sur le portefeuille est alors :

$$L_n = \sum_{i=1}^n A_i \cdot D_i$$

Dans ce modèle, la rentabilité standardisée des actifs (V_i) peut être exprimée en fonction d'un facteur unique de risque systématique (\mathcal{G}) et d'un facteur de risque spécifique diversifiable ε_i (facteur de risque idiosyncratique). Dans la littérature, on retrouve deux versions équivalentes de ce modèle :

$$\begin{cases} V_i = \rho_i \mathcal{G} + \sqrt{1 - \rho_i^2} \cdot \varepsilon_i & \text{(Gordy)} \\ V_i = \sqrt{\rho_i} \mathcal{G} + \sqrt{1 - \rho_i} \cdot \varepsilon_i & \text{(Vasicek)} \end{cases} \quad \forall i, \varepsilon_i \perp \mathcal{G}, \varepsilon_i \perp \varepsilon_j \quad i \neq j, \rho_i > 0, \mathcal{G}, \varepsilon_i \rightarrow \Phi$$

Obtention de l'expression de la rentabilité standardisée

Dans le modèle de Gordy-Vasicek avec l'hypothèse de normalité des facteurs de risque systématique et spécifique on écrit :

$$R_i = \psi_i \varepsilon_i + \mathcal{G} w_i \rightarrow \mathcal{N}(0, 1)$$

De ce fait, on a :

$$\begin{aligned} \mathbb{E}(R_i) &= \psi_i \mathbb{E}(\varepsilon_i) + \mathbb{E}(\mathcal{G}) w_i = 0 \\ \mathbb{V}(R_i) &= w_i^2 \mathbb{V}(\mathcal{G}) + \psi_i^2 \mathbb{V}(\varepsilon_i) + 2w_i \psi_i \text{Cov}(\mathcal{G}, \varepsilon_i) \\ &= w_i^2 + \psi_i^2 = 1 \end{aligned}$$

On déduit que $\psi_i = \sqrt{1 - w_i^2}$. On obtient bien la formule :

$$R_i = \rho_i \mathcal{G} + \sqrt{1 - \rho_i^2} \cdot \varepsilon_i$$

A base de cette expression, Gordy calcule la probabilité de défaut conditionnelle au facteur de risque systématique de la manière suivante :

$$\begin{aligned} p_i(g) &= \mathbb{P}(V_i \leq \gamma_i | \mathcal{G} = g) \\ &= \mathbb{P}(\rho_i \mathcal{G} + \sqrt{1 - \rho_i^2} \cdot \varepsilon_i \leq \gamma_i | \mathcal{G} = g) \\ &= \mathbb{P}\left(\varepsilon_i \leq \frac{\gamma_i - \rho_i \mathcal{G}}{\sqrt{1 - \rho_i^2}} | \mathcal{G} = g\right) \\ &= \Phi\left(\frac{\gamma_i - \rho_i g}{\sqrt{1 - \rho_i^2}}\right) \end{aligned}$$



On cherche donc à calculer $\mathbb{E}[L_n]$ et la $VaR[L_n]$. On exprime alors les probabilités conditionnelles suivantes :

$$\mathbb{E}[L|\mathcal{G}] = \sum_{i=1}^n \bar{A}_i \cdot p_i(\mathcal{G}) \quad \text{avec } p_i(\mathcal{G}) = \mathbb{E}[D_i|\mathcal{G}] = \Phi\left(\frac{\gamma_i - \rho_i \mathcal{G}}{\sqrt{1 - \rho_i^2}}\right)$$

L'hypothèse de portefeuille infiniment granulaire entraîne que le nombre de contreparties est très grand (n grand) et donc que la variable aléatoire de perte L_n est induite uniquement par la variable \mathcal{G} .

$$L_n \simeq \mathbb{E}[L_n|\mathcal{G}] \implies VaR_\alpha(L_n) \simeq VaR_\alpha(\mathbb{E}[L_n|\mathcal{G}])$$

De ce fait, si le facteur systémique est égal à son propre quantile $1 - \alpha$, la perte est égale à la Value-at-Risk de niveau α .

$$VaR_\alpha(L) = \sum_{i=1}^n \bar{A}_i \cdot p_i(g_\alpha), \quad g_\alpha = \Phi^{-1}(1 - \alpha)$$

Preuve du calcul de la VaR

Sous l'hypothèse de portefeuille infiniment granulaire on a $L_n \simeq \mathbb{E}[L_n|\mathcal{G}]$, on déduit $\forall x$

$$\begin{aligned} \mathbb{P}(L_n \leq x) &= \mathbb{E}[\mathbb{1}_{(L_n \leq x)}] \\ &\simeq \mathbb{E}[\mathbb{1}_{(\mathbb{E}[L_n|\mathcal{G}] \leq x)}] \\ &= \mathbb{E}[\mathbb{1}_{(\sum_{i=1}^n \bar{A}_i \cdot p_i(\mathcal{G}) \leq x)}] \\ &= \mathbb{P}\left(\sum_{i=1}^n \bar{A}_i \cdot p_i(\mathcal{G}) \leq x\right) \end{aligned}$$

Supposant que $\mathcal{G} \rightarrow \mathbb{P}_{i,\mathcal{G}}^a$ est strictement décroissante et notant $g_\alpha = \Phi^{-1}(1 - \alpha)$ le quantile, on a :

$$\begin{aligned} \mathbb{P}\left(L_n \geq \sum_{i=1}^n \bar{A}_i \cdot p_i(g_\alpha)\right) &= \mathbb{P}\left(\sum_{i=1}^n \bar{A}_i \cdot p_i(\mathcal{G}) \geq \sum_{i=1}^n \bar{A}_i \cdot p_i(g_\alpha)\right) \\ &= \mathbb{P}(\mathcal{G} \leq g_\alpha) \\ &= 1 - \alpha \end{aligned}$$

Ainsi, la VaR de niveau α pour un portefeuille infiniment granulaire tel que le facteur systémique est égal à son quantile d'ordre $1 - \alpha$ est :

$$VaR_\alpha(L) = \sum_{i=1}^n \bar{A}_i \cdot p_i(g_\alpha)$$

^a. $\mathbb{P}_{i,\mathcal{G}}$: mesure de probabilité associée au facteur de risque systémique évalué sur l'emprunteur i .



En considérant dans la suite la copule gaussienne à un facteur pour le portefeuille de crédit :

$$\begin{aligned}
 VaR_\alpha(L) &= \mathbb{E}[L|\mathcal{G} = VaR_{1-\alpha}(\mathcal{G})] \\
 &= \mathbb{E} \left[\sum_{i=1}^n EAD_i \cdot LGD_i \cdot \mathbb{1}_{(V_i \leq \gamma_i)} | \mathcal{G} = q_{1-\alpha} \right] \\
 &= \sum_{i=1}^n EAD_i \cdot LGD_i \cdot \mathbb{P}(V_i \leq \gamma_i | \mathcal{G} = q_{1-\alpha}) \\
 &= \sum_{i=1}^n EAD_i \cdot LGD_i \cdot \Phi \left(\frac{\gamma_i - \rho_i q_{1-\alpha}}{\sqrt{1 - \rho_i^2}} \right) \quad , \quad \gamma_i = \Phi^{-1}(p_i)
 \end{aligned}$$

p_i étant la probabilité de défaut inconditionnelle à un an d'une firme.

3.2.1.3 Formule réglementaire

La formule réglementaire est issue du modèle de Gordy-Vasicek basé sur l'approche structurelle de **Merton**.

Elle est obtenue en retirant de l'estimation de la perte future (VaR), le montant de perte attendu.

$$\begin{aligned}
 RWA &= VaR_\alpha(L) - \mathbb{E}(L) \\
 &= \sum_{i=1}^n EAD_i \cdot LGD_i \left[\Phi \left(\frac{\Phi^{-1}(PD_i) - \rho_i q_{1-\alpha}}{\sqrt{1 - \rho_i^2}} \right) - PD_i \right]
 \end{aligned}$$

Afin d'intégrer certaines spécificités de la réalité économique, le comité de Bâle a pris comme décision de :

- ♦ définir $\alpha = 99,9\%$;
- ♦ considérer des coefficients ρ_i^2 qui dépendent de PD_i et varie entre 12% et 24%. Mieux la firme est notée, plus son coefficient de facteur systémique est élevé. Cela veut dire que la qualité de crédit des entreprises les mieux notées est plus corrélée au marché que les autres ;

$$\rho_i^2 = 0,12 \cdot \left(\frac{1 - e^{-50 \cdot PD_i}}{1 - e^{-50}} \right) + 0,24 \cdot \left(1 - \frac{1 - e^{-50 \cdot PD_i}}{1 - e^{-50}} \right)$$

- ♦ rajouter un facteur d'ajustement de maturité pour tenir compte du caractère plus risqué des prêts de long terme. Pour une PD donnée, les crédits de long-terme sont plus risqués que les crédits de court-terme ;

$$MatAdj(M_i, PD_i) = \frac{1 + (M_i - 2,5)b_i}{1 - 1,5b_i} \quad \text{avec} \quad b_i = (0,11852 - 0,05478 \cdot \ln(PD_i))^2$$

- ♦ rajouter un facteur de **12,5** pour ramener le tout à un ratio de capital de 8% ($12,5 = \frac{1}{8\%}$) ;



- ✦ considérer sur la base des études quantitatives d'impact (QIS) un facteur multiplicatif de **1,06**. Ce facteur multiplicatif qui prend le nom de *scaling factor* est rajouté pour atteindre l'objectif de maintien du niveau global des fonds propres au niveau souhaité.

Ainsi, on obtient la formule réglementaire de Bâle II d'évaluation du risque de crédit par la formule suivante :

$$RWA_i = EAD_i \cdot LGD_i \left[\Phi \left(\frac{\Phi^{-1}(PD_i) + \rho_i \Phi^{-1}(0,999)}{\sqrt{1 - \rho_i^2}} \right) - PD_i \right] \frac{1 + (M_i - 2,5)b_i}{1 - 1,5b_i} \cdot 12,5 \cdot 1,06$$

Au final, la formule du *RWA* préconise pour couvrir les pertes inattendues de calculer la *VaR* à 99,9% (le quantile à 99,9% de la perte) à laquelle on retranche la perte attendue. Cette idée est un principe général de la gestion des risques.

3.3 ETUDE ET ANALYSE DE LA FORMULE RÉGLEMENTAIRE DE BÂLE II

La fonction réglementaire de calcul du *RWA* peut être découpée en deux parties. La première composée du montant de l'exposition (*EAD*) et la seconde qui mesure l'amplitude du risque vis-à-vis de l'exposition (*RW*). Le niveau de risque *RW* est le produit de la perte en cas de défaut (*LGD*) et d'une fonction prenant en paramètre la probabilité de défaut (*PD*) et la maturité (*M*). La fonction est la suivante :

$$RWA = EAD \cdot LGD \cdot \underbrace{\left[\Phi \left(\frac{\Phi^{-1}(PD) + \rho \cdot \Phi^{-1}(0,999)}{\sqrt{1 - \rho^2}} \right) - PD \right] \frac{1 + (M - 2,5)b}{1 - 1,5b} \cdot 12,5 \cdot 1,06}_{\text{fonction d'intérêt : } RW}$$

Nous nous intéressons dans notre analyse à la fonction d'intérêt qui matérialise l'amplitude du risque vis-à-vis d'une exposition. On rappelle que les paramètres ρ et b sont tous deux fonctions de la *PD* et définis plus haut.

3.3.1 Discussion sur la formule réglementaire

La formule réglementaire d'évaluation du risque de crédit a fait l'objet de critiques depuis sa mise en place par le comité de Bâle. Cette section est inspirée de DHIMA 2020 et a pour but de faire un résumé des principales critiques énoncées à son encontre.

L'approche IRBA vise à donner une certaine latitude aux institutions financières dans l'évaluation du niveau de risque couru. Cette approche dans sa forme avancée donne libre cours aux banques à l'estimation des paramètres de risque, mais reste opaque à la méthode d'agrégation de ces paramètres. En effet, la formule mathématique permettant de calculer le niveau minimal de fonds propres à l'aide des paramètres estimés est fournie par le régulateur.



Cette formule réglementaire étant basée sur le modèle de **Gordy-Vasicek** découle d'hypothèses difficilement vérifiables en pratique.

D'abord, le non-respect de l'hypothèse de portefeuille parfaitement granulaire (parfaite diversification) entraîne une sous-évaluation du risque. Si le portefeuille n'est pas assez diversifié, il présente une concentration qui rend les pertes plus élevées en cas de défaut (contagion du défaut). La fonction d'évaluation du capital réglementaire issue de la fonction de perte inattendue conditionnelle n'est alors plus adaptée. Les banques étant conscientes de cette limite font usage d'un reporting particulier des grands risques pour évaluer le risque de concentration de leur portefeuille. Toutefois, ces reportings se déclinent en stress tests où les paramètres de risque des contreparties sont choqués par rapport à leur risque spécifique. Procéder de la sorte s'avère insuffisant lorsque les portefeuilles sont très concentrés du fait des sources plus larges du risque spécifique (facteurs de risque commun et individuel). La version du modèle de Gordy-Vasicek utilisée par le comité de Bâle se base sur une PD moyenne reflétant uniquement le risque commun.

Ensuite, l'approche monofactorielle du risque commun découlant de la remarque précédente limite grandement les sources de risque commun à la seule conjoncture macroéconomique. Il serait plus vraisemblable de déterminer pour chaque secteur les facteurs de risque systématique impactants. Cette approche a le mérite d'inclure les principaux facteurs de risque communs ayant un impact à moyen terme sur les contreparties d'un secteur. Elle nécessiterait alors que les facteurs de sensibilité des secteurs soient estimés et publiés suivant des intervalles réguliers et correspondants aux mouvements de l'économie. Le facteur de risque systématique inclus actuellement dans la formule réglementaire date de la mise en place de Bâle II et ne reflète plus l'état de l'économie dans son ensemble.

Enfin, la probabilité de défaut conditionnelle dans le modèle de Gordy-Vasicek est une fonction de la probabilité de défaut et du facteur de risque commun. D'un point de vue théorique, répercuter sur la probabilité de défaut inconditionnelle un choc macroéconomique lors des stress tests est incohérent avec l'inconditionnalité dans le modèle de Gordy-Vasicek. Une confusion entre probabilités de défaut conditionnelle et inconditionnelle semble parfois commise dans les estimations.

Ces limites nous poussent à remettre en cause la forme de la fonction réglementaire de calcul du risque de crédit. Une alternative serait de faire appel à une généralisation multifactorielle de la formule réglementaire. Cette solution bien que pertinente induit une complexité plus grande dans l'évaluation du risque de crédit. Ainsi, nous proposons de construire une fonction alternative à celle actuelle qui aura l'intérêt de fournir une interprétation et un monitoring plus aisés du risque de crédit malgré la sous-évaluation du risque de crédit.



3.3.2 Représentation graphique de la fonction réglementaire

La formule réglementaire d'évaluation du risque de crédit peut être vue comme le produit de deux quantités : le montant de l'exposition d'une part, et l'ampleur du risque d'autre part. Dans cette section, nous essayons donc d'obtenir un aperçu visuel de la fonction d'intérêt (RW). En considérant un ensemble de données purement théorique basé sur les plages de réalisation des paramètres de la fonction réglementaire, on réussit à construire une représentation graphique de l'ensemble des possibles du RW .

3.3.2.1 Construction du jeu de données théorique

Les données théoriques que nous simulons ont pour première utilité la construction graphique de la formule réglementaire. De ce fait, le jeu de données est construit de sorte à parcourir un éventail assez large des combinaisons de valeurs possibles. Ainsi, nous générons des **séquences** de données de la manière suivante :

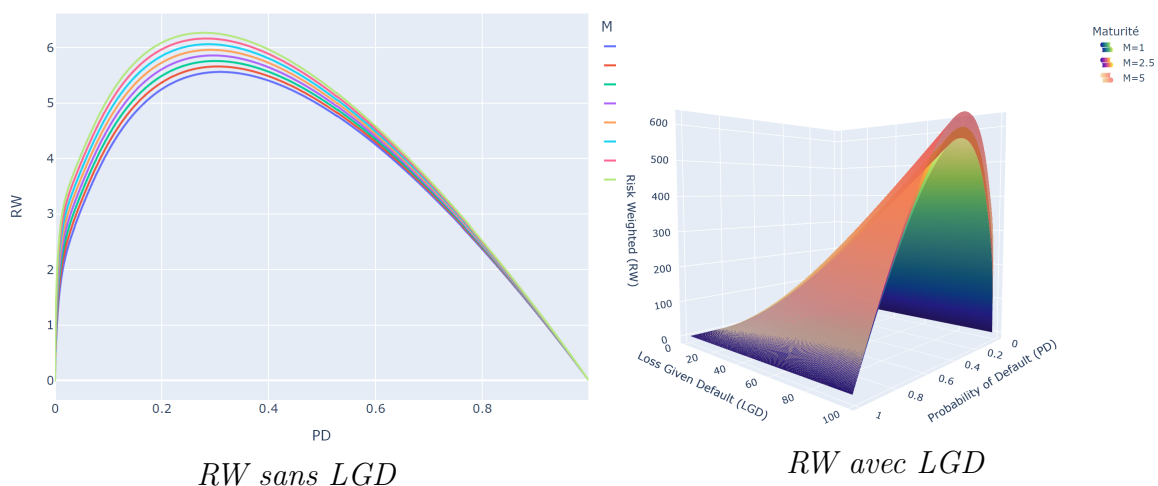
- PD : de 0,1% à 100% par pas de 0,1% (999 valeurs)²
- LGD : de 1% à 100% par pas de 1% (100 valeurs)
- M : de 1 à 5 par pas de 0,5 (8 valeurs)

La base de données est alors obtenue par l'ensemble des combinaisons possibles des séquences générées. On obtient une base composée de 799200 emprunts.

3.3.2.2 Analyse graphique

Dans un premier temps, nous ne considérons pas la LGD dans la formule. On peut remarquer que la fonction est plus étalée à droite qu'à gauche. Ainsi, pour des valeurs faibles de PD , une variation sensible fait plus rapidement varier le RW que pour des valeurs plus élevées de la PD .

FIGURE 3.3 – Représentation de la fonction réglementaire en 2D et 3D



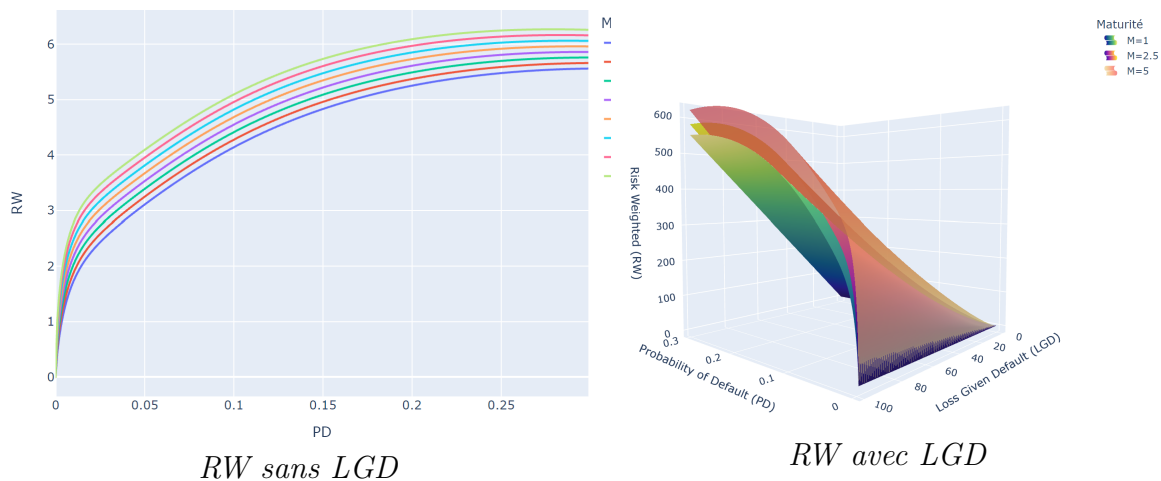
2. On ne génère pas de valeur nulle pour la probabilité de défaut en raison du plancher réglementaire fixé à 0,03%. De plus, la fonction réglementaire n'admet pas de valeur nulle pour la PD .



Aussi, en considérant dans un second temps, la représentation en trois dimensions (3D) du RW avec la prise en compte de la LGD, on retrouve des résultats similaires que précédemment. Comme attendu, le RW varie positivement à une augmentation de LGD et de la maturité.

Toutefois, ces représentations révèlent qu'après un seuil autour de 0,3 pour la PD, on observe une baisse du RW. Cela s'explique par le fait que cette partie de la fonction n'a aucun intérêt. En effet, la probabilité de défaut la plus élevée octroyée est de 0,27. Au-delà, on se retrouve dans le défaut et aucun calcul n'est nécessaire, car le risque encouru est directement de 1. De ce fait, nous reprenons les représentations en nous limitant à une valeur de 0,3 pour la PD.

FIGURE 3.4 – Fonction réglementaire pour les valeurs de PD sain



On obtient cette fois-ci une fonction strictement croissante en ses paramètres, signe d'une mesure de risque de crédit cohérente.

En résumé

La formule réglementaire d'évaluation du risque de crédit est issue du modèle monofactoriel de Gordy-Vasicek. Ce dernier basé sur l'approche structurelle de Merton permet d'une formule fermée quantifiant la perte extrême que peut enregistrer la banque compte tenu d'un seuil fixé. Bien que ce modèle soit pertinent au vu de la distribution des pertes, elle souffre de quelques limites qui en réduisent l'intérêt. Toutefois, les alternatives proposées dont le passage au modèle multifactoriel induit une complexité plus grande dans l'évaluation du risque de crédit. Nous retenons donc de construire à l'aide de la formule réglementaire, une version alternative qui aura le mérite d'offrir des axes d'interprétation des effets des paramètres beaucoup plus aisés. Le chapitre suivant s'attelle à la modélisation alternative du risque de crédit à partir de la formule réglementaire.

MODÉLISATION ALTERNATIVE DU RISQUE DE CRÉDIT

APPROXIMATION DE LA FONCTION RÉGLEMENTAIRE

Ce chapitre fait une revue des approches que nous avons retenues pour l'approximation de la fonction réglementaire. D'abord, nous présentons le formalisme mathématique de toutes les approches. Ensuite, nous les appliquons sur un jeu de données théoriques afin de ressortir des conclusions préliminaires.

4.1	Approche analytique	40
4.1.1	Un intermède théorique : formule de Taylor	40
4.1.2	Proposition sur la formule réglementaire	41
4.1.3	Effet marginal et contribution	42
4.1.4	Application	42
4.2	Approche par modélisation classique	46
4.2.1	Régression linéaire	46
4.2.2	Régression polynomiale	49
4.2.3	Discussion des résultats et extensions	49
4.3	Approche par machine learning	50
4.3.1	Random Forest	50
4.3.2	Boosting : Gradient boosting et XGBoost	51
4.4	Approche par deep learning : Réseau de neurones	53
4.4.1	Présentation et formalisme mathématique	53
4.4.2	Application aux données simulées	54



Dans ce chapitre, toutes les applications sont réalisées sur le jeu de données théorique utilisé pour la représentation de la formule réglementaire (figure 3.4). La construction de cette base est présentée dans la section 3.3.2.1, page 36. On ne retient à partir d'ici que les observations ayant une $PD \leq 0.3$, soit 240000 observations.



4.1 APPROCHE ANALYTIQUE

La fonction de calcul du RW étant clairement définie, une première approche serait de déterminer à l'aide de cette fonction les effets marginaux de chaque paramètre.

