

# DE L'IMPORTANCE DES MODÈLES EN ASSURANCE ET DE LEUR FIABILITÉ À L'HEURE DES STRESS-TESTS CLIMATIQUES ET DE LA PLACE CONSIDÉRABLE DE L'APPRENTISSAGE STATISTIQUE

**MOTS CLES : MODELES, REVUE, STRESS TESTS CLIMATIQUES, HYPOTHESES, DONNEES, PROJECTIONS**

## INTRODUCTION

Les modèles constituent un outil central des processus prévisionnels et décisionnels des organismes assureurs. Pour s'assurer qu'ils restent adaptés à la réalité de l'activité de ces derniers, à une réglementation inflationniste, aux pratiques du marché et aux évolutions technologiques, ils doivent faire l'objet d'audits réguliers.

Dans un environnement incertain où de nouveaux risques émergent, où le nombre de données à disposition est en forte croissance, et où les moyens de traiter ce nombre sont de plus en plus répandus, les modèles s'intègrent désormais au cœur des processus prévisionnels et décisionnels des organismes assureurs. Ils apparaissent comme un outil central pour de nombreux exercices stratégiques tels que la construction d'un Business Plan, le lancement de produits et/ou de partenariats ou encore l'ajustement des allocations d'actifs.

### Des revues de modèles de plus en plus fréquentes

Pour assurer la fiabilité et l'efficacité du modèle, et alimenter les échanges avec le régulateur ou l'auditeur interne, les revues de modèles sont de plus en plus fréquentes. Ces revues doivent permettre de rendre compte des domaines privilégiés d'utilisation des modèles ainsi que leurs limites, et de proposer une démarche pour en améliorer l'implémentation par les équipes opérationnelles et l'appropriation par les décideurs. Au-delà de l'adéquation avec la réglementation, l'objectif de la revue est de tendre vers un outil de pilotage conduisant à la meilleure prise de décision, comme le requiert la directive Solvabilité 2 à propos des travaux sur la gestion des risques.

En outre, la matérialisation de nouveaux risques, tels que les risques de durabilité (environnementaux et sociaux), nécessite de faire évoluer les modèles pour intégrer cette nouvelle dimension impactant l'activité d'assurance, tant sur les engagements pris que sur les investissements. Ainsi, la survenance plus fréquente d'événements climatiques extrêmes, ou encore l'évolution de la valeur des actifs financiers liée à la transition écologique, élargissent les données et hypothèses à intégrer, non seulement pour améliorer la gestion du risque, mais aussi pour s'inscrire dans le nouveau cadre réglementaire et les demandes du superviseur. Les modèles d'apprentissage statistique, ont également des caractéristiques distinctes qui doivent être prises en compte spécifiquement lors de la revue des modèles.

La revue d'un modèle s'articule autour de nombreuses thématiques, les principales sont les données d'entrée, les méthodes et les hypothèses implémentées, les tests de validation effectués, l'utilisation du modèle et sa documentation.

### Les données d'entrée

Les données en entrée du modèle constituent un axe important de la revue. Il s'agit de vérifier qu'elles sont bien documentées, que leur accès et leur qualité sont robustes dans le temps, et qu'elles sont traçables (dictionnaire de données, schémas de cheminement des données, etc.). Il est également important d'analyser leur gouvernance, c'est-à-dire les rôles et responsabilités des opérationnels fournisseurs de ces données.

La qualité des données est évaluée selon plusieurs critères. Par exemple, dans le cadre des évaluations prudentielles soumises à Solvabilité 2, l'exactitude, l'exhaustivité et le caractère approprié seront les critères à analyser. Il s'agit

aussi de vérifier la cohérence temporelle en effectuant une analyse des variations des données à fréquence régulière.

Concrètement, les démarches menées dans ce cadre pourront porter sur les points suivants :

- // le traçage de la vie de la donnée, de l'opération de vente, à l'utilisation dans les modèles, en passant par son évolution dans les systèmes d'information au gré des migrations informatiques ;
- // la réalisation de tests de cohérence sur les données (cohérence des dates, des montants garantis avec les données contrat par exemple) ;
- // le suivi du processus assurant la qualité des données : tableau de bord des erreurs détectées, notes de procédures décrivant les contrôles à effectuer notamment.