4.1.1 Un intermède théorique : formule de Taylor

Soit f une fonction régulière (infiniment dérivable sur son domaine de définition) définie sur un ouvert U de \mathbb{R}^n .

Formule de Taylor-Young à l'ordre 2 : Pour tout point $h \in \mathbb{R}^n$, tel que $x_0 + h \in U$, on peut écrire :

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x_0) + \frac{1}{2} \sum_{i,j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) + o(\|h\|^2).$$

où $o(\|h\|^2)$ est une fonction qui tend vers 0 quand h tend vers 0.

On adopte la notation suivante pour simplification :

$$D^k f(a)(h)^k = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a) h_{i_1} \dots h_{i_k}$$

Supposons que f est une fonction définie sur un ouvert U de \mathbb{R}^n à valeurs dans \mathbb{R}^p et soit x_0 un point de U

Formule de Taylor-Young : si f est k fois différentiable en x_0 , alors

$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + \dots + \frac{1}{k!} D^k f(x_0)(h)^k + o(\|h\|^k).$$

Formule de Taylor avec reste intégral : si f est de classe C^{k+1} sur U et si le segment $[x_0, x_0 + h]$ est contenu dans U , on a :

$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + \dots + \frac{1}{k!} D^k f(x_0)(h)^k + \int_0^1 \frac{(1-t)^k}{k!} D^{k+1} f(x_0 + th)(h)^{k+1} dt$$

Formule de Taylor avec reste de Lagrange : si de plus $\max_{y \in [x_0, x_0 + h]} \|D^{n+1} f(y)\| \leq M$, alors

$$\left\| f(x_0 + h) - f(x_0) - \sum_{k=1}^n \frac{1}{k!} D^k f(x_0)(h)^k \right\| \leq \frac{M}{(n+1)!} \|h\|^{n+1}$$

Les formules précédentes permettent d'approcher de manière analytique l'effet de chaque



paramètre sur une fonction vérifiant les différentes hypothèses. La formule avec reste de Lagrange permet de quantifier l'erreur d'approximation. On fera recours à cette formule pour évaluer la qualité d'approximation par méthode analytique dans la suite du document.

4.1.2 Proposition sur la formule réglementaire

En nous basant sur les formules de Taylor, nous proposons dans cette section, des approximations des effets des paramètres de la formule réglementaire. On rappelle que la formule réglementaire se définit comme suit :

$$RWA = EAD \times RW$$

$$RW = \psi(PD, LGD, M)$$

$$RW = LGD \left[\Phi \left(\frac{\Phi^{-1}(PD) + \sqrt{R} \cdot \Phi^{-1}(0,999)}{\sqrt{1-R}} \right) - PD \right] \frac{1 + (M - 2,5)b}{1 - 1,5b} \cdot 12,5 \cdot 1,06$$

Avec $R = 0,12 \cdot \left(\frac{1-e^{-50 \cdot PD}}{1-e^{-50}} \right) + 0,24 \cdot \left(1 - \frac{1-e^{-50 \cdot PD}}{1-e^{-50}} \right)$ et $b = (0,11852 - 0,05478 \cdot \ln(PD))^2$.

On pose $X = (PD, LGD, M)$ et x une réalisation de X . Pour $PD \in]0,1]$, $LGD \in \mathbb{R}^+$ et $M \in \mathbb{R}^+$ on montre de manière triviale que la fonction ψ est C^∞ .

Soit $D =]0,1] \times \mathbb{R}^+ \times \mathbb{R}^+$, pour $x_0 \in D$ et tout $h \in \mathbb{R}^3$ tel que $x_0 + h \in D$, on obtient la relation suivante en appliquant la formule de Taylor avec reste de Lagrange au RW :

$$\left\| \Delta RW - \widehat{\Delta RW} \right\| = \left\| \psi(x_0 + h) - \psi(x_0) - \sum_{k=1}^n \frac{1}{k!} D^k \psi(x_0) (h)^k \right\| \leq \frac{T}{(n+1)!} \|h\|^{n+1}$$

Une démonstration en annexe donne une forme de la borne T . On suppose de plus dans la suite que $\|h\| \rightarrow 0$. Cette dernière hypothèse assure que $\frac{T}{(n+1)!} \|h\|^{n+1} \rightarrow 0$. On admet donc les formules de différentielles totales suivantes :

↔ A l'ordre 1 :

$$\Delta RW_1 \approx \frac{\partial RW}{\partial PD} \Delta PD + \frac{\partial RW}{\partial LGD} \Delta LGD + \frac{\partial RW}{\partial M} \Delta M$$

↔ A l'ordre 2 :

$$\begin{aligned} \Delta RW_2 \approx & \frac{\partial RW}{\partial PD} \Delta PD + \frac{\partial RW}{\partial LGD} \Delta LGD + \frac{\partial RW}{\partial M} \Delta M + \frac{1}{2} \cdot \frac{\partial^2 RW}{\partial PD^2} \Delta PD^2 \\ & + \frac{\partial^2 RW}{\partial PD \partial LGD} \Delta PD \cdot \Delta LGD + \frac{\partial^2 RW}{\partial PD \partial M} \Delta PD \cdot \Delta M + \frac{\partial^2 RW}{\partial LGD \partial M} \Delta LGD \cdot \Delta M \end{aligned}$$

$$\text{avec } \frac{\partial^2 RW}{\partial LGD^2} = \frac{\partial^2 RW}{\partial M^2} = 0$$



4.1.3 Effet marginal et contribution

Soient X et Y deux variables aléatoires.

4.1.3.1 Effet marginal

On entend par effet marginal de X sur Y , la variation de la variable Y induite par une « faible » perturbation de la variable X . Soit $EM_{X/Y}$ l'effet marginal de X sur Y .

Il est défini de la manière suivante :

$$EM_{X/Y} = \frac{\partial Y}{\partial X}$$

4.1.3.2 Contribution

Soit X_1 et X_2 deux variables aléatoires telles que $Y = f(X_1, X_2)$ où f est C^∞ . D'après la formule de Taylor, on a la relation suivante : $\Delta Y \approx \frac{\partial Y}{\partial X_1} \Delta X_1 + \frac{\partial Y}{\partial X_2} \Delta X_2$.

En employant les notations précédentes en référence aux effets marginaux, on réécrit la formule de Taylor de la manière suivante : $\Delta Y \approx EM_{X_1/Y} \cdot \Delta X_1 + EM_{X_2/Y} \cdot \Delta X_2$.

On peut alors définir la contribution des dérivées partielles ($C_{X/Y}$) comme :

$$C_{X/Y} = EM_{X/Y} \cdot \frac{\Delta X}{\Delta Y}$$

Suivant la formule de Taylor, la relation $C_{X_1/Y} + C_{X_2/Y} = 1$ est toujours vérifiée.

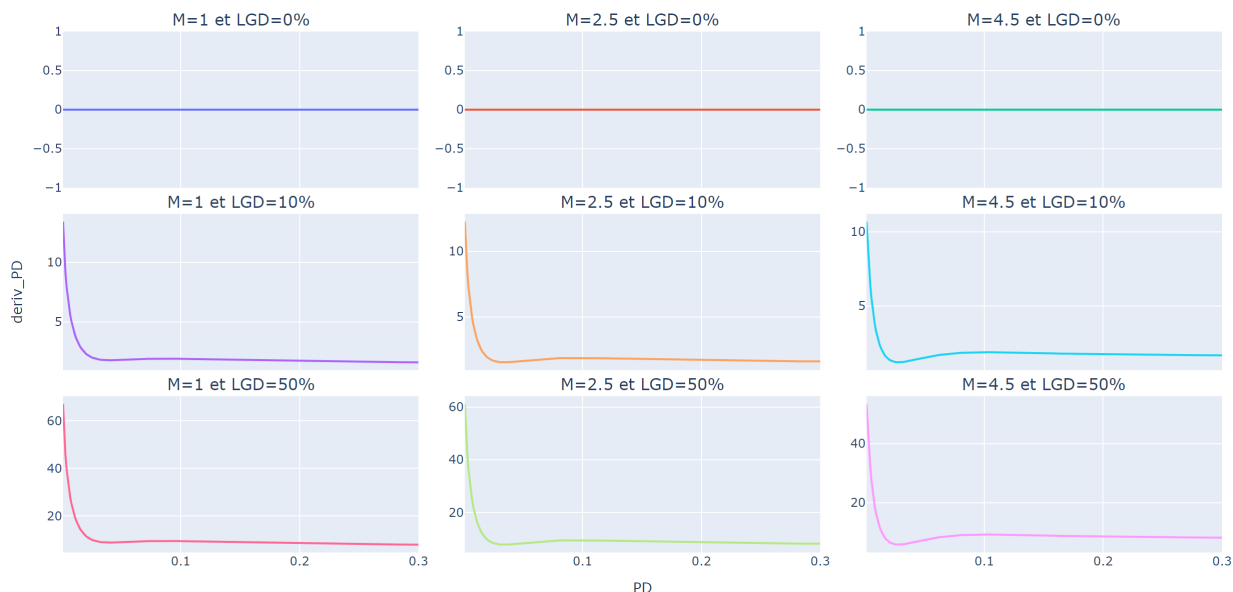
4.1.4 Application

Nous reprenons les données simulées qui ont servi à l'analyse de la fonction réglementaire. Dans cette section, nous appliquons les résultats précédents à l'estimation des effets marginaux des paramètres de la formule réglementaire. Dans un premier temps, nous présentons ces effets marginaux avant de faire une comparaison de la qualité d'approximation aux deux premiers ordres. Enfin, nous ferons une illustration de la contribution de chaque paramètre dans la variation du RW. Les formules de calcul des différents effets marginaux sont présentées en annexe.

L'effet marginal de la probabilité de défaut (PD) sur le RW ($\frac{\partial RW}{\partial PD}$) dépend des valeurs prises simultanément par la maturité et le taux de pertes en cas de défaut (LGD). Le graphique 4.1 montre clairement cette dépendance. De manière générale, on remarque que l'effet marginal de la PD décroît rapidement en fonction de la PD pour se stabiliser à un niveau assez faible. L'amplitude de variation est accentuée par des valeurs plus grandes de la maturité et de la perte en cas de défaut. L'effet d'une faible variation de la PD est d'autant plus fort que la valeur de PD originelle est faible. Ainsi, une dégradation de la qualité de crédit d'un emprunteur très bien noté présente un risque plus accru que la dégradation d'un emprunteur déjà mal noté.

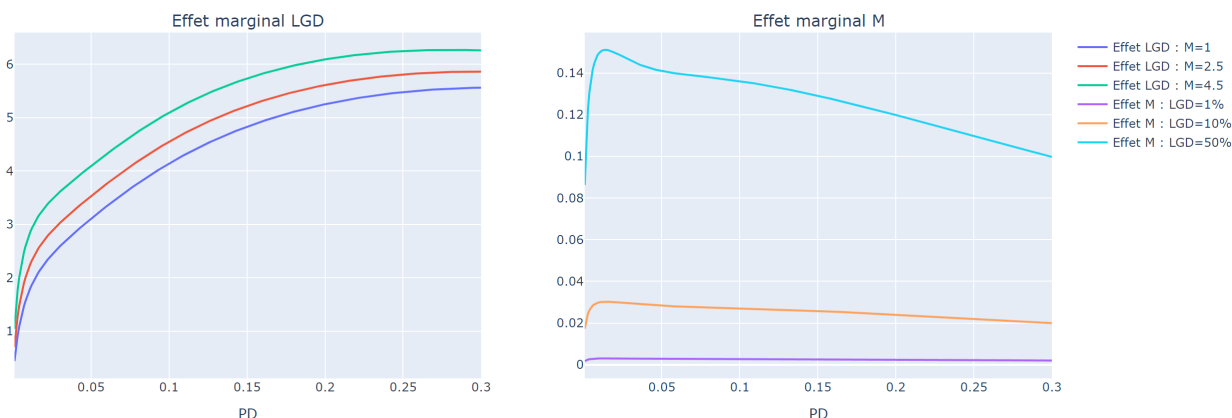


FIGURE 4.1 – Effet marginal de la PD



Sans surprise, l'effet marginal de la LGD ($\frac{\partial RW}{\partial LGD}$) reprend exactement la même allure que la courbe de RW sans prise en compte de la LGD présentée à la section précédente. L'effet marginal de la LGD est donc logarithmique. Une dégradation de la qualité de recouvrement d'un mauvais emprunteur¹ présente un risque plus accru que la dégradation de la qualité de recouvrement d'un bon emprunteur.

FIGURE 4.2 – Effet marginal de la LGD et de la maturité



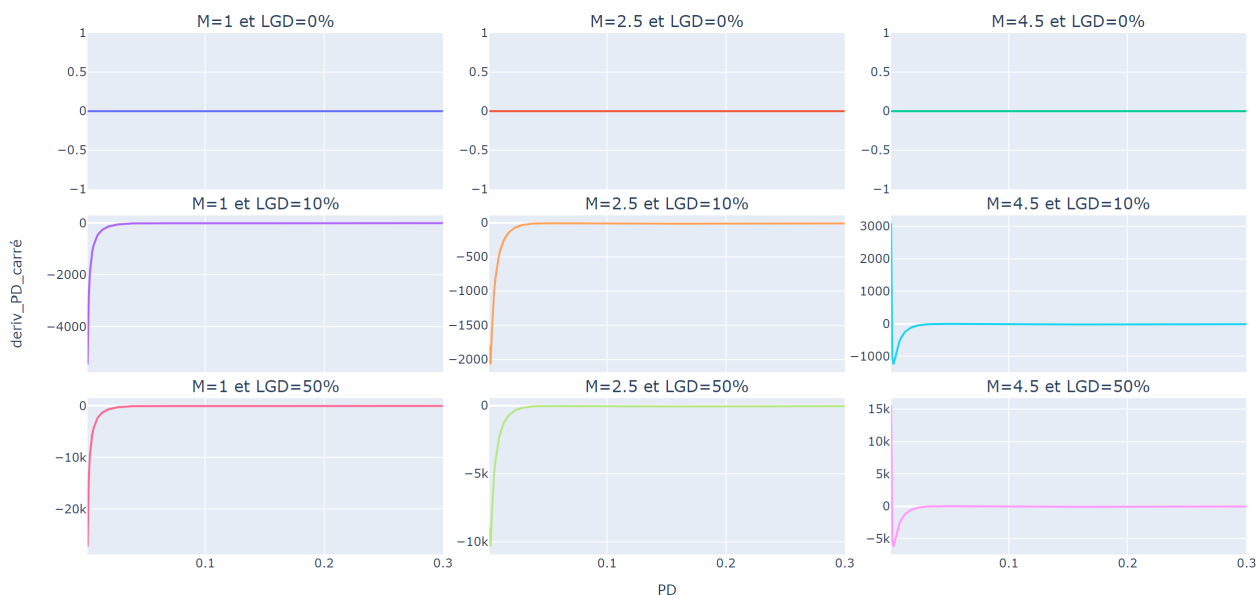
En ce qui concerne l'effet marginal de la maturité ($\frac{\partial RW}{\partial M}$), elle est fonction uniquement de la PD et de la LGD. On observe d'abord une croissance accélérée à mesure que la PD augmente avant d'avoir un effet inverse pour de fortes valeurs de la PD. La LGD influence l'amplitude de l'effet.

À l'ordre 2, on retrouve également l'effet du carré de la PD ($\frac{\partial^2 RW}{\partial PD^2}$) et les effets joints de la PD-LGD ($\frac{\partial^2 RW}{\partial PD \partial LGD}$) et PD-M ($\frac{\partial^2 RW}{\partial PD \partial M}$). Même si le carré de la PD a un effet négatif (d'atténuation) sur le RW, il s'estompe très rapidement.

1. Emprunteur mal noté.



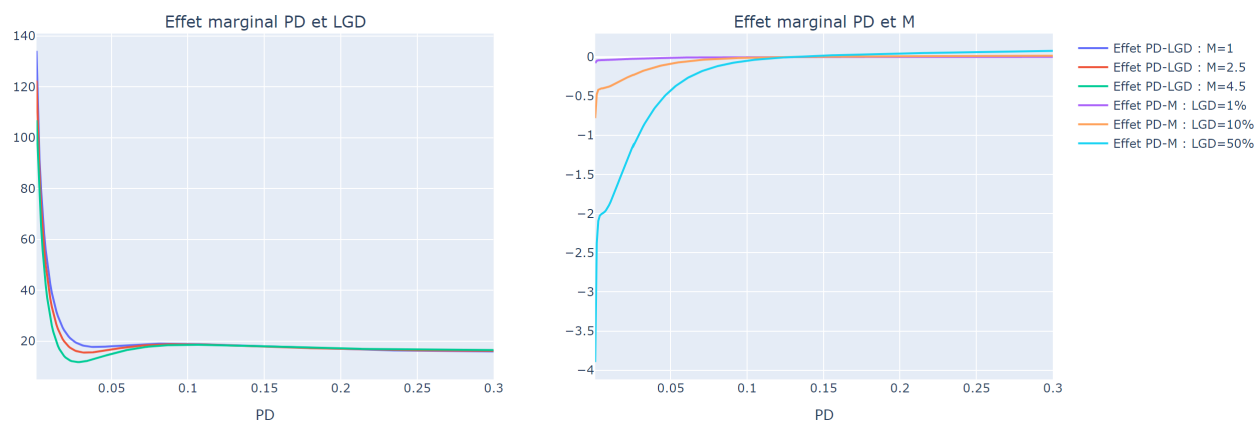
FIGURE 4.3 – Effet marginal de la PD^2



L'effet joint de la PD et de la LGD est positif et décroît rapidement jusqu'à une valeur approximative de 10% pour la PD. Le niveau minimum atteint par cet effet est accru par la valeur prise par la maturité. En effet, plus la maturité est grande, plus faible sera l'effet joint.

En ce qui concerne l'effet joint de la probabilité de défaut et de la maturité, il est négatif et croissant jusqu'à une valeur approchée de 10% pour la PD où il se stabilise.

FIGURE 4.4 – Effet marginal joint de la PD-LGD et PD-M



La question qui se pose à présent est l'intérêt de l'ajout des effets joints dans l'approximation de la formule réglementaire. Même si d'un point de vue théorique, l'ajout d'un ordre supplémentaire est censé améliorer les approximations, on se doit de le vérifier en pratique et d'en quantifier tout l'intérêt. On se sert de deux métriques (Root Mean Square Error - **RMSE** et Mean Absolute Error - **MAE**) pour quantifier l'erreur commise dans l'estimation de la variation de RW induite par une variation des paramètres.

Le tableau 4.1 présente les résultats pour des variations précises de PD (0,1%, 1% et 10%), de LGD (0,1%, 1%) et de maturité (0,1%, 1%).



TABLE 4.1 – Comparaison des approximations dans l’approche analytique

Variations	RMSE ordre 1	MAE ordre 1	RMSE ordre 2	MAE ordre 2
Delta / PD = 0.1%, LGD = 0.1%, M = 0.1%	0.0013	0.0289	0.0013	0.0289
Delta / PD = 1.0%, LGD = 0.1%, M = 0.1%	0.0130	0.0916	0.0129	0.0912
Delta / PD = 10.0%, LGD = 0.1%, M = 0.1%	0.1362	0.2966	0.1233	0.2821
Delta / PD = 0.1%, LGD = 1.0%, M = 0.1%	0.0013	0.0288	0.0013	0.0290
Delta / PD = 1.0%, LGD = 1.0%, M = 0.1%	0.0130	0.0915	0.0130	0.0916
Delta / PD = 10.0%, LGD = 1.0%, M = 0.1%	0.1360	0.2963	0.1245	0.2835
Delta / PD = 0.1%, LGD = 0.1%, M = 1.0%	0.0013	0.0289	0.0013	0.0289
Delta / PD = 1.0%, LGD = 0.1%, M = 1.0%	0.0130	0.0916	0.0129	0.0912
Delta / PD = 10.0%, LGD = 0.1%, M = 1.0%	0.1363	0.2967	0.1234	0.2822
Delta / PD = 0.1%, LGD = 1.0%, M = 1.0%	0.0013	0.0287	0.0013	0.0290
Delta / PD = 1.0%, LGD = 1.0%, M = 1.0%	0.0130	0.0915	0.0131	0.0916
Delta / PD = 10.0%, LGD = 1.0%, M = 1.0%	0.1361	0.2964	0.1247	0.2836

Les résultats révèlent par ailleurs que pour de fortes variations de la LGD, les approximations ne sont plus très précises. L’hypothèse formulée en amont qui nous permet l’application de la formule de Taylor est fortement remise en cause.

Enfin, l’approche analytique permet de répondre à la question suivante : « quelle est la contribution de chaque paramètre dans la variation du RW ? »

Les tableaux 4.2 et 4.3 suivants présentent respectivement les contributions pour des variations simultanées de tous les paramètres de 1% et de 25%. Ces contributions étant fonction des valeurs initiales de PD , LGD et M , on a choisi également des valeurs arbitraires des paramètres pour la mesure des contributions. On note clairement que la LGD est le paramètre qui contribue le plus à l’explication de la valeur du RW (~ 55% et 65%). Ensuite, pour de faibles valeurs de la PD, la maturité contribue plus. Cet effet est inversé pour des valeurs élevées de la PD. On remarque également que pour des variations fortes (25%), les effets joints sont beaucoup plus élevés.

TABLE 4.2 – Contribution des dérivées partielles pour une variation simultanée de 1% de tous les paramètres

Valeurs initiales				Variation du RW par approche analytique				Contribution des paramètres dans la variation du RW																	
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD2	contrib	PD	LGD	contrib	PD	M	contrib	LGD	M	
0.001	0.1	1.0	0.043978	0.000746	0.000749	17.905922	58.706219	23.025405	0.000179	0.179059	-0.010402	0.230254													
0.100	0.1	1.0	0.413990	0.006306	0.006320	29.967928	65.502928	4.307516	0.029968	0.299679	-0.001485	0.043075													
0.001	0.5	1.0	0.219891	0.003732	0.003746	17.905922	58.706219	23.025405	0.000179	0.179059	-0.010402	0.230254													
0.100	0.5	1.0	2.069952	0.031531	0.031601	29.967928	65.502928	4.307516	0.029968	0.299679	-0.001485	0.043075													
0.001	0.1	2.5	0.069852	0.001252	0.001257	9.737964	55.550662	34.293681	0.000097	0.097380	-0.015493	0.342937													
0.100	0.1	2.5	0.454827	0.007109	0.007126	26.380463	63.823837	9.550657	0.026380	0.263805	-0.003293	0.095507													
0.001	0.5	2.5	0.349258	0.006261	0.006287	9.737964	55.550662	34.293681	0.000097	0.097380	-0.015493	0.342937													
0.100	0.5	2.5	2.274135	0.035544	0.035631	26.380463	63.823837	9.550657	0.026380	0.263805	-0.003293	0.095507													

TABLE 4.3 – Contribution des dérivées partielles pour une variation simultanée de 25% de tous les paramètres

Valeurs initiales				Variation du RW par approche analytique				Contribution des paramètres dans la variation du RW																	
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD2	contrib	PD	LGD	contrib	PD	M	contrib	LGD	M	
0.001	0.1	1.0	0.043978	0.018660	0.020357	16.472957	54.008111	21.182740	0.004118	4.118239	-0.239240	5.295685													
0.100	0.1	1.0	0.413990	0.157654	0.166409	28.454419	62.194749	4.089968	0.711361	7.113605	-0.035256	1.022492													
0.001	0.5	1.0	0.219891	0.093301	0.101786	16.472957	54.008111	21.182740	0.004118	4.118239	-0.239240	5.295685													
0.100	0.5	1.0	2.069952	0.788272	0.832045	28.454419	62.194749	4.089968	0.711361	7.113605	-0.035256	1.022492													
0.001	0.1	2.5	0.069852	0.031305	0.034587	8.850710	50.489284	31.169087	0.002213	2.212678	-0.352026	7.792272													
0.100	0.1	2.5	0.454827	0.177721	0.188635	24.915191	60.278819	9.020177	0.622880	6.228798	-0.077756	2.255044													
0.001	0.5	2.5	0.349258	0.156523	0.172937	8.850710	50.489284	31.169087	0.002213	2.212678	-0.352026	7.792272													
0.100	0.5	2.5	2.274135	0.888603	0.943173	24.915191	60.278819	9.020177	0.622880	6.228798	-0.077756	2.255044													



4.2 APPROCHE PAR MODÉLISATION CLASSIQUE

4.2.1 Régression linéaire

4.2.1.1 Principe général

Le modèle de régression linéaire est l'un des modèles les plus simples en raison de sa construction. Ce modèle cherche à prédire une variable quantitative y à partir d'un ensemble de variables explicatives (qualitatives ou quantitatives). Soient $(x_j)_{1 \leq j \leq p}$, les p variables quantitatives supposées quantitatives. L'objectif du modèle linéaire est d'estimer les coefficients $(\beta_j)_{0 \leq j \leq p}$ les plus robustes tels que $\hat{y} = \beta_0 + \sum_{j=1}^p \beta_j x_j$ approxime au mieux la variable y . En considérant ε l'erreur d'approximation commise, le modèle se présente sous la forme $y = \hat{y} + \varepsilon = \beta_0 + \sum_{j=1}^p \beta_j x_j + \varepsilon$. Soient $Y = (y_i)_{1 \leq i \leq n}$ et $X = (x_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ les données sur n individus. Sous forme matricielle, le modèle se réécrit $Y = X\beta + \varepsilon$. La méthode des moindres carrés ordinaires (MCO) permet de trouver les coefficients $(\beta_j)_{0 \leq j \leq p}$ en minimisant l'erreur commise par le modèle.

$$\beta = \operatorname{argmin}_{(\beta_j)_{0 \leq j \leq p}} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2$$

On trouve finalement $\hat{\beta} = (X'X)^{-1} X'Y$.

4.2.1.2 Interprétation du modèle linéaire

Le modèle linéaire présente la qualité d'être simple et facilement interprétable. Cette propriété lui vaut une forte utilisation dans la littérature. Nous nous intéressons à l'interprétation dans le cas de variables explicatives quantitatives. Considérant la variable x_j , une augmentation d'une unité de cette variable entraîne une variation de β_j unité de l'estimation finale. Pour obtenir une interprétation en termes de pourcentage, une réécriture du modèle est nécessaire. En l'occurrence, en prenant le modèle définit comme suit : $\ln(y) = \beta_0 + \sum_{j=1}^p \beta_j \ln(x_j) + \varepsilon$, l'interprétation se fait en termes de pourcentage.

4.2.1.3 La régularisation

La régularisation consiste à ajouter un terme de pénalité aux valeurs fortes des paramètres afin d'éviter le surapprentissage. En imposant la nullité de certains coefficients, elle permet de rendre les modèles parcimonieux. La fonction de coût L à minimiser est l'erreur quadratique :

$$\forall (y, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^n, \quad L(y, \hat{y}) = \frac{1}{2n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{2n} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2$$



La régularisation ajoute un terme de correction dans la fonction de coût. La forme de ce terme de correction (proportionnel à la norme des coefficients) donne lieu à plusieurs formes de régularisation. On retrouve principalement le LASSO (Least Absolute Shrinkage and Selection Operator) avec la norme \mathbb{L}^1 et le RIDGE avec la norme \mathbb{L}^2 . La fonction de coût prend la forme :

- Régularisation \mathbb{L}^1 (Lasso) :

$$L(y, \hat{y}) = \frac{1}{2n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \|\beta\|_1 = \frac{1}{2n} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

- Régularisation \mathbb{L}^2 (Ridge) :

$$L(y, \hat{y}) = \frac{1}{2n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \|\beta\|_2 = \frac{1}{2n} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2$$

Le paramètre $\lambda \in \mathbb{R}$ est un paramètre de contrôle sur la force de la régularisation. Plus il sera grand, plus le modèle sera pénalisé et plus on aura de paramètres nuls. On se sert de la validation croisée pour déterminer la valeur optimale de ce paramètre.

4.2.1.4 Application sur les données simulées

Dans un premier temps, nous faisons appel au modèle linéaire pour l'approximation de la formule réglementaire. Ce modèle présente tout l'intérêt d'être simple et facilement interprétable. De plus, nous obtenons une nouvelle formule de calcul du RW. Les résultats sont contenus dans le tableau 4.4.

TABLE 4.4 – Résultats du modèle linéaire sur données simulées

Dep. Variable :	RW	R-squared :	0.932			
Model :	OLS	Adj. R-squared :	0.932			
No. Observations :	192000	Prob (F-statistic) :	0.00			
Df Model :	3	Log-Likelihood :	-98860			
AIC :	1.977e+05	BIC :	1.978e+05			
	coef	std err	t	P > t 	[0.025	0.975]
const	-1.2451	0.003	-377.603	0.000	-1.252	-1.239
PD	5.9778	0.011	560.168	0.000	5.957	5.999
LGD	4.8513	0.003	1515.317	0.000	4.845	4.858
M	0.1252	0.001	155.231	0.000	0.124	0.127

Le modèle précédent établit la relation suivante :

$$RW = 5.9778 \cdot PD + 4.8513 \cdot LGD + 0.1252 \cdot M - 1.2451$$



On en tire un ensemble de conclusions qui relatent les limites du modèle linéaire dans l'approximation de la formule réglementaire :

- Le signe négatif de la constante entraîne une prédiction du risque négative dans le cas où tous les paramètres seraient nuls. Ce résultat est incohérent avec la réalité que représente le RW.
- La LGD n'a à travers cette formule qu'un effet additif sur l'évaluation du risque. En effet, le modèle suggère qu'un niveau de perte en cas de défaut nul n'implique pas systématiquement que le risque encouru soit à son tour nul. L'état absorbant $\mathbf{0}$ de la LGD n'est pas retranscrit par ce modèle.
- Enfin, ce modèle suppose des effets constants des paramètres sur le risk weighted. Or, nos résultats précédents nous révèlent que les effets des paramètres sont fonction des valeurs de départ.

Ces limites font de ce modèle, une représentation erronée du risque de crédit même s'il présente une qualité d'ajustement très satisfaisante. En raison des valeurs prises par les paramètres, nous nous aidons d'une régression avec application de la fonction logarithmique ($\ln(RW) = \beta_0 + \beta_1 \ln(PD) + \beta_2 \ln(LGD) + \beta_3 \ln(M)$) pour quantifier aisément les effets fixes. Pour des raisons de stabilité, la régression a été effectuée sur les deals avec un RW strictement différent de 0.

TABLE 4.5 – Résultats du modèle linéaire (fonction logarithmique) sur données simulées

Dep. Variable :	ln(RW)	R-squared :	0.997			
Model :	OLS	Adj. R-squared :	0.997			
No. Observations :	190082	Prob (F-statistic) :	0.00			
Df Model :	3	Log-Likelihood :	2.8913e+05			
AIC :	-5.783e+05	BIC :	-5.782e+05			
	coef	std err	t	P > t 	[0.025	0.975]
const	2.1128	0.000	5300.288	0.000	2.112	2.114
ln(PD)	0.3195	0.000	2540.784	0.000	0.319	0.320
ln(LGD)	1.0000	0.000	7612.757	0.000	1.000	1.000
ln(M)	0.1380	0.000	549.369	0.000	0.137	0.138

Sans surprise, une variation d'1% de la LGD entraîne une variation similaire du RW. Le caractère multiplicatif de la LGD est bien retranscrit. En ce qui concerne la probabilité de défaut, une variation de celle-ci de 1% engendre une variation moins importante de 0,32% du RW. On retient que la LGD a un effet trois fois plus important que celui de la PD sur le risque.



4.2.2 Régression polynomiale

4.2.2.1 Principe

Un modèle de régression polynomiale est une extension d'un modèle de régression classique $Y = f(X) + \varepsilon$. Dans le cas de la régression polynomiale, on suppose que la fonction f est de la forme :

$$f(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_m x^m \text{ avec } m \geq 2$$

C'est une approche statistique qui vise à modéliser des relations de forme non linéaire entre $Y \in \mathbb{R}$ et les variables explicatives $X \in \mathbb{R}^p$. On parle de modèles de régression polynomiale car ils intègrent des polynômes dans les équations. Ces modèles appartiennent toujours à la famille des modèles linéaires étant donné qu'ils sont linéaires en les paramètres du modèle $(\beta_0, \dots, \beta_p)$.

4.2.2.2 Application sur les données simulées

Les résultats du modèle linéaire précédent nous ont permis de mettre en exergue trois limites. Le passage à la régression polynomiale permet de résoudre l'une de ces limites. Il s'agit des effets marginaux des paramètres. En particulier, l'effet marginal de la probabilité de défaut à l'aide du modèle polynomial de degré 2 est donné par : **9.7468 – 61 · PD**. D'autre part, les modèles polynomiaux permettent d'avoir une estimation des effets joints.

TABLE 4.6 – Résultats du modèle polynomial (degré 2) sur données simulées

Dep. Variable :	RW	R-squared :	0.991			
Model :	OLS	Adj. R-squared :	0.991			
No. Observations :	192000	Prob (F-statistic) :	0.00			
Df Model :	9	Log-Likelihood :	91963.			
AIC :	-1.839e+05	BIC :	-1.838e+05			
	coef	std err	t	P > t 	[0.025	0.975]
const	-0.5333	0.003	-159.570	0.000	-0.540	-0.527
PD	9.7468	0.020	495.686	0.000	9.708	9.785
LGD	2.3411	0.006	398.757	0.000	2.330	2.353
M	0.0233	0.002	12.797	0.000	0.020	0.027
PD²	-31.0034	0.051	-608.322	0.000	-31.103	-30.904
PD · LGD	12.0648	0.014	882.066	0.000	12.038	12.092
PD · M	-0.1502	0.003	-43.543	0.000	-0.157	-0.143
LGD²	-0.0017	0.005	-0.360	0.719	-0.011	0.007
LGD · M	0.2527	0.001	244.190	0.000	0.251	0.255
M²	-0.0001	0.000	-0.362	0.718	-0.001	0.000

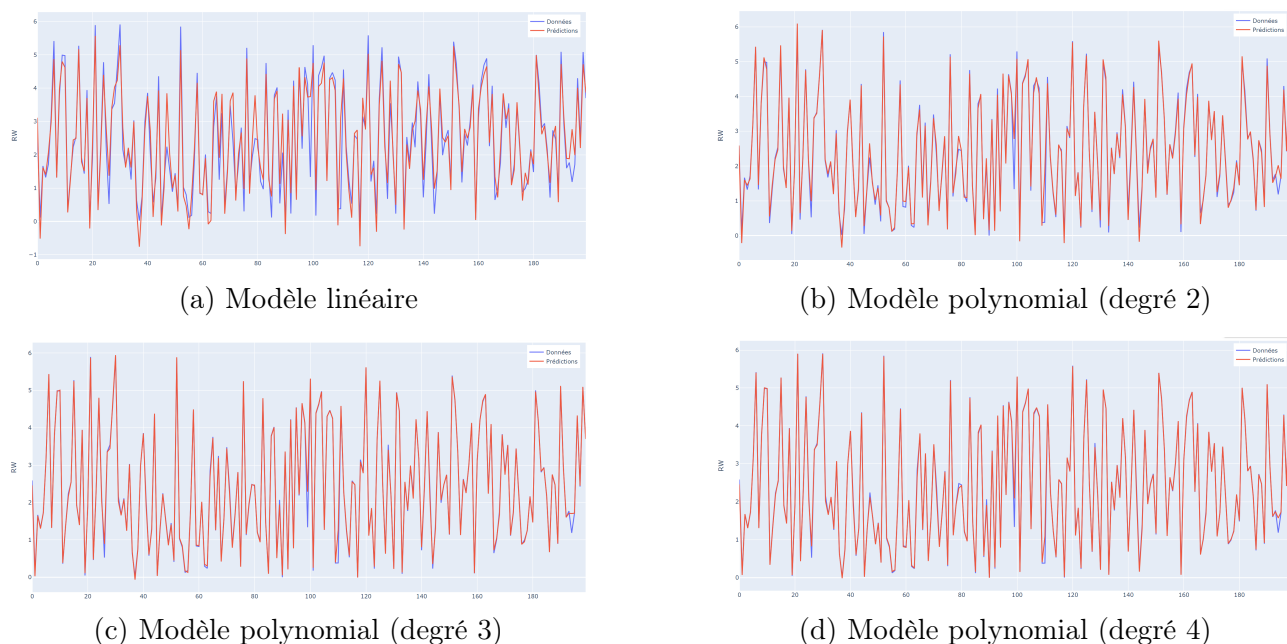
4.2.3 Discussion des résultats et extensions

Un tour d'horizon des résultats précédents sur le modèle linéaire et les modèles polynomiaux (en annexe) montrent que la qualité d'ajustement est fonction croissante du nombre de degrés



retenu. Toutefois, l'interprétabilité du modèle en pâtit. De plus, la forme de ces modèles entraînent de la multicolinéarité et donc du surapprentissage.

FIGURE 4.5 – Prédiction du RW par modélisation classique



Les graphiques montrent clairement que les modèles ajustent assez bien les données. Toutefois, ces modèles souffrent d'une limite majeure : **la prédiction de RW négative**.

4.3 APPROCHE PAR MACHINE LEARNING

4.3.1 Random Forest

4.3.1.1 Présentation

Les forêts aléatoires (Random Forest) est un modèle d'ensemble inspiré des arbres de décision. Introduit par Breiman, les Random Forest sont une extension des méthodes de Bagging. Le bagging est une méthode de réduction de variance dans l'estimation d'un modèle de prédiction. Il crée B sous-ensembles d'entraînement à l'aide d'un échantillonnage Bootstrap et construit des arbres sur chacun des échantillons aléatoires. La prédiction issue de cette méthode est :

$$\hat{y} = \frac{1}{B} \sum_{b=1}^B \hat{f}_b(x)$$

La méthode de Bootstrap permet d'obtenir un modèle qui réduit la variance sans pour autant accroître le biais.

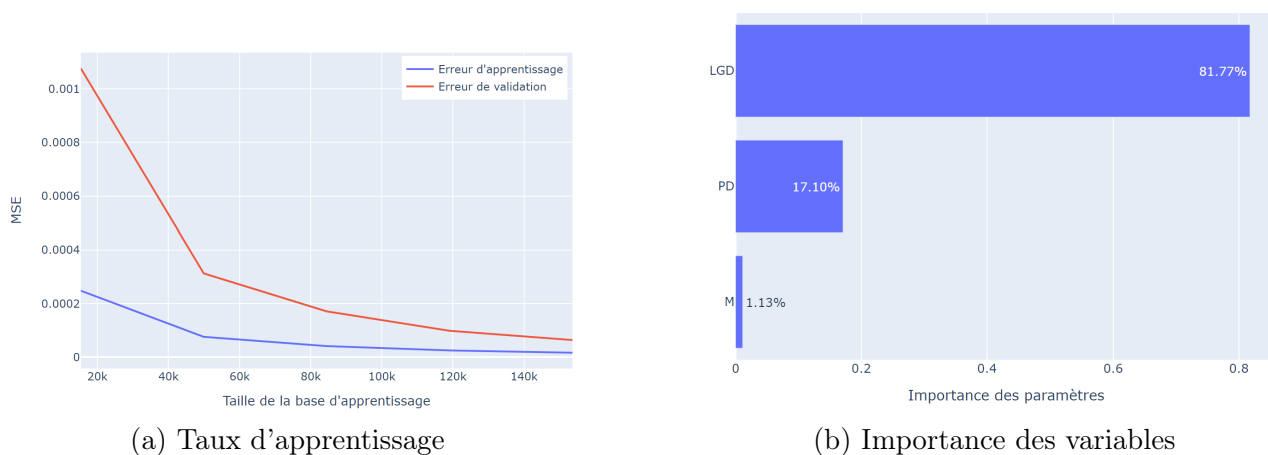
$$\sigma_f = \sqrt{\frac{\sum_{b=1}^B (\hat{f}_b(x) - \hat{y})^2}{B - 1}}$$



4.3.1.2 Application sur les données simulées

On ne remarque pas de surapprentissage du Random Forest à partir du graphique 4.6. Ce dernier suggère une décroissance continue de l'erreur d'apprentissage en fonction de la taille des données. Notre modèle ne se calque pas sur des caractéristiques propres de la base d'apprentissage. Il semble donc robuste pour une analyse plus ouverte. De plus, les valeurs prises par l'erreur de prédiction moyenne (très proches de 0) montrent que le Random Forest est une alternative judicieuse pour l'approximation de la formule réglementaire. En effet, dans la pratique, la mise en place d'échelle de notation des tiers ainsi que de taux de pertes groupés rend l'ensemble des valeurs possibles de ces paramètres limité. Un modèle fondé sur l'arbre de décision peut alors fournir des résultats très intéressants.

FIGURE 4.6 – Random Forest sur données simulées



Ces conclusions (en concordance avec les résultats préliminaires) placent la LGD comme contributeur à 82% de la prédiction faite. Ce qui correspond à une importance relative 4,8 fois plus importante que la probabilité de défaut.

4.3.2 Boosting : Gradient boosting et XGBoost

4.3.2.1 Principe et algorithme de descente de gradient

Le boosting est une technique ensembliste qui agrège des classifieurs faibles construits étape par étape tout en corrigeant les poids affectés aux individus. Pour optimiser les poids affectés, les méthodes de boosting se basent sur l'algorithme de descente de gradient. Il s'agit d'une technique itérative qui permet d'approcher la solution d'un problème d'optimisation convexe. Soit une fonction $f : \mathbb{R}^p \rightarrow \mathbb{R}$ deux fois différentiable qu'on souhaite minimiser. Pour $x \in \mathbb{R}^p$, on définit le gradient de f en x comme :

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_j}(x) \right)_{1 \leq j \leq p}$$



L'algorithme part d'une valeur initiale $x_{(0)}$ et calcule de manière itérative les $x_{(1)}, x_{(2)}, \dots$ jusqu'à l'optimum. Ainsi, avec un seuil $\epsilon > 0$, on définit $x_{(k+1)} = x_{(k)} - \alpha_k \nabla f(x_{(k)})$ tant que $\|\nabla f(x_{(k)})\| > \epsilon$. Dans notre cas, nous cherchons à minimiser la fonction de coût du modèle. Notons $l(y^{(i)}, f(x^{(i)}))$, la fonction de coût utilisée. On cherche à minimiser $L(y, f) = \sum_{i=1}^n l(y^{(i)}, f(x^{(i)}))$ par rapport aux paramètres de la fonction f à l'aide de l'algorithme de descente de gradient. A l'étape k de l'algorithme, on obtient la formule :

$$f_k(x) = f_{k-1}(x) - \alpha \nabla L(y, f)$$

$$f_k(x_i) = f_{k-1}(x^{(i)}) - \alpha \frac{\partial l(y^{(i)}, f(x^{(i)}))}{\partial f(x^{(i)})}, \quad 1 \leq i \leq n$$

4.3.2.2 Algorithme de Gradient Boosting

Nous présentons l'algorithme de Gradient Boosting dans le cas de la régression. L'algorithme suit ces étapes :

- initialisation : $f_0(x) = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^n L(y^{(i)}, \theta)$
- pour k allant de 1 à K :
 - calcul des pseudos résidus : $r_k^{(i)} = - \left. \frac{\partial L(y^{(i)}, f(x^{(i)}))}{\partial f(x^{(i)})} \right|_{f(x)=f_{k-1}(x)}$, $1 \leq i \leq n$
 - ajustement d'un classifieur $h_k(x)$
 - calcul du prochain classifieur : $f_k(x) = f_{k-1}(x) + \theta_k h_k(x)$
avec $\theta_k := \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^n L(y^{(i)}, f_{k-1}(x^{(i)}) + \theta h_k(x^{(i)}))$
- en sortie, on obtient f_K qui est le modèle final

4.3.2.3 XGBoost

Le modèle XGBoost est une optimisation de l'algorithme du Tree Gradient Boosting². C'est un algorithme d'apprentissage supervisé qui combine de manière séquentielle plusieurs algorithmes pour une meilleure prédiction de la variable cible. Le caractère séquentiel de l'algorithme le différencie des forêts aléatoires. L'algorithme du XGBoost applique une régularisation à la fonction de coût pour tenir compte de la complexité du modèle. A chaque itération, le nouveau modèle prédit la partie résiduelle de l'itération précédente. La particularité du XGBoost est la réduction des temps de calcul dus à l'utilisation d'un grand nombre de variables.

$$\hat{y}_b^{(i)} = \sum_{k=0}^b \hat{r}_k^{(i)} = \hat{y}_{b-1}^{(i)} + \hat{r}_b^{(i)} \quad \text{et} \quad \hat{y}_0^{(i)} = \hat{h}(x^{(i)}), \quad 1 \leq b \leq B, \quad 1 \leq i \leq n$$

2. Cas particulier du Gradient Boosting où les modèles h_k sont des arbres de décisions.

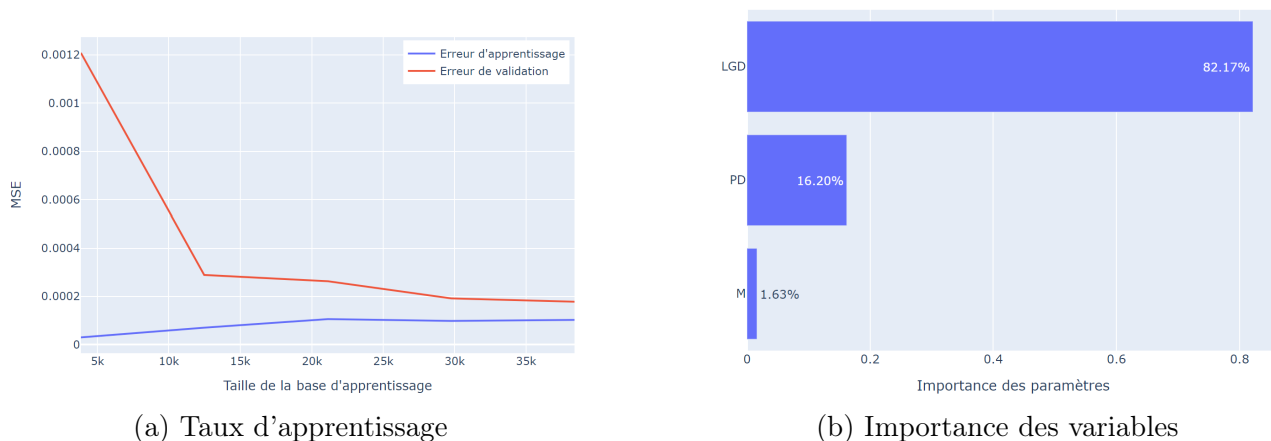


4.3.2.4 Application sur données simulées

Le XGBoost présente certaines caractéristiques semblables au Random Forest. Ces caractéristiques en font un prétendant judicieux à l'approximation de la formule réglementaire.

Les résultats sont présentés par le graphique 4.7. Les contributions de chacun des paramètres sont équivalentes au Random Forest.

FIGURE 4.7 – XGBoost sur données simulées



4.4 APPROCHE PAR DEEP LEARNING : RÉSEAU DE NEURONES

4.4.1 Présentation et formalisme mathématique

Un réseau de neurones est une combinaison de neurones artificiels³ dans le but de simuler l'action d'un neurone biologique. Un "neurone formel" (ou simplement "neurone") est une fonction algébrique non linéaire et bornée, dont la valeur dépend de paramètres appelés coefficients ou poids.