Par ailleurs, **au-delà des données historiques de l'organisme assureur, la modélisation de certains risques tels que le dérèglement climatique, ou l'utilisation de techniques de Data Science, permettent aujourd'hui d'élargir le spectre des données utilisées.** Le caractère relativement nouveau du risque climatique dans le pilotage de l'activité d'assurance, ainsi que les évolutions attendues dans le futur, non reflétées sur l'historique, nécessitent d'enrichir les données de l'assureur avec des informations extérieures aux systèmes d'informations. Certaines sont en *open data*, elles peuvent être structurées ou non (données Météo France<sup>1</sup>, rapports GIEC et autres organismes). Le contrôle de la qualité de ces données externes *a priori* non maîtrisées et retraitées apparaît alors d'autant plus nécessaire pour être en mesure de les exploiter efficacement.

### Les méthodes et hypothèses

Concernant les hypothèses, il s'agit de vérifier leur stabilité dans le temps, leur mise à jour

régulière et leurs limites de validité, afin d'éviter toute dérive de modèle (notion de *model drift* particulièrement suivie dans les modèles de *machine learning*). Elles peuvent être calibrées à partir de données d'expérience, de bases réglementaires, de données du marché ou de jugements d'experts. **La crise écologique en cours est de nature à remettre en cause les tendances passées et in fine les modalités de fixation des hypothèses.**

Concernant le choix des méthodes, la complexité et l'implémentation opérationnelle constituent des enjeux cruciaux de la revue, de même que la validation des approximations effectuées, d'autant plus quand les modèles sont d'une grande complexité, parfois telle qu'ils sont qualifiés de « boîte noire ». Pour les hypothèses et les méthodes, une attention particulière doit être portée sur la cohérence avec les éléments contractuels et réglementaires et/ou les pratiques du marché.

Des études statistiques peuvent ainsi être menées pour apprécier l'adéquation des hypothèses avec le risque sous-jacent. Par exemple, une analyse des variables explicatives du risque peut être conduite : étude des corrélations entre les variables, régression logistique et test de significativité des variables. Dans le cadre de la revue d'un modèle de gestion actif/passif (ALM), l'auditeur peut par exemple mettre en parallèle les règles de gestion internes (traduisant l'appétence au risque et les limites de risques) en termes de placements et les stratégies financières effectives implémentées dans le modèle. L'auditeur peut également être amené à comparer les clauses contractuelles et les *management actions* et l'implémentation effective dans le modèle. Les contraintes réglementaires font également l'objet d'une analyse par l'auditeur (par exemple : frontière des contrats sous Solvabilité 2, plafond de la provision d'égalisation du code des assurance).

---

<sup>1</sup> Dans une série de notes consacrées aux données climatiques disponibles en open data, le

DataLab de GALEA présente les données [Météo France](#) et [Drias](#).

L'émergence de nouveaux risques nécessite d'ajouter de nouvelles dimensions concernant les hypothèses et la méthode. Par exemple, la modélisation du risque climatique nécessite de repenser :

- // L'horizon de projection, en cohérence avec une dynamique d'évolution climatique de long terme ;
- // La liste des hypothèses en entrée, par exemple en ce qui concerne la valorisation des actifs, qui peuvent désormais intégrer les données issues de projections pour l'étude du risque de transition ;
- // Les nouvelles management actions, liées à de nouveaux comportements des dirigeants, par exemple, d'éventuelles résiliations de contrats ou de nouvelles pratiques de souscription et de revalorisation des primes en cas de déviation structurelle de certains portefeuilles ;
- // La structure du modèle, qui peut nécessiter d'intégrer des modules intermédiaires, élaborés par des experts tels que des climatologues, permettant de retranscrire et d'adapter des trajectoires climatiques de référence (scenarios du GIEC et déclinaisons NGFS) aux enjeux financiers et organisationnels pour l'organisme assureur.

### Les tests de validation

Plusieurs tests doivent être effectués pour valider un modèle, recensés dans une liste aussi exhaustive que possible. Différentes techniques peuvent être utilisées : le *backtesting*, l'analyse de cas tests, la réalisation de sensibilités ou « scénarios stressés », l'analyse discriminante ou validation croisée, etc. Pour valider les tests menés, il convient également de définir des seuils de tolérance au-delà desquels il sera considéré que le modèle n'est plus valide.