4.4.1.1 Perceptron simple

Si on considère un vecteur de données d'entrée $x = (x_1, \dots, x_p) \in \mathbb{R}^p$, un vecteur de poids $(w_1, \dots, w_p) \in \mathbb{R}^p$. Le perceptron simple renvoie en sortie $y(x) = G(w_0 + \sum_{i=1}^p w_i x_i)$ avec $G(\cdot)$ la fonction d'activation. Dans le cas de la régression, la fonction d'activation est généralement l'identité. Nous retenons $\forall u \in \mathbb{R}, G(u) = u$. Ainsi, on retient comme sortie du perceptron simple dans le cas de la régression :

$$f(x) = w_0 + \sum_{i=1}^p w_i x_i$$

3. Fonction algébrique non linéaire dont la valeur dépend de paramètres appelés poids.



4.4.1.2 Apprentissage d'un perceptron

La question qui se pose maintenant que l'on connaît la structure d'un perceptron simple est le choix optimal des poids w_1, \dots, w_p . On reprend les notations précédentes pour désigner les n observations $(x^{(i)}, y^{(i)})$. L'idée adoptée dans le cadre du neurone est le fait de réajuster les poids à chaque nouvelle observation. Cette idée se base sur l'algorithme de descente de gradient avec une fonction de coût L .

En partant de poids aléatoires $(w_j^{(0)})_{0 \leq j \leq p}$, à chaque nouvelle itération, les nouvelles sont définies de la manière suivante :

$$\begin{aligned} w_j^{(i+1)} &= w_j^{(i)} - \eta \frac{\partial L}{\partial w_j} (y^{(i)}, f(x^{(i)})) \\ &= w_j^{(i)} - \eta \frac{\partial L}{\partial w_j} \left(y^{(i)}, G \left(\sum_{k=1}^p w_k x_k^{(i)} + w_0 \right) \right), 0 \leq j \leq p \end{aligned}$$

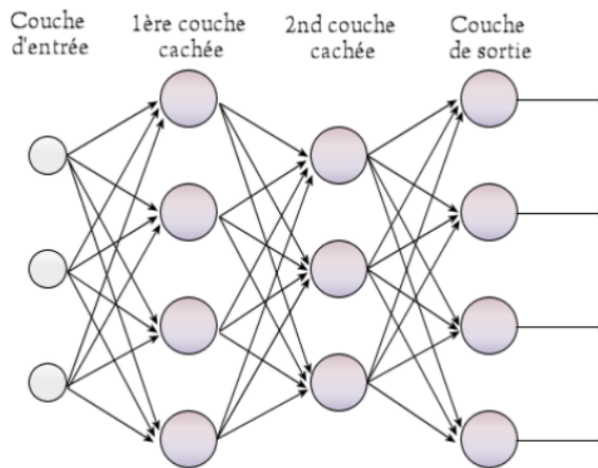
Si on considère la fonction de coût $L(y, y') = \frac{1}{2}(y - y')^2$, l'équation précédente se réécrit :

$$w_j^{(i+1)} = w_j^{(i)} - \eta (y^{(i)} - f(x^{(i)})) x_j^{(i)}, \quad 0 \leq j \leq p$$

4.4.1.3 Perceptron multicouche

L'appellation réseau de neurones vient du fait qu'on empile plusieurs couches intermédiaires appelées couches cachées entre les couches d'entrée et de sortie.

FIGURE 4.8 – Architecture d'un perceptron multicouche



Source : DELCAILLAU 2019

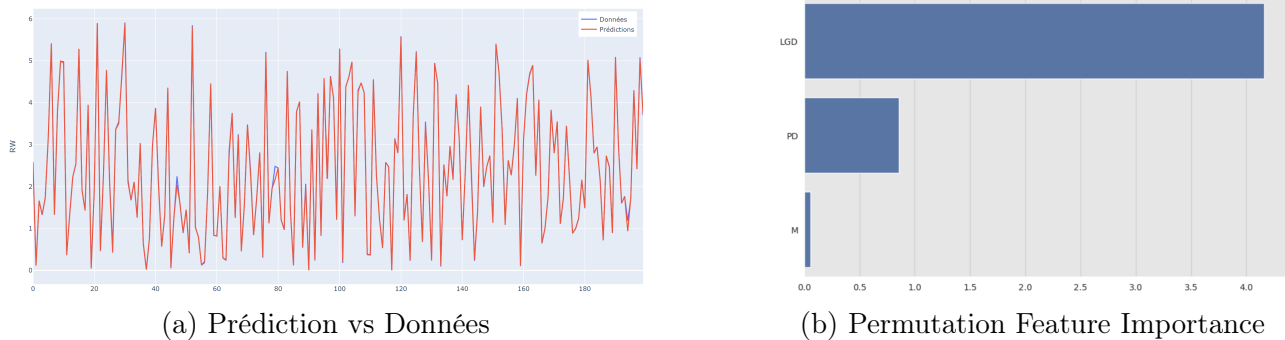
4.4.2 Application aux données simulées

L'utilisation des réseaux de neurones pour la résolution de notre problématique a pour but d'en faire un benchmark des modèles d'arbres construits précédemment. Nous utilisons ici une version de l'importance des variables qui est présentée dans le chapitre suivant. Son



interprétation en effet relatif est similaire à celles vues plus haut.

FIGURE 4.9 – Réseau de neurones sur données simulées



(a) Prédiction vs Données

(b) Permutation Feature Importance

Les résultats obtenus nous confirment que les effets observés sont vraisemblables. En effet, la LGD a un effet 4,7 fois plus important que la PD.

En résumé

Pour analyser les effets des paramètres de la formule réglementaire d'évaluation du risque de crédit, deux principales visions sont possibles. La première, "analytique" se sert de la théorie mathématique (formule de Taylor) pour proposer une formule fermée des effets marginaux. Les résultats obtenus sur données théoriques reflètent bien les propriétés que doit posséder une formule d'évaluation de risque cohérente. On retient que la LGD explique à plus de 50% les variations du RW. Toutefois, cette vision reposant sur l'hypothèse de variation infime des paramètres, elle peut être limitée dans son application. Pour répondre à cette problématique, la deuxième vision, "modélisation alternative" du risque de crédit est utile. Le modèle d'approximation retenu se doit d'avoir une qualité d'ajustement satisfaisante, de bonnes propriétés d'interprétation et une reproductibilité des principales caractéristiques d'une mesure de risque cohérente. Ainsi, une application du modèle linéaire classique révèle qu'il n'est pas un candidat adéquat étant donné qu'il prédit des valeurs négatives de RW. Le prochain chapitre sert à la sélection d'un modèle alternatif d'évaluation du risque de crédit compte tenu de certains points d'attention.

INTERPRÉTABILITÉ DES MODÈLES ET APPLICATIONS

Dans ce chapitre, nous présentons les outils d'interprétation des modèles de boîtes noires et nous les appliquons à des données réelles issues du portefeuille de la Société Générale.

5.1	Qualité d'ajustement des modèles	57
5.1.1	Le biais	57
5.1.2	Critère RMSE (Root Mean Square Error)	57
5.1.3	Critère MAE (Mean Absolute Error)	57
5.1.4	Indicateurs normalisés	58
5.1.5	Le R^2 et le R^2 ajusté	58
5.1.6	Application aux données simulées	59
5.2	Méthodes d'interprétation des modèles de boîtes noires	59
5.2.1	Graphique de dépendance partielle (PDP)	59
5.2.2	Graphique des effets locaux accumulés (ALE)	61
5.2.3	Importance des variables	62
5.2.4	Interaction entre les variables	64
5.2.5	Modèles de substitution locaux : LIME	65
5.2.6	SHAP	66
5.3	Application à des données réelles : échantillon du portefeuille de la Société Générale	67
5.3.1	Base de travail et note sur la confidentialité	67
5.3.2	Analyse statistique succincte de la base de travail	68
5.3.3	Approche analytique	70
5.3.4	Modélisation	72



5.1 QUALITÉ D'AJUSTEMENT DES MODÈLES

Lors de la sélection d'un modèle, il se pose alors la question de sa fiabilité et de sa pertinence. En effet, un modèle a pour but de se rapprocher le plus possible du phénomène que l'on souhaite étudier. Pour ce faire, il est nécessaire d'utiliser des indicateurs de qualité permettant de classer les modèles candidats. Nous présentons dans la suite les indicateurs les plus connus pour l'évaluation de la qualité d'ajustement d'un modèle.

5.1.1 Le biais

L'un des premiers indicateurs est le **biais**. Il s'agit de la moyenne arithmétique simple de tous les écarts. Certains modèles comme la régression linéaire imposent que le modèle soit sans biais dans leur forme la plus classique. C'est un critère qui renseigne si le modèle a tendance à sous-estimer ou surestimer systématiquement les prédictions.

$$bias = \frac{1}{n} \sum_{i=1}^n \left(y^{(i)} - \hat{f}(x^{(i)}) \right)$$

5.1.2 Critère RMSE (Root Mean Square Error)

De par sa construction, le biais souffre d'un défaut majeur, il somme des valeurs algébriques. Ainsi, il ne renseigne pas sur l'amplitude des écarts et donc la capacité du modèle à se rapprocher des vraies valeurs. Pour arriver à cette fin, on se sert du **Root Mean Square Error (RMSE)**.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(y^{(i)} - \hat{f}(x^{(i)}) \right)^2}$$

Cet indicateur possède une version sans l'utilisation de la racine carrée, il s'agit du **Mean Square Error (MSE)**.

$$MSE = \frac{1}{n} \sum_{i=1}^n \left(y^{(i)} - \hat{f}(x^{(i)}) \right)^2$$

Plus la valeur du RMSE (MSE) est proche de zéro, meilleur est le modèle en termes d'exactitude.

5.1.3 Critère MAE (Mean Absolute Error)

L'erreur absolue moyenne (MAE) est un indicateur très utilisé dans la théorie pour évaluer la qualité d'un modèle. Cet indicateur est considéré comme étant plus fidèle vu qu'il est moins influencé par les erreurs de prédictions extrêmes. De plus, n'étant pas une fonction strictement croissante de la taille de l'échantillon, le MAE est un outil complémentaire au RMSE.

$$MAE = \frac{1}{n} \sum_{i=1}^n \left| y^{(i)} - \hat{f}(x^{(i)}) \right|$$



5.1.4 Indicateurs normalisés

5.1.4.1 Critère de Nash-Sutcliffe

Le critère de Nash-Sutcliffe est un indicateur de performance qui permet d'estimer la capacité du modèle à reproduire un comportement observé. Il a la particularité d'être un indicateur normalisé et donc d'avoir une interprétation relative aux données.

$$NS = 1 - \frac{MSE}{\sigma_y^2} = 1 - \frac{\sum_{i=1}^n (y^{(i)} - \hat{f}(x^{(i)}))^2}{\sum_{i=1}^n (y^{(i)} - \bar{y})^2}$$

Plus il est proche de 1, meilleure est l'adéquation compte tenu des données. En général, une valeur supérieure à 0,7 est suffisante pour considérer le modèle comme satisfaisant.

5.1.4.2 Critère RSR

Le critère **RMSE standard deviation ratio (RSR)** est un indicateur similaire au critère de Nash-Sutcliffe mais au contraire basé sur le RMSE.

$$RSR = \frac{RMSE}{\sigma_y} = \sqrt{\frac{\sum_{i=1}^n (y^{(i)} - \hat{f}(x^{(i)}))^2}{\sum_{i=1}^n (y^{(i)} - \bar{y})^2}}$$

Sa construction rend son interprétation plus aisée. Prenant des valeurs strictement positives, plus il est proche de 0, plus il est admis que le modèle ajuste bien les données.

5.1.5 Le R^2 et le R^2 ajusté

Cet indicateur mesure la part de variabilité expliquée par le modèle. Pour le calculer, on se base sur la somme des carrés des résidus (SSE) et la somme totale des carrés (SST) :

$$SSE = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \text{et} \quad SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad \text{avec} \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

A l'aide de ces grandeurs, on détermine le coefficient R^2 appelé coefficient de détermination de la manière suivante :

$$R^2 = 1 - \frac{SSE}{SST} = \frac{SST - SSE}{SST}, \quad 0 \leq R^2 \leq 1$$

Plus la valeur du R^2 est élevée, plus le modèle ajuste correctement les données. De par sa définition, cet indicateur est croissant par rapport au nombre de paramètres du modèle. Il suffit de continuer à rajouter des variables (même ne contenant aucune information) pour continuer à accroître le R^2 . Ce constat en fait un indicateur limité pour l'évaluation de la qualité d'un modèle. On fera plus souvent appel au R^2 ajusté (\tilde{R}^2) qui pénalise les modèles avec beaucoup



de paramètres. Il est défini par :

$$\tilde{R}^2 = R^2 - (1 - R^2) \frac{p}{n - p - 1}$$

5.1.6 Application aux données simulées

Nous reprenons ici les différents modèles construits dans le chapitre précédent. Nous comparons leur qualité d'ajustement à l'aide des indicateurs présentés précédemment. On remarque dans un premier temps que tous les modèles présentent un fort niveau d'ajustement aux données. En l'occurrence, les deux modèles de machine learning utilisés obtiennent un coefficient de détermination de **1**. Compte tenu de l'ordre de grandeur des erreurs, nous nous référons en priorité au MAE pour le choix du meilleur modèle. Le **Random Forest** présente les meilleurs résultats hormis le biais qu'il engendre. On conclut qu'il engendre une sous-évaluation globale du risque par rapport à la formule réglementaire. En dehors du XGBoost et des modèles polynomiaux d'ordres supérieurs à 2, tous les autres modèles partagent cette même dernière particularité.

TABLE 5.1 – Comparaison des modèles sur données simulées

	Modèle	Biais	RMSE	MAE	RSR	R2
<i>Modèles classiques</i>	Modèle linéaire	0.0031	0.4036	0.2973	0.259615	0.9326
	Modèle polynomial (degré 2)	0.0006	0.1486	0.0996	0.095394	0.9909
	Modèle polynomial (degré 3)	-0.0001	0.0700	0.0352	0.044721	0.9980
	Modèle polynomial (degré 4)	0.0000	0.0570	0.0279	0.036056	0.9987
<i>Modèles de ML</i>	Random Forest	0.0007	0.0063	0.0036	0.000000	1.0000
	XGBoost	0.0000	0.0110	0.0079	0.000000	1.0000
<i>Modèle de DeepL</i>	Réseaux de neurones	0.0061	0.0376	0.0178	0.024495	0.9994

Avant l'application sur les données de la Société Générale, nous retiendrons le Random Forest comme modèle approximatif de la formule réglementaire pour les illustrations des méthodes d'interprétation des modèles de boîte noire.

5.2 MÉTHODES D'INTERPRÉTATION DES MODÈLES DE BOITES NOIRES

5.2.1 Graphique de dépendance partielle (PDP)

5.2.1.1 Principe général

Le *Partial Dependence Plot* (PDP) introduit par Friedman en 2001, est une méthode d'interprétation des modèles de boîte noire qui a pour objectif de montrer l'effet marginal d'une ou plusieurs variables explicatives sur la prédiction faite par le modèle. Considérons une base d'apprentissage $Z^{(i)} = (X^{(i)}, Y^{(i)}) \in \mathbb{R}^p \times \mathbb{R}$ pour $1 \leq i \leq n$ et \hat{f} le modèle entraîné.



On considère une partition des variables en deux ensembles : $X_S = (X_1, X_2)$ les variables dont on souhaite connaître l'effet et $X_C = (X_3, \dots, X_p)$ les variables restantes. La fonction de dépendance partielle est définie par la formule :

$$\hat{f}_{x_S}(x_S) = \mathbb{E}_{X_C} \left[\hat{f}(x_S, X_C) \right] = \int \hat{f}(x_S, x_C) d\mathbb{P}_{X_C}(x_C)$$

Il faut noter que cette expression diffère de l'espérance conditionnelle de X_S obtenue par Monte-Carlo de la manière suivante :

$$\hat{f}_{x_S}(x_S) \simeq \frac{1}{n} \sum_{i=1}^n \hat{f}(x_S, x_C^{(i)})$$

5.2.1.2 Algorithme d'estimation de la PDP

La méthode d'estimation de la PDP repose sur la non-corrélation entre les variables X_C et X_S . La procédure d'estimation est donnée par l'algorithme suivant :

- Input : base d'apprentissage $(x_j^{(i)})_{1 \leq i \leq n, 1 \leq j \leq p}$, le modèle \hat{f} , la variable à expliquer (x_1)
- Pour i allant de 1 à n :
 - copie de la base d'apprentissage en remplaçant x_1 par $x_1^{(i)}$
 - calcul de $\hat{f}(x_1^{(i)}, x_C^{(k)})$ pour $k = 1, \dots, n$
 - calcul de $\hat{f}_{x_1}(x_1^{(i)}) \simeq \frac{1}{n} \sum_{k=1}^n \hat{f}(x_1^{(i)}, x_C^{(k)})$
- Output : le graphique des points $(x_1^{(i)}, \hat{f}_{x_1}(x_1^{(i)}))$ pour $i = 1, \dots, n$ appelé graphique de dépendance partielle (PDP)

5.2.1.3 Avantages et inconvénients de la PDP

Le principal avantage du PDP est sa simplicité d'interprétation et d'implémentation. De plus, par ce graphique, on peut estimer l'importance des variables et les interactions entre elles. Toutefois, ce graphique repose sur l'hypothèse d'indépendance entre les variables nécessaires au calcul de la moyenne. Cette hypothèse n'étant pas nécessairement vérifiée en pratique, pose un réel problème dans la fiabilité des résultats obtenus. Un autre problème du PDP est le fait de masquer les effets hétérogènes.

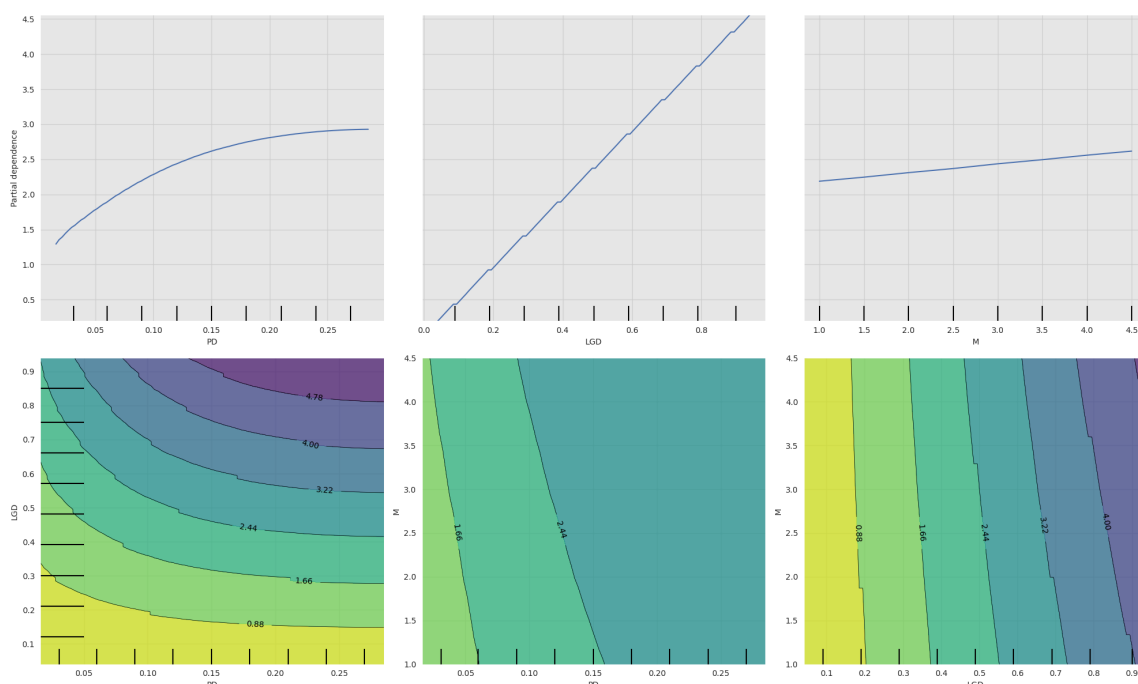
5.2.1.4 Application au Random Forest

Nos données simulées présentent la particularité d'avoir une corrélation nulle entre les différentes variables (voir figure A.1, page XII). Cette propriété rend robuste l'application de la PDP pour l'estimation des effets marginaux moyens. Le graphique 5.1 présente les résultats attendus. L'effet de la probabilité de défaut est logarithmique. En effet, plus la PD est élevée, moins grande est l'impact de sa variation sur la variation du RW tous les autres paramètres



supposés constants. Ce résultat rejoint les résultats obtenus avec l'approche analytique. Aussi, on retrouve que les effets marginaux de la LGD et de la maturité sont linéaires et croissants.

FIGURE 5.1 – Partial Dependence Plot sur données simulées (Random Forest)



5.2.2 Graphique des effets locaux accumulés (ALE)

5.2.2.1 Principe général et formalise mathématique

Le graphique des ALE (*Accumulated Local Effects*) vient en correction du PDP dans le cas de variables explicatives corrélées. L'ALE décrit l'influence des variables sur la prédiction du modèle en moyenne. Pour neutraliser l'effet de la corrélation, cette méthode calcule l'effet moyen des différences de prédiction. En d'autres termes, si nous voulons l'effet de la valeur x de la variable X , on définit en premier l'intervalle autour de x au sein duquel on calcul l'effet de toutes les valeurs prises par x . Enfin, on calcule la moyenne de ces effets. Ainsi, il faut définir une borne $z_{0,1}$ pour délimiter l'intervalle sur lequel on va faire la moyenne des différences de prédiction.

$$\hat{f}_{X_S, ALE}(x_S) = \int_{z_{0,1}}^{x_S} \mathbb{E}_{X_C|X_S} \left[\hat{f}^{(S)}(X_S, X_C) | X_S = x_S \right] dz_S = \int_{z_{0,1}}^{x_S} \int_{x_C} \hat{f}^{(S)}(x_S, x_C) P(x_C|x_S) dx_C dz_S$$

où $\hat{f}^{(S)}(x_S, x_C) = \frac{\partial \hat{f}(x_S, x_C)}{\partial x_S}$ est le gradient.

5.2.2.2 Estimation de l'ALE

En pratique, pour comprendre le comportement de la variable numérique x_j , on commence par diviser cette variable en plusieurs intervalles : $[z_{0,j}, z_{1,j}[$, $[z_{1,j}, z_{2,j}[$, \dots , $[z_{k_j(x)-1,j}, z_{k_j(x),j}]$ où



$k_j(x)$ est le nombre d'intervalles et $z_{0,j} < z_{1,j} < \dots < z_{k_j(x),j}$. Pour $k \in \{1, \dots, k_j(x)\}$, on note $N_j(k)$ l'ensemble des individus de la base d'apprentissage pour lesquels x_j est dans l'intervalle k et $n_j(k)$ le nombre d'individus dans $N_j(k)$. L'ALE au point $x \in \mathcal{X}$ associée à x_j est estimée par la formule :

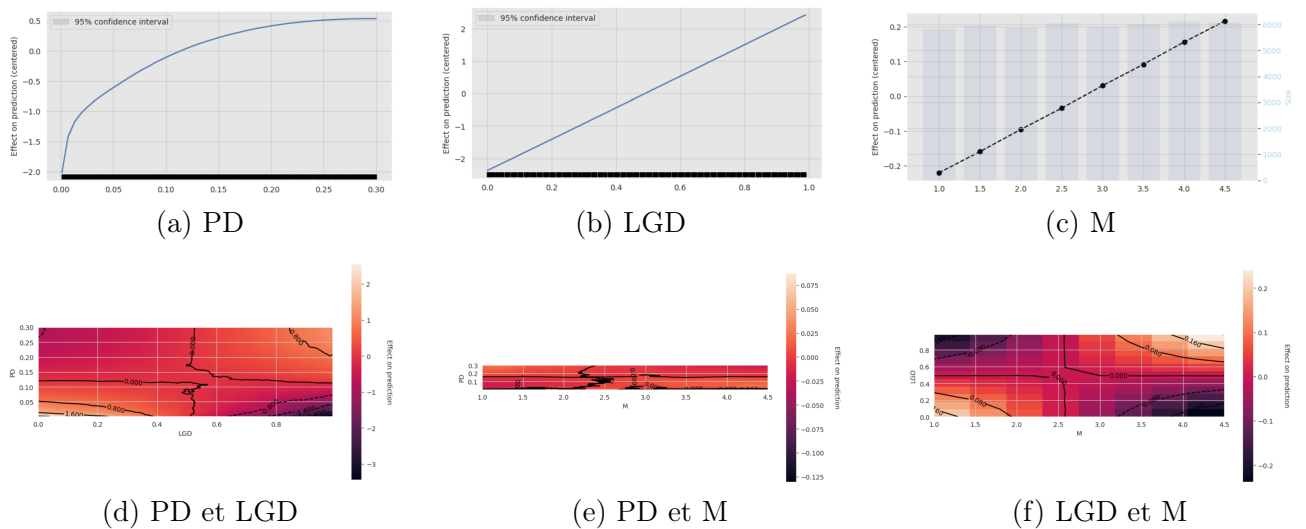
$$\hat{f}_{j,ALE}(x) = \sum_{k=1}^{k_j(x)} \frac{1}{n_j(k)} \sum_{i: x_j^{(i)} \in N_j(k)} \left[f(z_{k,j}, x_j^{(i)}) - f(z_{k-1,j}, x_j^{(i)}) \right]$$

Le terme d'effets locaux accumulés s'explique du fait qu'on somme la différence de prédiction sur tous les intervalles. En général, on retient les différents quantiles de la distribution de x_j comme intervalle pour le calcul de l'ALE. Cela possède l'intérêt d'avoir un nombre d'individus équivalents dans chaque intervalle.

5.2.2.3 Application au Random Forest

Compte tenu de nos données d'entrée, les résultats obtenus par le graphique des effets locaux accumulés se doivent d'être assez similaires à ceux obtenus à l'aide de la PDP. En effet, on peut clairement noter que les effets individuels sont similaires à ceux précédemment obtenus. On obtient donc une analyse du modèle dans sa globalité.

FIGURE 5.2 – ALE sur données simulées (Random Forest)



5.2.3 Importance des variables

5.2.3.1 Principe de la méthode

L'importance d'une variable ou *Permutation Feature Importance* (PFI) est le fait de considérer une variable comme étant d'autant plus importante que l'erreur de prédiction du modèle augmente après avoir permuté les valeurs de cette variable. En effet, si un changement des valeurs d'une variable impacte énormément les prédictions d'un modèle on peut conclure à



une forte dépendance entre le modèle et les valeurs prises par cette variable. Soit f la fonction que l'on cherche à prédire et $L(y, f(x)) = \frac{1}{2n} \sum_{i=1}^n (y^{(i)} - f(x_j^{(i)}))^2$ la fonction de coût d'intérêt.

La procédure pour le calcul de l'importance des variables est la suivante :

- évaluation de l'erreur de base : $err_1 = L(y, f(x))$
- pour $j = 1, \dots, p$:
 - on effectue une permutation σ de $\{1, \dots, n\}$. On définit la nouvelle matrice de données

$$\forall i \in \{1, \dots, n\}, \forall k \in \{1, \dots, p\} \quad \tilde{x}_k^{(i)} = \begin{cases} x_k^{(i)} & \text{si } k \neq j \\ x_k^{(\sigma(i))} & \text{si } k = j \end{cases}$$
 - on estime l'erreur commise sur le nouveau jeu de données : $err_j = L(y, f(\tilde{x}))$
 - on calcule l'importance de la variable x_j par : $FI(j) = \frac{err_j}{err_1}$
- en sortie de l'algorithme on a $FI(1), \dots, FI(p)$ par ordre décroissant

$$PFI_S = \frac{\mathbb{E}[L(\hat{f}(\tilde{X}_S, X_C), Y)]}{\mathbb{E}[L(\hat{f}(X), Y)]}$$

5.2.3.2 Avantages et inconvénients

La méthode d'importance des variables a plusieurs avantages parmi lesquels :

- une interprétation aisée ;
- une vision globale du modèle au même titre que les coefficients du modèle linéaire ;
- une comparaison facile des différents modèles ;
- la non-nécessité de ré-entraîner le modèle.

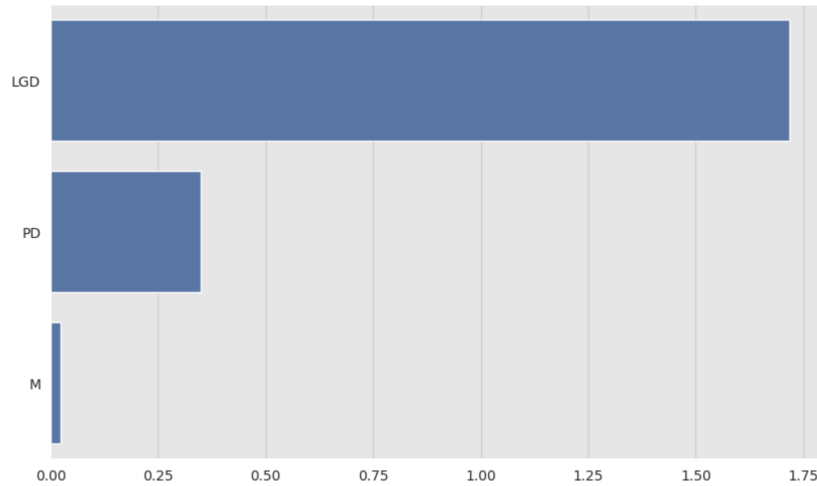
Dans l'autre sens, la présence d'une variable corrélée diminue l'importance de la variable considérée. Aussi, le résultat obtenu par l'algorithme est assez changeant en raison du hasard introduit par les permutations.

5.2.3.3 Application au Random Forest

Le graphique 5.3 montre l'importance de chaque paramètre dans la prédiction du risque. Il nous renseigne sur les contributions relatives des paramètres. En particulier, on peut noter que c'est la LGD qui explique principalement le niveau de risque affecté. Il a un effet cinq fois plus important que celui de la probabilité de défaut. Ainsi, les décideurs se doivent d'accorder une attention très particulière à la construction des modèles de LGD. Une erreur d'estimation a un impact direct très important sur le niveau de risque estimé. Aussi, on peut noter que la maturité a un effet très limité sur les prédictions. Ce résultat est confortant dans le sens où son ajout n'est qu'un choix réglementaire fait pour discriminer les crédits de long terme vis-à-vis des crédits de court-terme.



FIGURE 5.3 – Permutation Feature Importance sur données simulées (Random Forest)



5.2.4 Interaction entre les variables

5.2.4.1 Principe général

On parle d'interaction entre les variables lorsque les prédictions ne sont pas seulement dues aux effets individuels des variables, mais aussi à des termes supplémentaires représentant l'effet d'interaction. Dans le cas du modèle de régression linéaire, le terme d'interaction est représenté par le produit des variables concernées $X_i X_j (i \neq j)$.

5.2.4.2 H-statistique de Friedman

H-statistique sert à mesurer l'interaction entre les variables. On note PD_j la fonction de dépendance associée à la variable x_j et PD_{-j} la fonction de dépendance associée à toutes les variables sauf la variable x_j . La fonction de dépendance est estimée par la relation : $PD_S \simeq \frac{1}{n} \sum_{i=1}^n \hat{f}(x_S, x_C^{(i)})$. En cas d'absence d'interaction entre les variables x_j et x_k on a la relation :

$$PD_{j,k}(x_j, x_k) = PD_j(x_j) + PD_{-j}(x_{-j})$$

Dans le cas où x_j n'a d'interaction avec aucune des variables, la prédiction d'une entrée x vérifie :

$$\hat{x} = PD_j(x_j) + PD_{-j}(x_{-j})$$

Le premier coefficient $H_{j,k}$ introduit par Friedman mesure la quantité de variance expliquée par l'interaction entre x_j et x_k .

$$H_{j,k}^2 = \frac{\sum_{i=1}^n \left[PD_{j,k}(x_j^{(i)}, x_k^{(i)}) - PD_j(x_j^{(i)}) - PD_k(x_k^{(i)}) \right]^2}{\sum_{i=1}^n PD_{j,k}^2(x_j^{(i)}, x_k^{(i)})}$$

Cette statistique varie entre 0 et 1. En cas de non-interaction, elle prend la valeur 0.



La seconde statistique introduite par Friedman permet de mesurer l'effet d'une variable avec toutes les autres variables utilisées par le modèle \hat{f} . La statistique est donnée par la formule :

$$H_j^2 = \frac{\sum_{i=1}^n \left[\hat{f}(x^{(i)}) - PD_j(x_j^{(i)}) - PD_{-j}(x_{-j}^{(i)}) \right]^2}{\sum_{i=1}^n \hat{f}^2(x^{(i)})}$$

La H-statistique peut servir comme un test statistique de présence d'interaction entre les variables.

5.2.5 Modèles de substitution locaux : LIME

La méthode LIME est une méthode d'interprétation locale qui consiste à utiliser un modèle de substitution (M_2) qui approche au mieux le modèle initial (M_1) plus complexe.

Modèle de substitution : Dans la littérature, on retrouve les modèles de substitutions globaux et locaux. Les premiers approchent les modèles de boîtes noires dans leur globalité alors que les modèles locaux ne fournissent un modèle de substitution que sur une sélection d'observations.

On crée un nouveau dataset \tilde{X} en appliquant une légère perturbation au dataset initial X . Une pondération du dataset en fonction de sa proximité avec le dataset initial est utilisé pour construire le dataset pondéré. Le nouveau modèle M_2 fournit une bonne approximation locale du modèle. La fonction \hat{g} associée au modèle M_2 est la solution du problème d'optimisation :

$$\hat{g} = \underset{g \in G}{\operatorname{argmin}} [J(f, g, \pi_x) + \Omega(g)]$$

avec J la fonction de coût, π_x une mesure de proximité définissant la taille du voisinage autour de x et Ω une fonction traduisant la complexité d'un modèle.

La méthode LIME suit les étapes suivantes :

- on permute les variables explicatives ;
- pour chaque permutation, on évalue la prédiction à l'aide du modèle M_1 ;
- on construit un indice de similarité qui mesure la ressemblance entre les permutations et les données initiales ;
- on construit le modèle plus simple M_2 à l'aide de variables explicatives ;
- les poids des variables explicatives dans le modèle M_2 servent à expliquer le modèle M_1 .



5.2.6 SHAP

5.2.6.1 Valeur de Shapley en théorie des jeux

La décision prise par un modèle est l'objet d'une discussion de toutes les variables qui entrent en jeu. Ces dernières jouent des rôles différents dans la prédiction du modèle. La valeur de Shapley a pour rôle de quantifier le rôle joué par chacune de ces variables. Son principe est inspiré de la théorie des jeux. On considère un jeu $J = (\{1, \dots, p\}, v)$ de coalition où $P = \{1, \dots, p\}$ est un ensemble de p joueurs et $v : S(P) \rightarrow \mathbb{R}$ est une fonction caractéristique telle que $v(\emptyset) = 0$ et $S(P)$ est l'ensemble des sous-ensembles de P . Un sous-ensemble de joueurs $S \in S(P)$ est une coalition et P est la grande coalition. Le but du jeu est d'identifier de la plus juste des manières, l'importance de chaque joueur dans le gain total. La valeur de Shapley $\phi = (\phi_1, \dots, \phi_p)$ qui distribue le gain total $v(\{1, \dots, p\})$ est donnée par la relation :

$$\forall i \in \{1, \dots, p\}, \quad \phi_i(v) = \sum_{S \in S(\{1, \dots, p\} \setminus i)} \frac{(p - |S| - 1)! |S|!}{p!} (v(S \cup i) - v(S))$$

5.2.6.2 Méthode SHAP

La méthode de SHAP vise à utiliser les valeurs de Shapley pour expliquer les prédictions des modèles de boîtes noires. On se place toujours dans le cadre où l'on prédit une variable $Y \in \mathbb{R}$ à l'aide de p variables $X \in \mathbb{R}^p$. On note par \hat{f} la fonction associée au modèle. On considère aussi $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_p)$ l'instance pour laquelle on veut expliquer la prédiction. Soit $S = \{i_1, \dots, i_s\} \subset \{1, \dots, p\}$ (avec $s \in \{1, \dots, p\}$) et notons $\Delta^{\tilde{x}}$ la différence de prédiction induite par l'instance $\tilde{x} : \Delta^{\tilde{x}}(S) = \mathbb{E} \left[\hat{f}(X_1, \dots, X_p) | X_{i_1} = x_{i_1}, \dots, X_{i_s} = x_{i_s} \right] - \mathbb{E} \left[\hat{f}(X_1, \dots, X_p) \right]$. On construit donc un jeu de coalition $(\{1, \dots, p\}, \Delta^{\tilde{x}})$ comme définit précédemment. La contribution de la variable explicative x_j est donc déduite de la formule :

$$\phi_j(\Delta^{\tilde{x}}) = \sum_{S \in S(\{x_1, \dots, x_p\} \setminus \{x_j\})} \frac{(p - |S| - 1)! |S|!}{p!} (\Delta^{\tilde{x}}(S \cup x_j) - \Delta^{\tilde{x}}(S))$$

La méthode de SHAP a le mérite de présenter un fondement mathématique solide. Cette propriété lui confère une certaine fiabilité. Toutefois, en cas de forte corrélation entre les variables, les explications fournies peuvent s'avérer erronées. Il est proposé dans la littérature pour corriger ce problème de supposer les variables explicatives comme suivant une distribution gaussienne afin d'utiliser la théorie des copules pour y inclure le facteur de dépendance.



5.3 APPLICATION À DES DONNÉES RÉELLES : ÉCHANTILLON DU PORTEFEUILLE DE LA SOCIÉTÉ GÉNÉRALE

Les divers résultats présentés précédemment proviennent d'un jeu de données théorique. Bien qu'ils reflètent les réalités de la formule réglementaire, ils ne sont pas représentatifs du portefeuille de la Société Générale (SG). Cette section a pour but de reconstruire les différents modèles en se basant sur un échantillon des deals de la SG pour obtenir des effets plus précis.

5.3.1 Base de travail et note sur la confidentialité

5.3.1.1 Contexte

Le portefeuille de la SG comporte une grande typologie de clients présentant des réalités et caractéristiques différentes. Ces différences se matérialisent par des échelles de PD et des valeurs de LGD différentes. On retrouve entre autres les corporates, les institutions financières, les souverains et les PME. Du fait de la particularité des PME, la formule réglementaire de calcul du RWA appliquée sur ce périmètre a été légèrement modifiée par le comité de Bâle.

Nous décidons dans la suite de nos travaux, de ne retenir que le portefeuille de la SG exclusion faite des PME. Ce choix est fait en vue de garder une certaine cohérence avec l'ensemble des travaux précédemment réalisés.

5.3.1.2 Nature des données retenues

Les données retenues représentent des lignes de contrat (prêt contracté avec un tiers). Ainsi, nous nous intéressons aux caractéristiques du tiers en rapport avec le contrat. Il s'agit de sa probabilité de défaut (PD), de son taux de perte en cas de défaut (LGD), sa maturité (M) et du montant de l'exposition (EAD).

5.3.1.3 Processus de sélection des données

Notre base de travail est un échantillon du portefeuille de la SG. Nous avons effectué un tirage aléatoire sans remise de 10000 lignes de contrat tout en veillant à ce que notre échantillon possède la même structure que le portefeuille en termes de type de clients et produits. Ainsi, nous appliquons un tirage stratifié avec pour strate la variable issue de la combinaison entre client et produit. Le détail du processus de tirage est omis afin d'éviter toute extrapolation sur le portefeuille entier de la SG. Les résultats présentés dans la suite sont tout de même représentatifs du périmètre retenu. Un point d'attention est mis en exergue sur la confidentialité des chiffres obtenus par la suite.



5.3.2 Analyse statistique succincte de la base de travail

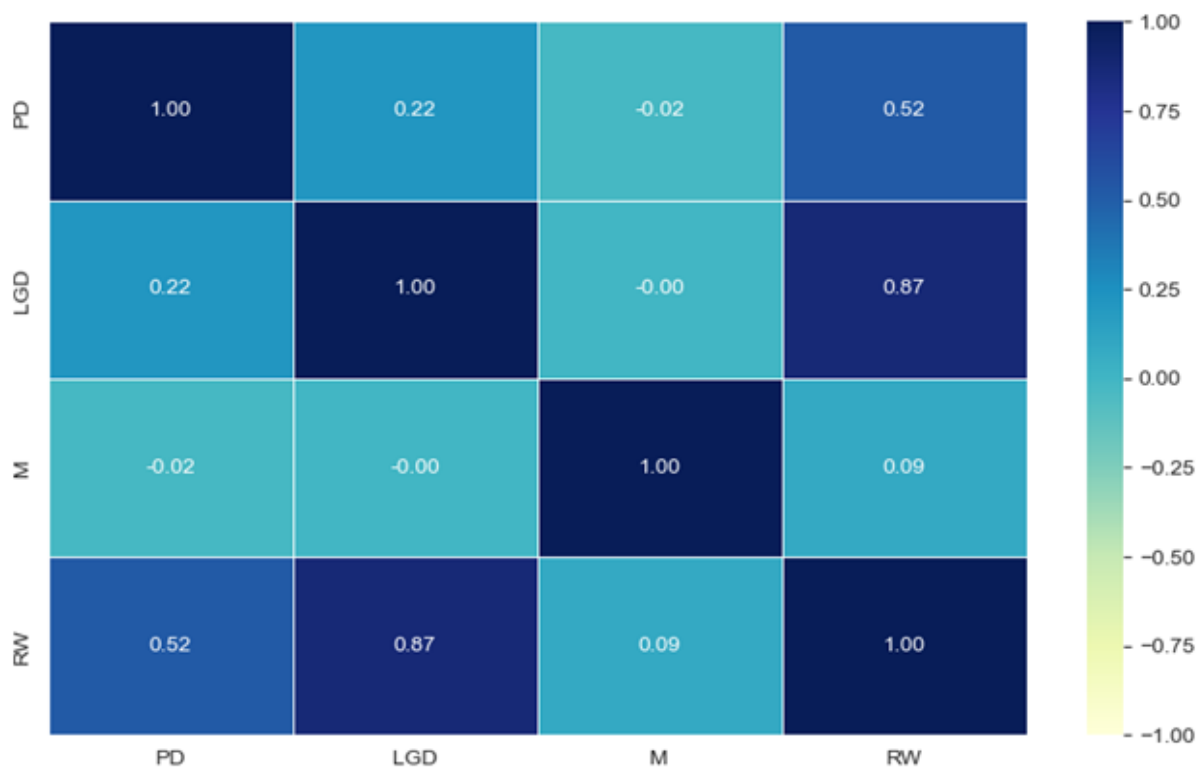
Nous nous intéressons ici aux statistiques descriptives qui peuvent apporter des informations nécessaires à la compréhension des interactions retranscrites par les modèles retenus. Ainsi, nous nous référons à la matrice de corrélation ainsi qu'aux histogrammes des différentes variables.

La matrice de corrélation de notre jeu de données (figure 5.4) nous révèle deux catégories d'informations importantes :

- ♦ **corrélation inter paramètres** : la probabilité de défaut et la perte en cas de défaut présentent une corrélation positive même si celle-ci est faible. Ce résultat est en concordance avec la théorie qui stipule qu'un tiers mal noté est plus susceptible de présenter un taux de recouvrement faible. Les efforts de la banque pour minimiser les pertes par le processus de garantie jouent assez bien leur rôle ici en rendant leur interdépendance faible. En ce qui concerne la maturité, elle ne présente pas de corrélation significative avec aucun des deux autres paramètres (PD et LGD). Cela laisse présager que la maturité des contrats ne présente pas d'incidence directe sur le risque couru par la SG.
- ♦ **corrélation avec la variable d'intérêt (RW)** : la LGD est la variable la plus fortement corrélée positivement au risque pondéré. Ce résultat suggère d'avance que la LGD puisse être la variable la plus déterminante dans le niveau de risque couru par la banque. Cette vision est pertinente vu que si la LGD d'un contrat est très faible, quel que soit le niveau de la PD, la perte engendrée dans le cas du défaut s'avérera minime et donc source d'un risque faible pour la banque.



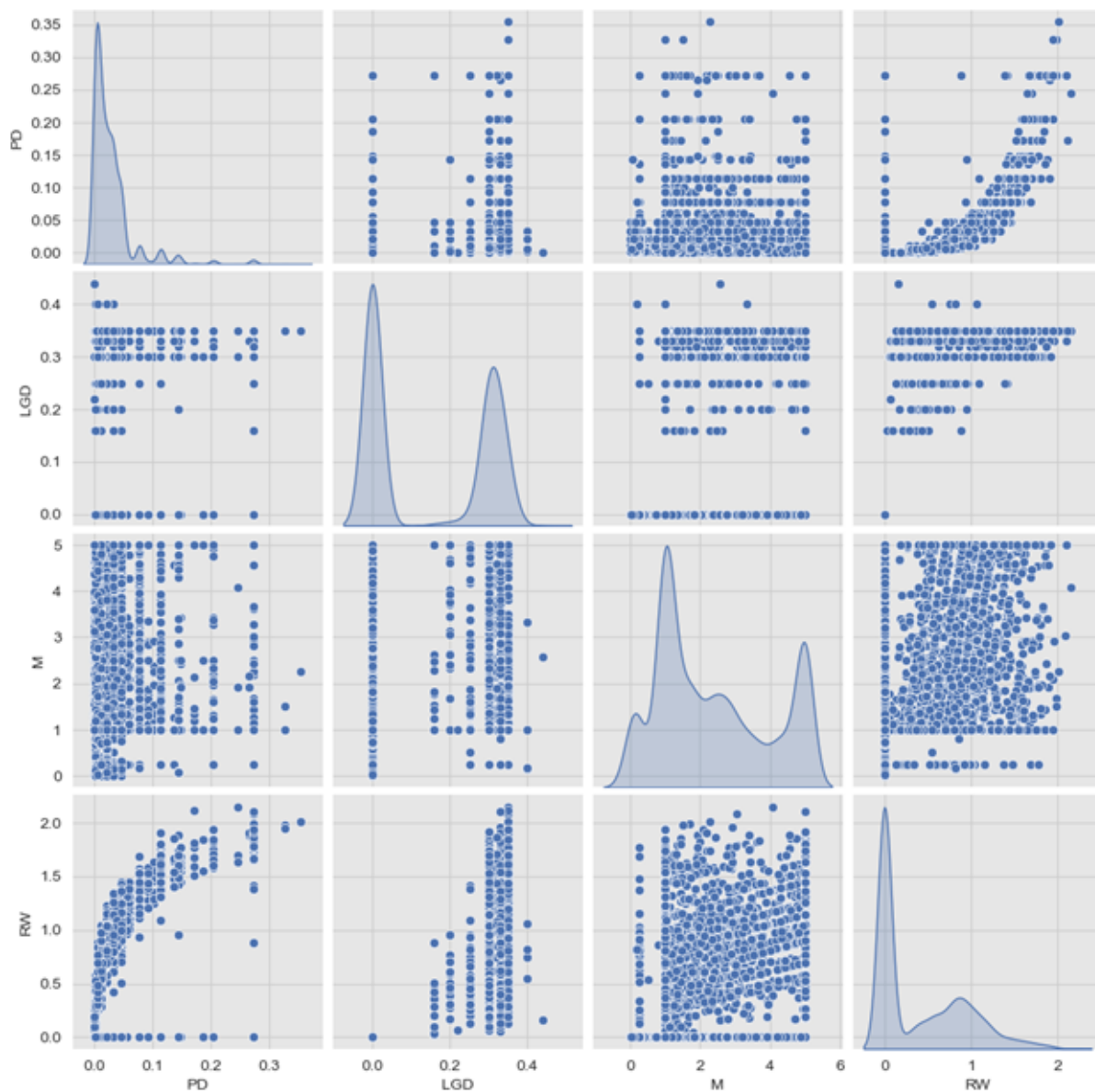
FIGURE 5.4 – Matrice de corrélation des données réelles



La figure 5.5 nous renseigne sur les distributions des nos données. Concernant la probabilité de défaut, on peut noter une forte concentration pour des valeurs proches de 0 et donc un étalement prononcé des données sur la droite. En effet, travaillant sur le portefeuille des corporates, des souverains et des institutions financières hormis les PME, notre base comporte un essentiel de tiers bien noté. Ce constat se réitère sur la LGD avec une forte concentration autour de la valeur 0. Cela implique de facto une concentration des valeurs de RW autour de 0. Le modèle que nous construirons par la suite devra impérativement retranscrire cet effet absorbant de la LGD. Enfin, on peut voir que pour la maturité, les données sont plus uniformément réparties malgré des pics à un an et cinq ans.