Par exemple, dans les modèles classiques, des cas tests peuvent être réalisés sur des calculs de provisions individuelles (par exemple, PM/PRC<sup>2</sup> en emprunteur), permettant de s'assurer d'une méthodologie cohérente avec les contraintes réglementaires et les pratiques du marché ; une revue du calcul du taux de rendement moyen des actifs peut être menée dans une projection risque-neutre avec détection d'erreurs sur le rendement de certains actifs, à partir d'indicateurs statistiques (moyenne, médiane, quantile, boîtes à moustache, ...) ; l'application de tests de martingalité des prix actualisés des actifs projetés permettant la validation de GSE risque-neutre ; ou une revue du nombre minimal de simulations d'un modèle ALM assurant sa stabilité, en étudiant la convergence et la distribution du *Best Estimate* au regard d'un benchmark de marché. Dans certains cas, des modèles simplifiés (« *challenger models* ») peuvent être élaborés pour tester plus grossièrement les résultats produits.

En apprentissage statistique, les techniques (train-validation-test ou validation croisée-test) et le choix des métriques (F1 score, AUC, et tant d'autres...) pour valider l'apprentissage sont d'une importance capitale.

Dans le cas des enjeux climatiques, les tests de validation s'avèrent d'autant plus importants qu'à l'heure actuelle les modèles doivent s'adapter en permanence à un environnement en évolution rapide.

### L'utilisation du modèle

L'utilisation d'un modèle fait également partie du périmètre de la revue. Un avis et des recommandations sont donnés sur l'ergonomie du modèle, la lecture des résultats, la facilité de paramétrage, ou encore l'utilisation transversale du modèle pour différents usages, justifiant son rôle central dans la prise de décision.

La revue de l'utilisation du modèle peut mettre en avant des améliorations concernant

---

<sup>2</sup> Provision mathématique / provision pour risque croissant

l'automatisation sur l'importation des données ou la production des résultats. Dans un contexte multinormes (droit français, Solvabilité 2, IFRS), l'optimisation des résultats pour alimenter les différents reportings réglementaires peut par exemple être analysée.

Une attention particulière est portée sur l'information communiquée aux **dirigeants** et aux **administrateurs** sur les modèles utilisés pour bâtir les éléments de *reporting* à partir desquels ils prendront leurs décisions. Des **formations spécifiques** sont parfois nécessaires compte tenu de la sophistication des modélisations intervenue ces dernières années, souvent sous l'impulsion de la réglementation (Solvabilité 2 et Bâle 3).

**L'évolution des modèles pour intégrer les risques de durabilité et climatiques nécessiteront des échanges réguliers avec les dirigeants et les administrateurs des organismes.**

#### Documentation

La documentation d'un modèle (notes techniques et guides d'utilisateur) doit être exhaustive, claire et disponible. La documentation la plus récente est à partager avec l'ensemble des utilisateurs du modèle. Les rôles et responsabilités pour la modification et la validation de la documentation doivent être correctement définis. Par ailleurs, plus la documentation est claire et précise, plus l'utilisation du modèle est facile, les résultats compréhensibles et le risque opérationnel réduit.

La revue d'un modèle est également l'occasion de donner un avis sur la gouvernance du modèle (processus de suivi du modèle, instances de validation et sujets abordés, indépendance des acteurs, fréquence des audits, ...).

La documentation pour les modèles d'apprentissage statistique se révèle cruciale, elle doit inclure les éléments permettant de rejouer le modèle sur l'ensemble de la chaîne de valeur de la donnée. Une partie est intégrée au code directement, sous chaque fonction. La

**capacité de reprise d'un code d'une personne tierce est limitée à la qualité de la documentation.**

#### Une action constructive

La revue de modèle doit être une action constructive et à forte valeur ajoutée, ancrée dans les actualités internes et externes qui impactent l'organisme. Ainsi, la démarche doit s'attacher au principe de proportionnalité, c'est-à-dire que l'ensemble des recommandations émises doivent mettre en parallèle les impacts sur le résultat, au regard du coût de mise en œuvre et de la nécessaire conformité.

Selon le besoin et le contexte, plusieurs niveaux de revue sont possibles : de la revue globale permettant une analyse macro du modèle et de son utilisation, à la revue approfondie sur les hypothèses, méthodes et résultats.

**Dans le cas des modèles spécifiquement développés pour anticiper et suivre les enjeux climatiques, la revue du modèle devra offrir un niveau de sécurité raisonnable aux dirigeants pour engager des actions constructives, sur des bases suffisamment robustes.**

<https://www.galea-associes.eu/>

//galea