FIGURE 5.5 – Matrice de relation des données réelles



On retient de cette analyse que la LGD induit fortement la valeur future du RW et présente un caractère absorbant à 0 que doit nécessairement retranscrire le modèle que nous retiendrons.

5.3.3 Approche analytique

Selon cette approche, les effets marginaux des paramètres dépendent des valeurs de départ de la PD, LGD et maturité. Afin d’analyser son application, nous avons sélectionné dix deals aléatoirement au sein de notre base de travail et les contributions ont été calculées en supposant des variations de 0,1%, 1%, 10% et 25%.

Pour une variation d’1% de tous les paramètres, on note une contribution moyenne de la LGD de 67,35%. La PD arrive ensuite avec une contribution moyenne de 15,64%. Pour la maturité, on peut noter une contribution très disparate. En effet, des valeurs très fortes (~ 5) ont un effet très important tandis que des valeurs très faibles (~ 1) ont des effets très faibles.



On note enfin que les effets joints ont des contributions marginales dans la variation du RW.

TABLE 5.2 – Contributions pour une variation simultanée de 1% de tous les paramètres sur données réelles

Valeurs initiales				Variation du RW par approche analytique				Contribution des paramètres dans la variation du RW																
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD	contrib	LGD	contrib	PD	M	contrib	LGD	M	
0.011000	0.30	4.741713	0.876740	0.013876		0.013915		5.987424		63.007103		30.722828		0.000659		0.059874		-0.041654						0.307228
0.005015	0.35	1.000000	0.430772	0.006492		0.006508		18.816831		66.189730		14.742436		0.000944		0.188168		-0.010826						0.147424
0.046123	0.33	1.000000	0.994219	0.013592		0.013623		19.932851		72.981623		6.861454		0.009194		0.199329		-0.011142						0.068615
0.011000	0.30	5.000000	0.900027	0.014310		0.014351		5.585545		62.715870		31.412264		0.000614		0.055855		-0.042589						0.314123
0.011000	0.16	5.000000	0.480014	0.007632		0.007654		5.585545		62.715870		31.412264		0.000614		0.055855		-0.042589						0.314123
0.137042	0.33	1.000000	1.544914	0.024515		0.024561		33.385286		62.901328		3.525964		0.045752		0.333853		0.000297						0.035260
0.032596	0.33	1.000000	0.878314	0.011691		0.011712		16.651991		74.989693		8.173030		0.005428		0.166520		-0.015891						0.081730
0.013200	0.30	1.497378	0.625997	0.008869		0.008886		14.084414		70.450627		15.274567		0.001859		0.140844		-0.023359						0.152746
0.005015	0.30	1.398025	0.401966	0.006195		0.006212		16.510928		64.709120		18.508445		0.000828		0.165109		-0.013591						0.185084
0.046123	0.30	1.000000	0.903835	0.012357		0.012384		19.932851		72.981623		6.861454		0.009194		0.199329		-0.011142						0.068615

Nous présentons aussi les contributions dans le cas d'une variation très forte des paramètres (25%). Ce cas n'étant pas en pratique très réaliste, il sert de test de robustesse de l'approche (hypothèse de variation quasi nulle).

TABLE 5.3 – Contributions pour une variation simultanée de 25% de tous les paramètres sur données réelles

Valeurs initiales				Variation du RW par approche analytique				Contribution des paramètres dans la variation du RW																
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD	contrib	LGD	contrib	PD	M	contrib	LGD	M	
0.011000	0.30	4.741713	0.876740	0.346890		0.371471		5.607070		59.004546		28.771145		0.015419		1.401768		-0.975207						7.192786
0.005015	0.35	1.000000	0.430772	0.162295		0.172505		17.747697		62.428965		13.904801		0.022251		4.436924		-0.255269						3.476200
0.046123	0.33	1.000000	0.994219	0.339809		0.358887		18.915623		69.257169		6.511295		0.218111		4.728906		-0.264332						1.627824
0.011000	0.30	5.000000	0.900027	0.357744		0.383425		5.226403		58.683331		29.392501		0.014373		1.306601		-0.996268						7.348125
0.011000	0.16	5.000000	0.480014	0.190797		0.204493		5.226403		58.683331		29.392501		0.014373		1.306601		-0.996268						7.348125
0.137042	0.33	1.000000	1.544914	0.612872		0.641643		31.948213		60.193735		3.374189		1.094565		7.987053		0.007116						0.843547
0.032596	0.33	1.000000	0.878314	0.292269		0.305832		15.943027		71.796982		7.825061		0.129920		3.985757		-0.380356						1.956265
0.013200	0.30	1.497378	0.625997	0.221717		0.232291		13.468961		67.372117		14.607108		0.044448		3.367240		-0.558449						3.651777
0.005015	0.30	1.398025	0.401966	0.154876		0.165417		15.500865		60.750515		17.376184		0.019434		3.875216		-0.318998						4.344046
0.046123	0.30	1.000000	0.903835	0.308917		0.326261		18.915623		69.257169		6.511295		0.218111		4.728906		-0.264332						1.627824

Une analyse comparative des tableaux 5.2 et 5.3 montrent que sur les contributions des effets individuels, baissent systématiquement lors du passage des variations simultanées à 25%. Les effets joints ont donc un impact plus marqué sur la variation du RW. Aussi, on peut remarquer que différence d'approximation entre le premier et le deuxième degré est plus importante pour la variation de 25%. Ainsi, pour des variations importantes des paramètres, il apparaît plus pertinent de recourir à l'approximation au deuxième degré de la formule réglementaire. Ce constat est renforcé par la qualité d'ajustement du second degré présenté dans le tableau 5.4. On peut clairement y voir que l'approximation au second ordre présente une bonne qualité d'ajustement en plus de plusieurs axes d'interprétation.



TABLE 5.4 – Comparaison des approximations dans l’approche analytique sur données réelles

Variations	RMSE ordre 1	MAE ordre 1	RMSE ordre 2	MAE ordre 2
Delta / PD = 0.1%, LGD = 0.1%, M = 0.1%	0.0002	0.0082	0.0002	0.0082
Delta / PD = 1.0%, LGD = 0.1%, M = 0.1%	0.0015	0.0259	0.0015	0.0259
Delta / PD = 10.0%, LGD = 0.1%, M = 0.1%	0.0157	0.0810	0.0146	0.0808
Delta / PD = 0.1%, LGD = 1.0%, M = 0.1%	0.0002	0.0083	0.0002	0.0082
Delta / PD = 1.0%, LGD = 1.0%, M = 0.1%	0.0015	0.0261	0.0015	0.0260
Delta / PD = 10.0%, LGD = 1.0%, M = 0.1%	0.0157	0.0816	0.0147	0.0811
Delta / PD = 0.1%, LGD = 0.1%, M = 1.0%	0.0002	0.0083	0.0002	0.0082
Delta / PD = 1.0%, LGD = 0.1%, M = 1.0%	0.0015	0.0260	0.0015	0.0260
Delta / PD = 10.0%, LGD = 0.1%, M = 1.0%	0.0157	0.0811	0.0146	0.0808
Delta / PD = 0.1%, LGD = 1.0%, M = 1.0%	0.0002	0.0087	0.0002	0.0083
Delta / PD = 1.0%, LGD = 1.0%, M = 1.0%	0.0015	0.0263	0.0015	0.0261
Delta / PD = 10.0%, LGD = 1.0%, M = 1.0%	0.0157	0.0816	0.0148	0.0812

5.3.4 Modélisation

5.3.4.1 Modélisation classique

Le modèle linéaire que nous présentons ici présente les mêmes limites que celui sur données simulées. En l’occurrence la prédiction de valeur négative et l’estimation d’un effet fixe pour chaque paramètre.

TABLE 5.5 – Résultats du modèle linéaire sur données réelles

Dep. Variable :	RW	R-squared :	0.883			
Model :	OLS	Adj. R-squared :	0.883			
No. Observations :	8000	Prob (F-statistic) :	0.00			
Df Model :	3	Log-Likelihood :	2730.3			
AIC :	-5453	BIC :	-5425			
	coef	std err	t	P > t 	[0.025	0.975]
const	-0.1705	0.004	-42.336	0.000	-0.178	-0.163
PD	4.5038	0.051	88.623	0.000	4.404	4.603
LGD	2.5694	0.013	203.789	0.000	2.545	2.594
M	0.0318	0.001	26.585	0.000	0.029	0.034

Ce modèle présente tout de même l’intérêt de sa simplicité et de son interprétabilité. La LGD a un effet plus que proportionnel au RW. Une variation d’1% de la LGD entraîne une variation de 1,11% du risque. Ce résultat différent de la parité parfaite obtenue sur données simulées pourrait provenir de la répartition non uniforme des valeurs de LGD (voir graphique 5.5, page 70).



TABLE 5.6 – Résultats du modèle linéaire (fonction logarithmique) sur données réelles

Dep. Variable :	ln(RW)	R-squared :	0.957
Model :	OLS	Adj. R-squared :	0.957
No. Observations :	4702	Prob (F-statistic) :	0.00
Df Model :	3	Log-Likelihood :	3736.3
AIC :	-7465	BIC :	-7439

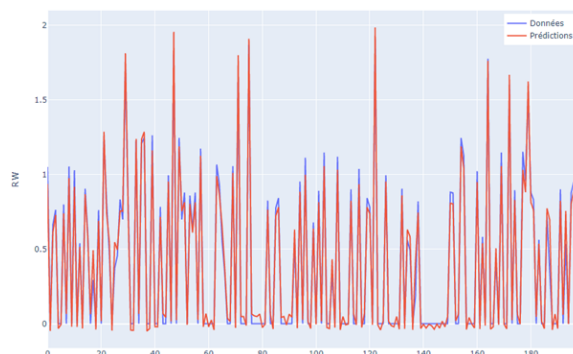
	coef	std err	t	P > t	[0.025	0.975]
const	2.3885	0.018	129.487	0.000	2.352	2.425
ln(PD)	0.3801	0.001	290.581	0.000	0.378	0.383
ln(LGD)	1.1137	0.016	69.697	0.000	1.082	1.145
ln(M)	0.2503	0.002	107.691	0.000	0.246	0.255

Le graphique 5.6 confronte les prédictions faites par les modèles classiques (linéaire et polynomiaux) et les valeurs réelles obtenues à l'aide de la formule réglementaire. Les conclusions sont similaires, plus on retient de degrés, mieux le modèle ajuste les données.

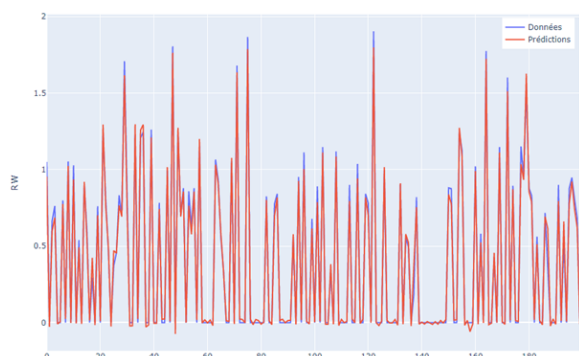
FIGURE 5.6 – Prédiction RW par modélisation classique sur données réelles



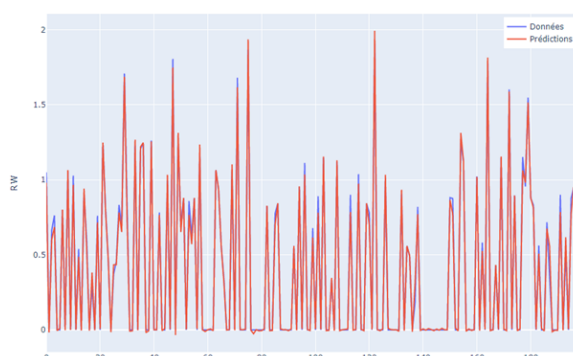
(a) Modèle linéaire



(b) Modèle polynomial (degré 2)



(c) Modèle polynomial (degré 3)



(d) Modèle polynomial (degré 4)

Pour un modèle aussi complexe que la régression polynomiale de degré 4, on préfère avoir recours à un modèle de boîte noire qui présente de meilleures performances.

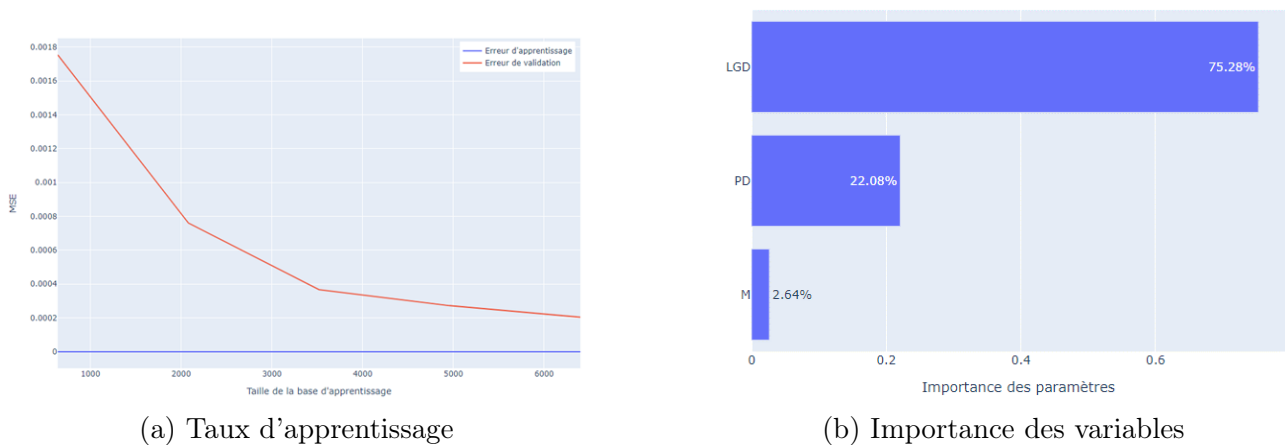


5.3.4.2 Modélisation par machine learning

Les modèles de machine learning basés sur des arbres de décisions arrivent à reproduire l'état absorbant 0 de la LGD. Cette propriété leur confère des performances plus poussées que les modèles polynomiaux. Le principal défaut reste leur interprétabilité. En se basant sur les fonctions de pertes, on arrive à estimer une version de l'importance des variables qui se rapproche assez de la PFI. Ici, ces importances prennent plutôt la forme de contribution dans l'explication de la valeur prédite du RW.

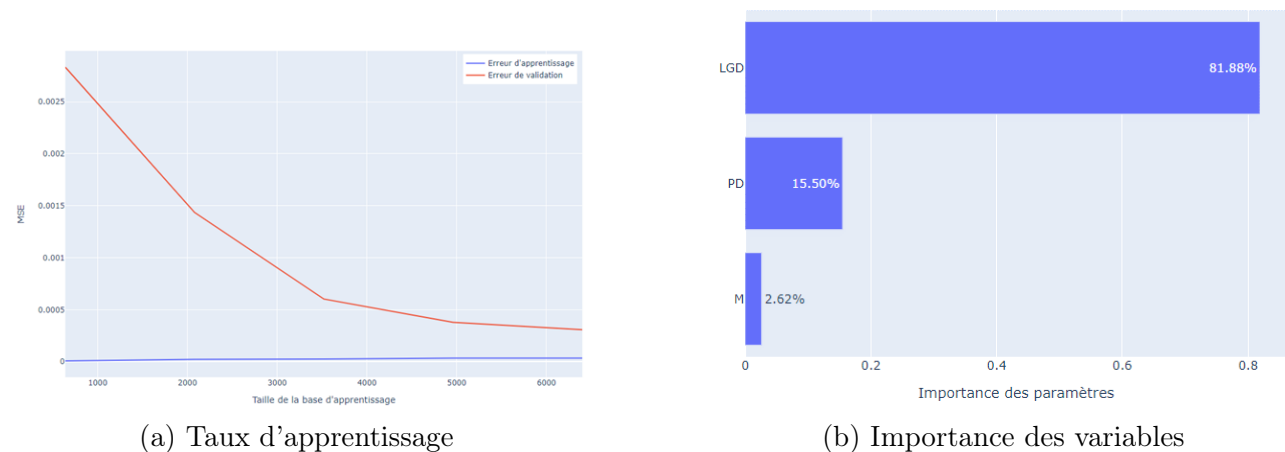
Le graphique 5.7 montre un ordre de contribution des paramètres similaires à ceux obtenus précédemment. La LGD contribue 3,4 fois plus que la PD dans la valeur de risque prédite par le Random Forest. La courbe d'évolution du taux d'apprentissage montre que le modèle a bien été entraîné sur une taille de données rendant stables les résultats. On peut dès lors considérer nos résultats comme valides.

FIGURE 5.7 – Random Forest sur données réelles



Les mêmes conclusions ressortent avec le XGBoost même s'il accorde un poids plus faible à la probabilité de défaut. Pour sa part, la LGD contribue 5,3 fois plus que la PD.

FIGURE 5.8 – XGBoost sur données réelles



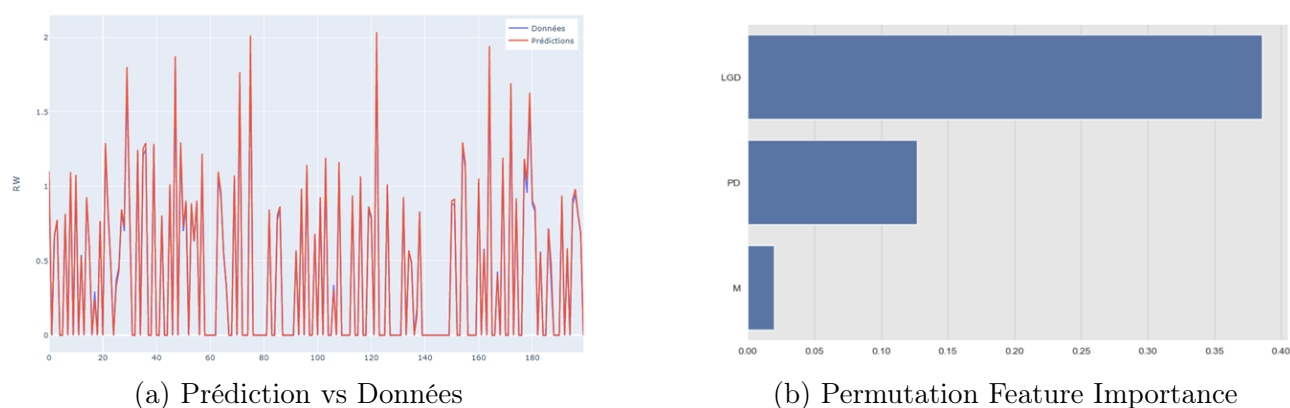


5.3.4.3 Modélisation par réseau de neurones

Dans la littérature, les modèles de réseau de neurones présentent des résultats très surprenants. C'est une classe de modèles qui arrive à reproduire les relations les plus complexes que l'on puisse retrouver. Cela explique son application massive ces dernières années. Malgré le fait que les modèles d'arbres semblent un choix judicieux pour notre problématique, nous construisons également un modèle de réseau de neurones afin de confronter les performances. La section suivante fait une comparaison des différents modèles et donne une vision plus poussée du modèle retenu.

On peut voir sur le graphique 5.9 que ce modèle ajuste très bien nos données. Ce modèle arrive à une conclusion similaire que le Random Forest en termes d'effet relatif des paramètres. En effet, il estime que la LGD a un effet 3,2 fois plus important que la PD dans la prédiction du risque.

FIGURE 5.9 – Réseau de neurones sur données réelles



5.3.4.4 Interprétation du modèle retenu

Comme sur données simulées, le Random Forest présente les meilleures performances. En dehors du réseau de neurones, tous les modèles estimés entraînent une sous-évaluation du risque. Même si ce biais est de l'ordre de 10^{-4} , il est quand même présent.

TABLE 5.7 – Comparaison des modèles sur données réelles

	Modèle	Biais	RMSE	MAE	RSR	R2
<i>Modèles classiques</i>	Modèle linéaire	0.0002	0.1738	0.1270	0.343948	0.8817
	Modèle polynomial (degré 2)	0.0005	0.0845	0.0556	0.167332	0.9721
	Modèle polynomial (degré 3)	0.0004	0.0576	0.0342	0.114018	0.9870
	Modèle polynomial (degré 4)	0.0009	0.0450	0.0240	0.088882	0.9921
<i>Modèles de ML</i>	Random Forest	0.0004	0.0101	0.0020	0.020000	0.9996
	XGBoost	0.0006	0.0117	0.0052	0.022361	0.9995
<i>Modèle de DeepL</i>	Réseaux de neurones	-0.0113	0.0315	0.0155	0.062450	0.9980

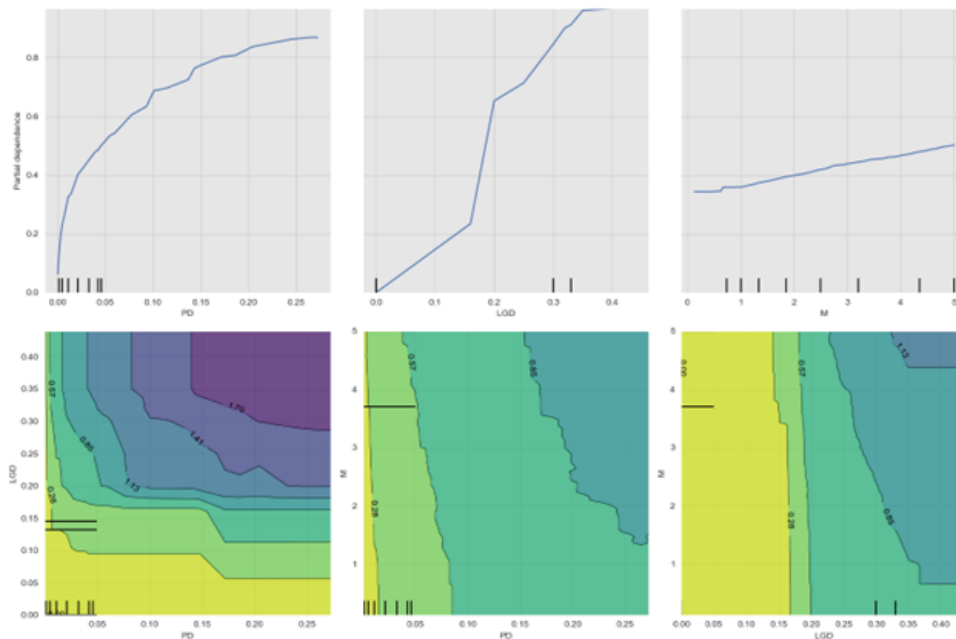
Nous retenons encore une fois le Random Forest comme modèle d'approximation de la formule réglementaire.



L'estimation des effets marginaux dans notre cas de figure par la Partial Dependence Plot (PDP) souffre d'une limite, l'indépendance des paramètres. Même si la corrélation entre la probabilité de défaut et la perte en cas de défaut est faible (voir graphique 5.4, page 69), elle existe. Ce constat qui a tout son sens d'un point de vue financier limite la fiabilité de la PDP dans l'évaluation des effets marginaux. Nous compléterons donc ces résultats par ceux fournis par la courbe des effets locaux accumulés (ALE).

Le graphique 5.10 nous présente des résultats globalement similaires à ceux obtenus sur données simulées. Cependant, on peut noter que les effets joints sont plus aléatoires que précédemment. En l'occurrence pour un taux de perte à 20%, une variation de la probabilité de défaut a beaucoup moins d'effet que pour un taux de perte à 35%. En somme, les graphiques 5.10 et A.3 (page XIII) donnent une estimation assez similaire des effets marginaux moyens individuels. La probabilité de défaut possède un effet logarithmique tandis que la perte en cas de défaut et la maturité possèdent eux, un effet plus ou moins linéaire.

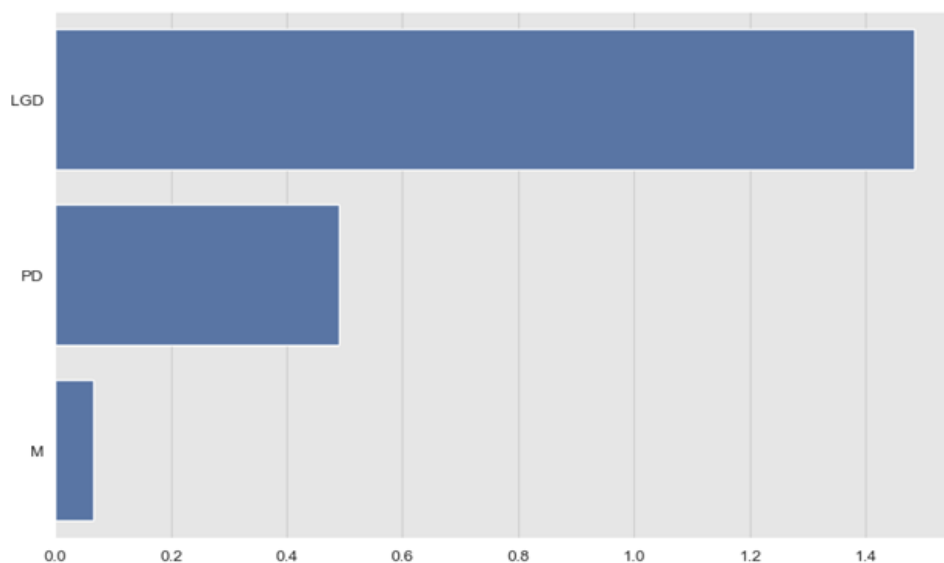
FIGURE 5.10 – Partial Dependence Plot sur données réelles (Random Forest)



Enfin, nous nous intéressons à une estimation des effets moyens de chacun des paramètres. L'importance des variables nous donne en plus de cette information, une estimation des contributions relatives de chacune d'entre elles. On retrouve des résultats familiers, une plus grande importance de la perte en cas de défaut dans l'évaluation du risque de crédit (graphique 5.11). De plus, cet effet est trois fois plus important que celui de la probabilité de défaut. La maturité quant à elle, a un effet assez marginal sur l'évaluation du risque, elle respecte sa propriété d'ajustement de maturité pour la prise en compte des effets de long terme.



FIGURE 5.11 – Permutation Feature Importance sur données réelles (Random Forest)



En résumé

Ce chapitre est revenu sur l'application de l'ensemble des travaux préliminaires sur un échantillon de données issu du portefeuille de la SG. Concrètement, nous avons travaillé à l'approximation de la formule réglementaire suivant une approche analytique et une autre sous forme de modélisation. De manière consensuelle, toutes les méthodes proposées ont ressorti que la LGD explique approximativement à hauteur de 70% le niveau de risque couru par la banque sur une ligne de contrat. Aussi, la LGD présente un caractère absorbant à sa valeur 0 qui induit une sélection plus fine des modèles adéquats d'approximation. Nos estimations ont ressorti le Random Forest comme étant le modèle qui répliquait le mieux les propriétés de la formule réglementaire. En somme, l'approximation de la formule réglementaire ainsi que les données de la SG (forte présence de LGD nulle dans le portefeuille) induisent une vigilance particulière sur les modèles de LGD construits. Ayant un effet trois fois plus important que la probabilité de défaut (feature importance), une mauvaise maîtrise de la LGD par la banque a un impact non négligeable sur l'estimation du niveau de risque.

CONCLUSION, LIMITES ET PERSPECTIVES

Dans ce mémoire, il a été question d'analyser la formule réglementaire d'évaluation du risque de crédit. Cette analyse vise à évaluer les effets des paramètres de la fonction. Il a donc été question de confronter plusieurs approches afin d'estimer de manière précise ces effets.

Nous pouvons conclure au terme de notre analyse que trois grandes possibilités existent pour répondre à cette problématique, chacune moyennant des intérêts et des limites.

La première possibilité fondée sur la formule de Taylor fournit une approche analytique permettant d'estimer à la fois les effets marginaux et les effets joints de chaque paramètre dans la prédiction du risque. Cette première tentative présente l'intérêt de proposer une formule fermée et claire de calcul de ces effets. Son implémentation dans un outil de monitoring est une chose aisée. Toutefois, la qualité de son approximation repose sur une hypothèse de variation quasi nulle des paramètres. Cette dernière remarque limite son applicabilité.

La seconde approche découle des modèles de régression classique à savoir la régression linéaire et celle polynomiale. Cette alternative fournit clairement une nouvelle formule d'évaluation du risque qui a le mérite d'être simple. Cette simplicité lui confère une facilité d'interprétation et d'estimation des effets fixes. L'atout majeur de cette approche en est aussi la principale faiblesse. En effet, l'estimation d'effet fixe quelque soit la valeur considérée des variables, entraîne la prédiction via ces modèles de niveau de risque négatif sur certaines réalisations de la probabilité de défaut et de la perte en cas de défaut. Le choix de cette solution est donc assimilable à un arbitrage entre facilité d'interprétation et difficulté d'approximation.

La troisième et dernière approche vient en réponse aux limites des deux précédentes. En effet, elle a recourt à des modèles de machine learning qui arrivent à apprendre les mécanismes de la formule réglementaire. Cette approche fournit les meilleurs résultats en termes d'approximation de la formule réglementaire d'autant qu'elle prédit parfaitement des valeurs positives. Cependant, la complexité de sa construction empêche une facile interprétation. Nous faisons donc appel à des méthodes alternatives de quantification des effets des paramètres. Nous retenons qu'elle fournit des effets variables aux différents paramètres et évalue la contribution de la LGD trois fois plus importante que celle de la PD dans l'évaluation du risque.



En définitive, même si l'approche par machine learning ne permet pas d'obtenir une nouvelle formule simple d'évaluation du risque de crédit, nous recommandons son utilisation étant donné les résultats chiffrés qu'elle apporte à l'estimation de l'effet des paramètres de la fonction réglementaire.

On retient de cette analyse que l'effort de modélisation des paramètres de la fonction réglementaire se doit de prendre en compte les possibles effets négatifs que pourrait engendrer une mauvaise estimation de la perte en cas de défaut. Une attention particulière doit être mise dans son estimation.

•→ **Limites et perspectives**

Dans cette étude, nous nous intéressons à l'effet des paramètres de la formule réglementaire en les considérant comme des données. Il serait judicieux dans une extension de construire une fonction d'évaluation du risque fondée sur les drivers des modèles de probabilité de défaut et de taux de pertes en cas de défaut. Cette extension permettrait de proposer une nouvelle formule réglementaire corrigeant les limites de celle actuelle.

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] ACPR, « Finalisation de Bâle III - En bref, » Banque de France, déc. 2017. DOI : https://www.bis.org/bcbs/publ/d424_inbrief_fr.pdf.
- [2] Basel Committee on BANKING SUPERVISION, « Working Paper on the Regulatory Treatment of Operational Risks, » Bank for International Settlements, sept. 2001. DOI : http://www.bis.org/publ/bcbs_wp8.pdf.
- [3] Basel Committee on BANKING SUPERVISION, « Comment les pondérations de risque diffèrent-elles parmi les banques? Étude comparée sur les portefeuilles « Entreprises » des banques françaises, » Bank for International Settlements, mars 2015. DOI : <https://acpr.banque-france.fr/sites/default/files/20150311-as42-comment-les-ponderations-de-risque-different-parmi-les-banques.pdf>.
- [4] Basel Committee on BANKING SUPERVISION, « Instructions for Basel III monitoring, » Bank for International Settlements, août 2017. DOI : https://www.bis.org/bcbs/qis/biiiimplmoninstr_aug17.pdf.
- [5] Sylvie BENZONI-GAVAGE, « Calcul différentiel et équations différentiels, » DUNOD, mars 2015, Université Lyon I.
- [6] Gerard BIAU, « Analysis of a Random Forests Model, » Journal of Machine Learning Research, 2012. DOI : <https://www.jmlr.org/papers/volume13/biau12a/biau12a.pdf>.
- [7] Gerard BIAU et Erwan SCORNET, « A Random Forest Guided Tour, » Sorbonne Université, DOI : <https://www.normalesup.org/~scornet/paper/test.pdf>.
- [8] Fabio Orlando Guacaneme CASTIBLANCO, « Estimation de modèle de risque de crédit avec des filtres à particules, » Université du Québec à Montréal, juin 2016.
- [9] Tianfeng CHAI et R. R. DRAXLER, « Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, » Geoscientific Model Development, juin 2014. DOI : <https://gmd.copernicus.org/books/7/1247/2014/gmd-7-1247-2014.pdf>.
- [10] Françoise CHARLES, « Des modèles internes pour asseoir les exigences en fonds propres des banques au titre du risque de crédit : de Bâle II à Bâle III, » Banque de France, 2015.
- [11] Christophe CHESNEAU, « Modèles de régression, » Université de Caen, oct. 2020.



- [12] Michel CROUHY, Dan GALAI et Robert MARK, « Risk Management, » McGraw-Hill, 2001.
- [13] Adele CUTLER, Richard CUTLER et John STEVENS, « Random Forests, » Utah State University, jan. 2011.
- [14] Dimitri DELCAILLAU, « Contrôle et Transparence des modèles complexes en actuariat, » EURIA, sept. 2019, Milliman. DOI : <https://www.institutdesactuaires.com/docs/mem/af306e55a66c5ce6bcf6696cd3df9e0a.pdf>.
- [15] Julien DHIMA, « Evolution des méthodes de gestion des risques dans les banques sous la réglementation de Bale III : une étude sur les stress tests macro-prudentiels en Europe, » Economies et finances, 2019, Université Panthéon-Sorbonne - Paris I. DOI : <https://tel.archives-ouvertes.fr/tel-02440557>.
- [16] Julien DHIMA, « Une analyse critique sur la finalisation de Bâle III, » BSI Economics, 2020. DOI : <http://www.bsi-economics.org/images/bale.pdf>.
- [17] Gérard DREYFUS, « Les réseaux de neurones, » sept. 1998, École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (ESPCI).
- [18] Darell DUFFIE et Kenneth SINGLETON, « Credit Risk : Pricing, Measurement and Management, » Princeton University Press, 2003.
- [19] François HAAS, « Vers un « continuum de marché » ? Modèles structurels et interactions entre marchés de crédit et d'actions, » Banque de France, juin 2003.
- [20] Andrew G. HALDANE et Vasileios MADOUROS, « La réglementation basée sur le risque : est-ce le moment pour un retour vers le futur ? » Revue d'économie financière, 2013, Association d'économie financière. DOI : <https://www.cairn.info/revue-d-economie-financiere-2013-4-page-93.htm>.
- [21] Jacob HALLMAN, « A comparative study on Linear Regression and Neural Networks for estimating order quantities of powder blends, » KTH Royal Institute of Technology, 2019. DOI : <http://www.diva-portal.org/smash/get/diva2:1383464/FULLTEXT01.pdf>.
- [22] Olivier Junior KARUSISI, « Risque de crédit, » Société Watson Wyatt, 2007.
- [23] Florian KASTNER, Benedikt JANSSEN, Frederik KAUTZ et Michael HÜBNER, « Exploring Deep Neural Networks for Regression Analysis, » PESARO, 2018.
- [24] Stephane LATHUILIÈRE, Pablo MESEJO, Xavier ALAMEDA-PINEDA et Radu HORAUD, « A Comprehensive Analysis of Deep Regression, » IEEE, sept. 2020. DOI : <https://arxiv.org/pdf/1803.08450.pdf>.
- [25] Andy LIAW et Matthew WIENER, « Classification and Regression by randomForest, » R News, déc. 2002. DOI : <https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf>.
- [26] François LONGIN, « Management bancaire : Le modèle de Merton, » ESSEC, 2014.



-
- [27] Gilles LOUPPE, « Understanding Random Forests : from theory to practice, » University of Liège, juill. 2014. DOI : <https://arxiv.org/pdf/1407.7502.pdf>.
- [28] Gilles MARCOU et Pierre JOST, « Cours de statistique, » Université de Strasbourg, 7 juill. 2011.
- [29] Harry MARKOWITZ, « Portfolio selection : Efficient Diversification of Investments, » Yale University Press, 1959.
- [30] Franck MORAUX et Olivier RENAULT, « 30 ans de modèles structurels de risque de défaut, » 2002. DOI : <https://perso.univ-rennes1.fr/franck.moraux/research/morauxrenault.pdf>.
- [31] Alexis NASR, « Regression polynomiale, » Aix Marseille Université,
- [32] Joseph SALMON, « Méthodes non paramétriques : Polynômes, splines et GAM, » Université de Montpellier,
- [33] Laurent SCHWARTZ, « Analyse II : Calcul différentiel et équations différentiels, » Hermann, Editeur des sciences et des arts, 1992.
- [34] Maroua SENHAJ, « Mesure du risque de contrepartie - Comparaison de différentes méthodologies de diffusion de facteurs de risques "actions", » ENSAE Paris, fév. 2019.
- [35] Claude TOUZET, « Les réseaux de neurones artificiels - Introduction au connexionnisme, » HAL Open science, juill. 1992, Collection de l'EERIE. DOI : <https://hal-amu.archives-ouvertes.fr/hal-01338010>.
- [36] Brian Hébert D. WHANNOU, « Intégration de la LGD des dossiers non-clos à l'estimation de la LGD des actifs sains, » ENSAE Paris, nov. 2016.
- [37] Xin XU, « Modélisation De La Probabilité De Défaut Dans Le Cadre de Stress Scénarii Bâle II, » ISUP, 2011.

◆ Formules

Dérivées partielles de la formule réglementaire

Pour des raisons de simplification, nous adoptons les notations suivantes :

$$RW = LGD \times B \times C \times 12,5 \times 1,06$$

$$B = \Phi(A) - PD$$

$$A = \frac{1}{\sqrt{1-R}} \cdot \Phi^{-1}(PD) + \sqrt{\frac{R}{1-R}} \cdot \Phi^{-1}(0,999)$$

$$C = \frac{1 + (M - 2,5) \cdot b}{1 - 1,5 \cdot b}$$

$$\frac{\partial RW}{\partial PD} = LGD \times (B' \times C + B \times C') \times 12,5 \times 1,06$$

$$\frac{\partial RW}{\partial LGD} = B \times C \times 12,5 \times 1,06$$

$$\frac{\partial RW}{\partial M} = LGD \times B \times D \times 12,5 \times 1,06$$

$$\frac{\partial^2 RW}{\partial PD^2} = LGD \times (B'' \times C + 2 \times B' \times C' + B \times C'') \times 12,5 \times 1,06$$

$$\frac{\partial^2 RW}{\partial PD \partial LGD} = (B' \times C + B \times C') \times 12,5 \times 1,06$$

$$\frac{\partial^2 RW}{\partial PD \partial M} = LGD \times (B' \times D + B \times D') \times 12,5 \times 1,06$$

$$\frac{\partial^2 RW}{\partial LGD \partial M} = B \times D \times 12,5 \times 1,06$$



Dérivées utiles au calcul des effets dans l'approche analytique

$$\left\{ \begin{array}{l} b = (0,11852 - 0,05478 \times \ln(PD))^2 \\ b' = -2 \frac{0,05478}{PD} (0,11852 - 0,05478 \times \ln(PD)) \\ b'' = 2 \frac{0,05478}{PD^2} (0,1733 - 0,05478 \times \ln(PD)) \end{array} \right. \quad \left\{ \begin{array}{l} R = 0,24 - 0,12 \cdot \frac{1 - e^{-50 \cdot PD}}{1 - e^{-50}} \\ R' = -\frac{0,12}{1 - e^{-50}} \cdot (50 \cdot e^{-50 \cdot PD}) \\ R'' = \frac{0,12}{1 - e^{-50}} \cdot (50^2 \cdot e^{-50 \cdot PD}) \end{array} \right.$$

$$\left\{ \begin{array}{l} B = \Phi[A] - PD \\ B' = A' \times \phi[A] - 1 \\ B'' = (A'' - A'^2 \cdot A) \times \phi[A] \end{array} \right. \quad \left\{ \begin{array}{l} D = \frac{b}{1 - 1,5 \cdot b} \\ D' = \frac{b'}{(1 - 1,5 \cdot b)^2} \end{array} \right. \quad \left\{ \begin{array}{l} C = \frac{1 + (M - 2,5) \cdot b}{1 - 1,5 \cdot b} \\ C' = \frac{b'(M - 1)}{(1 - 1,5 \cdot b)^2} \\ C'' = \frac{(M - 1)[b''(1 - 1,5b) + 3b'^2]}{(1 - 1,5 \cdot b)^3} \end{array} \right.$$

$$\left\{ \begin{array}{l} A = \frac{1}{\sqrt{1-R}} \cdot \Phi^{-1}(PD) + \sqrt{\frac{R}{1-R}} \cdot \Phi^{-1}(0,999) \\ A' = \frac{R'}{2\sqrt{(1-R)^3}} \cdot \left(\Phi^{-1}(PD) + \frac{1}{\sqrt{R}} \Phi^{-1}(0,999) \right) + \frac{1}{\sqrt{1-R}} \cdot \frac{1}{\phi[\Phi^{-1}(PD)]} \\ A'' = \frac{R'' + 1,5R'(1-R)}{2\sqrt{(1-R)^3}} \cdot \left(\Phi^{-1}(PD) + \frac{1}{\sqrt{R}} \Phi^{-1}(0,999) \right) - \frac{R'^2}{4\sqrt{R^3(1-R)^3}} \Phi^{-1}(0,999) \\ \quad + \frac{1}{\sqrt{1-R} \cdot \phi[\Phi^{-1}(PD)]} \left(\frac{R'}{1-R} + \frac{\Phi^{-1}(PD)}{\phi[\Phi^{-1}(PD)]} \right) \end{array} \right.$$



◆ Tableaux

TABLE A.1 – Contributions pour une variation de 0,1% de tous les paramètres sur données simulées

Valeurs initiales				Variation du RW par approche analytique			Contribution des paramètres dans la variation du RW																	
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD2	contrib	PD	LGD	contrib	PD	M	contrib	LGD	M
0.001	0.1	1.0	0.043978	0.000075	0.000075	17.964524	58.898351	23.100761	0.000018	0.017965	-0.001044	0.023101												
0.100	0.1	1.0	0.413990	0.000631	0.000631	30.027823	65.633845	4.316125	0.003003	0.030028	-0.000149	0.004316												
0.001	0.5	1.0	0.219891	0.000373	0.000373	17.964524	58.898351	23.100761	0.000018	0.017965	-0.001044	0.023101												
0.100	0.5	1.0	2.069952	0.003153	0.003154	30.027823	65.633845	4.316125	0.003003	0.030028	-0.000149	0.004316												
0.001	0.1	2.5	0.069852	0.000125	0.000125	9.774709	55.760278	34.423085	0.000010	0.009775	-0.001555	0.034423												
0.100	0.1	2.5	0.454827	0.000711	0.000711	26.438770	63.964905	9.571766	0.002644	0.026439	-0.000330	0.009572												
0.001	0.5	2.5	0.349258	0.000626	0.000626	9.774709	55.760278	34.423085	0.000010	0.009775	-0.001555	0.034423												
0.100	0.5	2.5	2.274135	0.003554	0.003555	26.438770	63.964905	9.571766	0.002644	0.026439	-0.000330	0.009572												

TABLE A.2 – Contributions pour une variation de 10% de tous les paramètres sur données simulées

Valeurs initiales				Variation du RW par approche analytique			Contribution des paramètres dans la variation du RW																	
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD2	contrib	PD	LGD	contrib	PD	M	contrib	LGD	M
0.001	0.1	1.0	0.043978	0.007464	0.007736	17.340268	56.851669	22.298024	0.001734	1.734027	-0.100734	2.229802												
0.100	0.1	1.0	0.413990	0.063062	0.064462	29.381863	64.221926	4.223276	0.293819	2.938186	-0.014562	0.422328												
0.001	0.5	1.0	0.219891	0.037320	0.038678	17.340268	56.851669	22.298024	0.001734	1.734027	-0.100734	2.229802												
0.100	0.5	1.0	2.069952	0.315309	0.322312	29.381863	64.221926	4.223276	0.293819	2.938186	-0.014562	0.422328												
0.001	0.1	2.5	0.069852	0.012522	0.013047	9.385153	53.538035	33.051204	0.000939	0.938515	-0.149313	3.305120												
0.100	0.1	2.5	0.454827	0.071088	0.072834	25.811225	62.446647	9.344573	0.258112	2.581122	-0.032221	0.934457												
0.001	0.5	2.5	0.349258	0.062609	0.065235	9.385153	53.538035	33.051204	0.000939	0.938515	-0.149313	3.305120												
0.100	0.5	2.5	2.274135	0.359441	0.364172	25.811225	62.446647	9.344573	0.258112	2.581122	-0.032221	0.934457												

TABLE A.3 – Contributions pour une variation de 0,1% de tous les paramètres sur données réelles

Valeurs initiales				Variation du RW par approche analytique			Contribution des paramètres dans la variation du RW																	
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD2	contrib	PD	LGD	contrib	PD	M	contrib	LGD	M
0.011000	0.30	4.741713	0.876740	0.001388	0.001388	6.002694	63.167789	30.801180	0.000066	0.006003	-0.004176	0.030801												
0.005015	0.35	1.000000	0.430772	0.000649	0.000649	18.859435	66.339593	14.775815	0.000095	0.018859	-0.001085	0.014776												
0.046123	0.33	1.000000	0.994219	0.001359	0.001360	19.973130	73.129098	6.875319	0.000921	0.019973	-0.001116	0.006875												
0.011000	0.30	5.000000	0.900027	0.001431	0.001431	5.599976	62.877899	31.493419	0.000062	0.005600	-0.004270	0.031493												
0.011000	0.16	5.000000	0.480014	0.000763	0.000763	5.599976	62.877899	31.493419	0.000062	0.005600	-0.004270	0.031493												
0.137042	0.33	1.000000	1.544914	0.002451	0.002452	33.441695	63.007609	3.531922	0.004583	0.033442	0.000030	0.003532												
0.032596	0.33	1.000000	0.878314	0.001169	0.001169	16.679806	75.114953	8.186682	0.000544	0.016680	-0.001592	0.008187												
0.013200	0.30	1.497378	0.625997	0.000887	0.000887	14.108589	70.571554	15.300785	0.000186	0.014109	-0.002340	0.015301												
0.005015	0.30	1.398025	0.401966	0.000620	0.000620	16.551372	64.867628	18.553783	0.000083	0.016551	-0.001362	0.018554												
0.046123	0.30	1.000000	0.903835	0.001236	0.001236	19.973130	73.129098	6.875319	0.000921	0.019973	-0.001116	0.006875												

TABLE A.4 – Contributions pour une variation de 10% de tous les paramètres sur données réelles

Valeurs initiales				Variation du RW par approche analytique			Contribution des paramètres dans la variation du RW																	
PD	LGD	M	RW	delta	RW1	delta	RW2	contrib	PD	contrib	LGD	contrib	M	contrib	PD2	contrib	PD	LGD	contrib	PD	M	contrib	LGD	M
0.011000	0.30	4.741713	0.876740	0.138756	0.142689	5.838894	61.444087	29.960688	0.006423	0.583889	-0.406211	2.996069												
0.005015	0.35	1.000000	0.430772	0.064918	0.066552	18.401144	64.727518	14.416758	0.009228	1.840114	-0.105867	1.441676												
0.046123	0.33	1.000000	0.994219	0.135923	0.138976	19.538823	71.538938	6.725819	0.090119	1.953882	-0.109216	0.672582												
0.011000	0.30	5.000000	0.900027	0.143098	0.147207	5.445228	61.140353	30.623141	0.005990	0.544523	-0.415192	3.062314												
0.011000	0.16	5.000000	0.480014	0.076319	0.078510	5.445228	61.140353	30.623141	0.005990	0.544523	-0.415192	3.062314												
0.137042	0.33	1.000000	1.544914	0.245149	0.249752	32.831484	61.857909	3.467475	0.449931	3.283148	0.002925	0.346747												
0.032596	0.33	1.000000	0.878314	0.116908	0.119078	16.378861	73.759697	8.038974	0.053389	1.637886	-0.156302	0.803897												
0.013200	0.30	1.497378	0.625997	0.088687	0.090379	13.847139	69.263772	15.017242	0.018278	1.384714	-0.229652	1.501724												
0.005015	0.30	1.398025	0.401966	0.061950	0.063637	16.117096	63.165628	18.066968	0.008083	1.611710	-0.132672	1.806697												
0.046123	0.30	1.000000	0.903835	0.123567	0.126342	19.538823	71.538938	6.725819	0.090119	1.953882	-0.109216	0.672582												



TABLE A.5 – Résultats du modèle polynomial (degré 3) sur données simulées

Dep. Variable :	RW	R-squared :	0.998			
Model :	OLS	Adj. R-squared :	0.998			
No. Observations :	192000	Prob (F-statistic) :	0.00			
Df Model :	19	Log-Likelihood :	2.3634e+05			
AIC :	-4.726e+05	BIC :	-4.724e+05			
	coef	std err	t	P > t	[0.025	0.975]
const	-0.1236	0.004	-31.244	0.000	-0.131	-0.116
PD	5.2015	0.032	163.417	0.000	5.139	5.264
LGD	1.2632	0.009	133.410	0.000	1.245	1.282
M	-0.0058	0.004	-1.562	0.118	-0.013	0.001
PD²	-44.5721	0.161	-276.621	0.000	-44.888	-44.256
PD · LGD	31.7366	0.039	813.692	0.000	31.660	31.813
PD · M	0.1033	0.011	9.072	0.000	0.081	0.126
LGD²	-0.0046	0.014	-0.319	0.750	-0.033	0.024
LGD · M	0.3015	0.003	88.344	0.000	0.295	0.308
M²	0.0002	0.001	0.139	0.890	-0.002	0.003
PD³	101.0462	0.316	319.795	0.000	100.427	101.666
PD² · LGD	-62.5939	0.083	-752.169	0.000	-62.757	-62.431
PD² · M	-0.3821	0.021	-18.222	0.000	-0.423	-0.341
PD · LGD²	0.0042	0.025	0.168	0.866	-0.045	0.053
PD · LGD · M	-0.3021	0.006	-53.626	0.000	-0.313	-0.291
PD · M²	0.0018	0.002	1.130	0.258	-0.001	0.005
LGD³	0.0042	0.009	0.490	0.624	-0.013	0.021
LGD² · M	-0.0006	0.002	-0.303	0.762	-0.004	0.003
LGD · M²	-0.0005	0.000	-1.062	0.288	-0.001	0.000
M³	-3.229e-05	0.000	-0.216	0.829	-0.000	0.000



TABLE A.6 – Résultats du modèle polynomial (degré 4) sur données simulées

Dep. Variable :		RW	R-squared :		0.999	
Model :		OLS	Adj. R-squared :		0.999	
No. Observations :		192000	Prob (F-statistic) :		0.00	
Df Model :		34	Log-Likelihood :		2.7607e+05	
AIC :		-5.521e+05	BIC :		-5.517e+05	
	coef	std err	t	P > t	[0.025	0.975]
const	-0.0801	0.008	-9.821	0.000	-0.096	-0.064
PD	5.3957	0.074	72.562	0.000	5.250	5.541
LGD	1.0180	0.022	46.237	0.000	0.975	1.061
M	-0.0014	0.012	-0.121	0.903	-0.024	0.022
PD ²	-82.1759	0.558	-147.311	0.000	-83.269	-81.083
PD · LGD	42.2599	0.131	321.861	0.000	42.003	42.517
PD · M	0.0992	0.046	2.173	0.030	0.010	0.189
LGD ²	-0.0173	0.049	-0.350	0.726	-0.114	0.080
LGD · M	0.2862	0.014	20.963	0.000	0.259	0.313
M ²	-0.0007	0.006	-0.114	0.909	-0.013	0.012
PD ³	430.9202	2.199	195.968	0.000	426.610	435.230
PD ² · LGD	-152.5291	0.510	-298.958	0.000	-153.529	-151.529
PD ² · M	-0.8226	0.142	-5.796	0.000	-1.101	-0.544
PD · LGD ²	0.0108	0.152	0.071	0.943	-0.288	0.310
PD · LGD · M	-0.1137	0.036	-3.183	0.001	-0.184	-0.044
PD · M ²	-0.0004	0.013	-0.032	0.974	-0.026	0.025
LGD ³	0.0092	0.059	0.157	0.875	-0.106	0.124
LGD ² · M	0.0073	0.013	0.571	0.568	-0.018	0.032
LGD · M ²	0.0005	0.004	0.132	0.895	-0.007	0.008
M ³	0.0002	0.002	0.139	0.889	-0.003	0.003
PD ⁴	-725.4002	3.400	-213.366	0.000	-732.064	-718.737
PD ³ · LGD	203.9777	0.890	229.119	0.000	202.233	205.723
PD ³ · M	2.1379	0.224	9.531	0.000	1.698	2.577
PD ² · LGD ²	0.0433	0.262	0.165	0.869	-0.471	0.557
PD ² · LGD · M	-0.7966	0.059	-13.481	0.000	-0.912	-0.681
PD ² · M ²	-0.0223	0.017	-1.306	0.192	-0.056	0.011
PD · LGD ³	-0.0519	0.080	-0.648	0.517	-0.209	0.105
PD · LGD ² · M	0.0162	0.018	0.914	0.361	-0.019	0.051
PD · LGD · M ²	0.0063	0.005	1.375	0.169	-0.003	0.015
PD · M ³	0.0006	0.001	0.428	0.668	-0.002	0.003
LGD ⁴	0.0041	0.028	0.149	0.882	-0.050	0.058
LGD ³ · M	-0.0026	0.006	-0.423	0.672	-0.014	0.009
LGD ² · M ²	-0.0013	0.002	-0.814	0.415	-0.004	0.002
LGD · M ³	-4.737e-05	0.000	-0.112	0.911	-0.001	0.001
M ⁴	-2.571e-05	0.000	-0.184	0.854	-0.000	0.000



TABLE A.7 – Résultats du modèle polynomial (degré 2) sur données réelles

Dep. Variable :	RW	R-squared :	0.974			
Model :	OLS	Adj. R-squared :	0.974			
No. Observations :	8000	Prob (F-statistic) :	0.00			
Df Model :	9	Log-Likelihood :	8755.6			
AIC :	-1.749e+04	BIC :	-1.742e+04			
	coef	std err	t	P> t 	[0.025	0.975]
const	-0.0414	0.003	-13.414	0.000	-0.047	-0.035
PD	3.2369	0.069	47.163	0.000	3.102	3.371
LGD	1.0282	0.059	17.395	0.000	0.912	1.144
M	0.0052	0.002	2.129	0.033	0.000	0.010
PD²	-26.4550	0.278	-95.327	0.000	-26.999	-25.911
PD · LGD	27.2955	0.184	148.524	0.000	26.935	27.656
PD · M	0.0990	0.016	6.368	0.000	0.069	0.130
LGD²	0.4614	0.180	2.557	0.011	0.108	0.815
LGD · M	0.2634	0.004	68.209	0.000	0.256	0.271
M²	-0.0014	0.000	-3.208	0.001	-0.002	-0.001



TABLE A.8 – Résultats du modèle polynomial (degré 3) sur données réelles

Dep. Variable :	RW	R-squared :	0.998			
Model :	OLS	Adj. R-squared :	0.998			
No. Observations :	192000	Prob (F-statistic) :	0.00			
Df Model :	19	Log-Likelihood :	2.3634e+05			
AIC :	-4.726e+05	BIC :	-4.724e+05			
	coef	std err	t	P > t 	[0.025	0.975]
const	-0.0193	0.003	-6.309	0.000	-0.025	-0.013
PD	1.9700	0.111	17.717	0.000	1.752	2.188
LGD	-0.3600	0.214	-1.679	0.093	-0.780	0.060
M	0.0060	0.004	1.348	0.178	-0.003	0.015
PD²	-35.7146	0.887	-40.262	0.000	-37.453	-33.976
PD · LGD	31.7366	0.039	813.692	0.000	31.660	31.813
PD · M	-0.0415	0.056	-0.737	0.461	-0.152	0.069
LGD²	7.1315	1.339	5.327	0.000	4.507	9.756
LGD · M	0.3644	0.034	10.653	0.000	0.297	0.431
M²	-0.0008	0.002	-0.378	0.706	-0.005	0.003
PD³	148.4164	2.459	60.347	0.000	143.595	153.237
PD² · LGD	-150.5767	1.790	-84.113	0.000	-154.086	-147.068
PD² · M	-1.9524	0.130	-14.992	0.000	-2.208	-1.697
PD · LGD²	3.7067	2.785	1.331	0.183	-1.752	9.165
PD · LGD · M	0.7758	0.087	8.944	0.000	0.606	0.946
PD · M²	0.0472	0.009	5.166	0.000	0.029	0.065
LGD³	-10.1377	2.165	-4.682	0.000	-14.382	-5.893
LGD² · M	-0.3846	0.093	-4.130	0.000	-0.567	-0.202
LGD · M²	-0.0002	0.002	-0.076	0.939	-0.005	0.005
M³	-0.0002	0.000	-0.631	0.528	-0.001	0.000



TABLE A.9 – Résultats du modèle polynomial (degré 4) sur données réelles

Dep. Variable :	RW	R-squared :	0.992			
Model :	OLS	Adj. R-squared :	0.992			
No. Observations :	8000	Prob (F-statistic) :	0.00			
Df Model :	34	Log-Likelihood :	13329			
AIC :	-2.659e+04	BIC :	-2.634e+04			
	coef	std err	t	P > t	[0.025	0.975]
const	-0.0121	0.003	-3.594	0.000	-0.019	-0.005
PD	1.5606	0.168	9.274	0.000	1.231	1.890
LGD	-2.7983	0.701	-3.992	0.000	-4.173	-1.424
M	0.0169	0.007	2.406	0.016	0.003	0.031
PD ²	-51.1250	2.187	-23.373	0.000	-55.413	-46.837
PD · LGD	103.7956	4.842	21.438	0.000	94.304	113.287
PD · M	0.1303	0.139	0.935	0.350	-0.143	0.404
LGD ²	32.3456	7.208	4.487	0.000	18.215	46.476
LGD · M	0.3063	0.169	1.816	0.069	-0.024	0.637
M ²	-0.0144	0.006	-2.548	0.011	-0.025	-0.003
PD ³	454.8052	12.047	37.751	0.000	431.189	478.421
PD ² · LGD	-435.7808	9.900	-44.017	0.000	-455.188	-416.374
PD ² · M	-0.7000	0.679	-1.031	0.303	-2.031	0.631
PD · LGD ²	-237.9483	28.169	-8.447	0.000	-293.166	-182.730
PD · LGD · M	0.2422	0.754	0.321	0.748	-1.236	1.721
PD · M ²	-0.1068	0.051	-2.085	0.037	-0.207	-0.006
LGD ³	-102.3996	25.154	-4.071	0.000	-151.708	-53.092
LGD ² · M	0.1566	0.840	0.186	0.852	-1.490	1.803
LGD · M ²	0.0005	0.026	0.018	0.985	-0.051	0.052
M ³	0.0047	0.002	2.724	0.006	0.001	0.008
PD ⁴	-1230.0922	23.301	-52.792	0.000	-1275.768	-1184.416
PD ³ · LGD	1026.1880	17.802	57.645	0.000	991.292	1061.084
PD ³ · M	24.2050	1.502	16.113	0.000	21.260	27.150
PD ² · LGD ²	59.4996	27.013	2.203	0.028	6.546	112.453
PD ² · LGD · M	-16.1514	1.028	-15.711	0.000	-18.167	-14.136
PD ² · M ²	-0.8160	0.090	-9.112	0.000	-0.992	-0.640
PD · LGD ³	359.1142	44.173	8.130	0.000	272.524	445.705
PD · LGD ² · M	1.8338	1.985	0.924	0.356	-2.057	5.725
PD · LGD · M ²	0.3008	0.057	5.278	0.000	0.189	0.413
PD · M ³	0.0255	0.006	4.049	0.000	0.013	0.038
LGD ⁴	109.9971	29.204	3.767	0.000	52.750	167.244
LGD ³ · M	-0.9591	1.240	-0.773	0.439	-3.390	1.472
LGD ² · M ²	0.0133	0.064	0.207	0.836	-0.113	0.139
LGD · M ³	-0.0022	0.002	-1.346	0.178	-0.005	0.001
M ⁴	-0.0005	0.000	-2.956	0.003	-0.001	-0.000



◆ Figures

FIGURE A.1 – Matrice de corrélation des données simulées

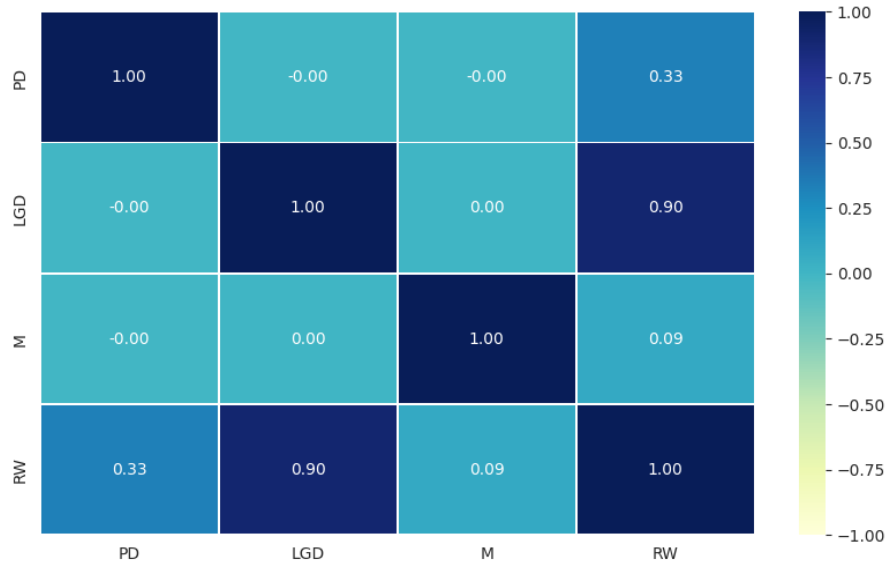


FIGURE A.2 – Matrice de relation des données simulées

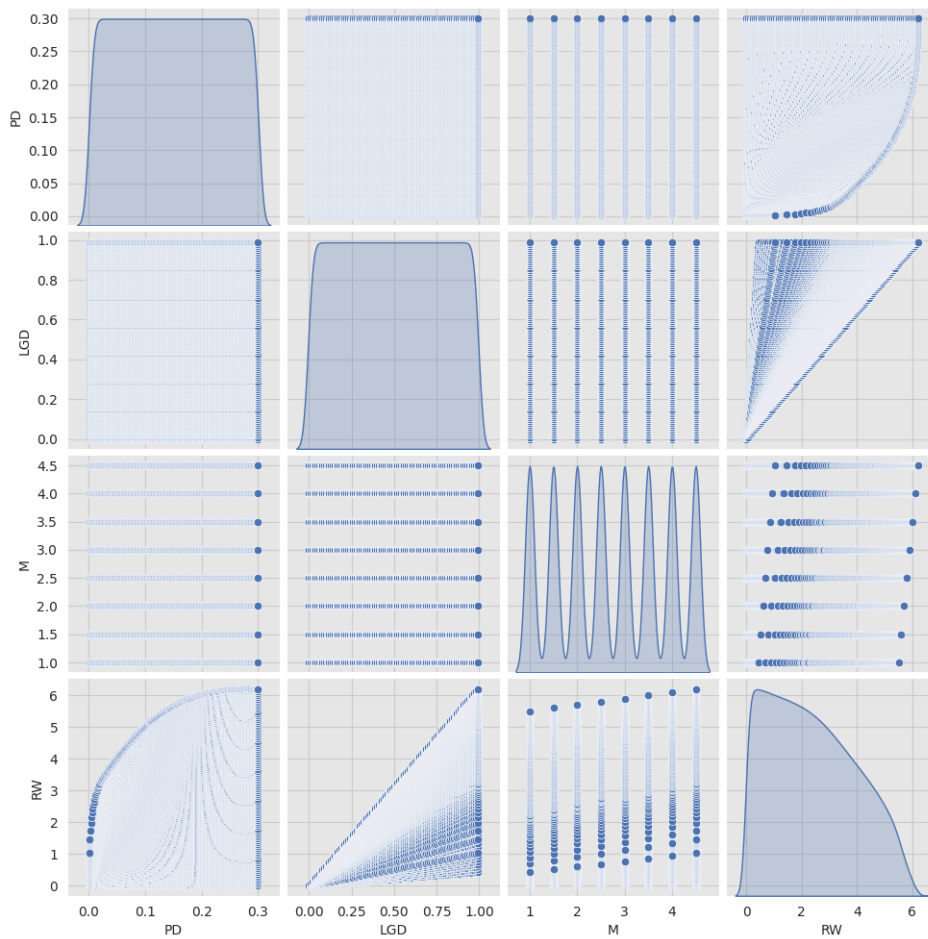




FIGURE A.3 – ALE sur données réelles (Random Forest)

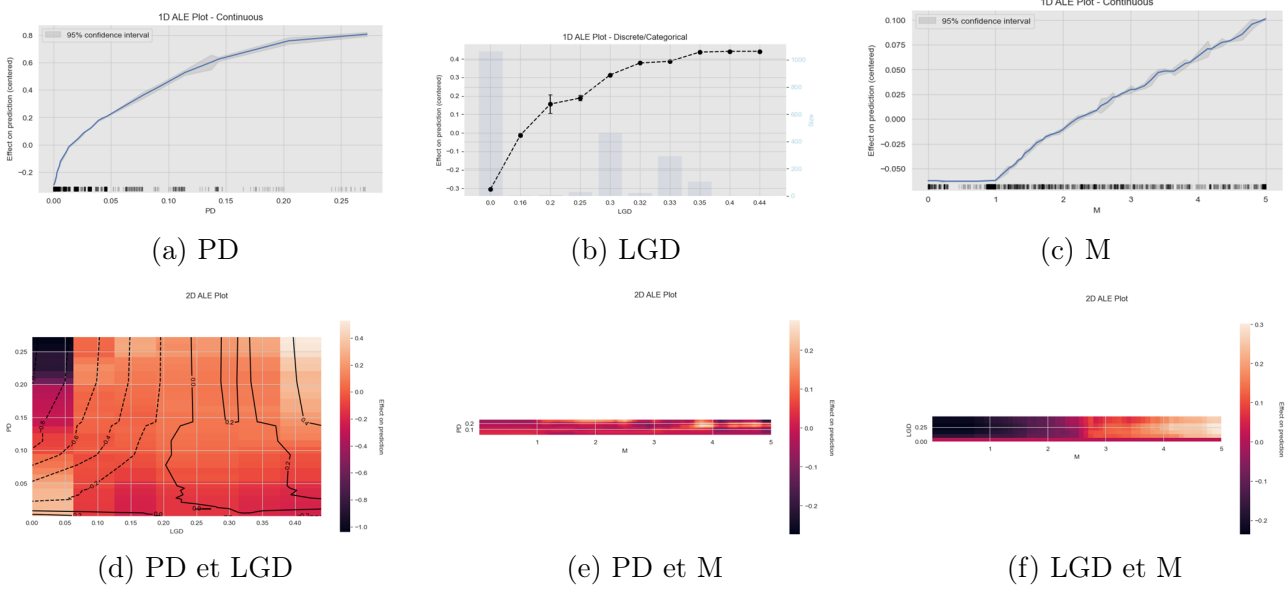


FIGURE A.4 – Modèles de machine learning sur données simulées : Prédiction vs Données

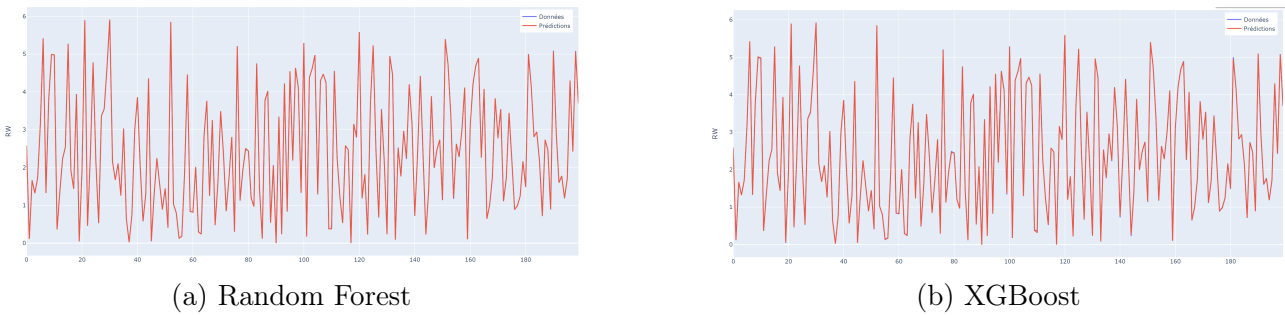


FIGURE A.5 – Modèles de machine learning sur données réelles : Prédiction vs Données

