

MPG Overview

Le Machine Learning est-il une solution envisageable pour les Reverse Stress tests ?

Contributeurs :

Cheryf Olayimika Laleyé
Senior Consultant

Christophe Bertet
Senior Advisor

Matthias Poirier
Founding Partner

Sur la base des travaux de :

Jean-Marc Mercier, PhD
Senior Research Advisor

Shohruh Miryusupov, PhD
Research Advisor

Depuis 2008, les stress tests n'ont cessé de s'imposer dans le pilotage des risques (marché, crédit, liquidité, contrepartie...), sur la base de scénarios historiques ou hypothétiques.

Bien que la plupart des banques aient implémenté aujourd'hui un cadre robuste de stress testing, les dispositifs actuels échouent encore à identifier de façon exhaustive tous les scénarios plausibles pouvant conduire à une perte colossale ou à un défaut.

Le Reverse Stress Test permet de combler ce défaut majeur et intervient comme un outil complémentaire au cadre de stress test traditionnel.

En pratique, la réalisation d'un test de résistance inversé présente de nombreuses difficultés en particulier sur les aspects méthodologiques, à savoir l'inversement de la fonction permettant le calcul des niveaux de pertes à partir des scénarios et le choix d'un scénario cohérent et plausible parmi ceux identifiés.

Ces problèmes peuvent être adressés par CoDPy, notre solution de Machine Learning qui fournit des résultats **interprétables** et **reproductibles**. Ce document résume les enjeux et les étapes clés nécessaires pour l'implémentation des tests de résistance inversés, notre solution, et nos résultats obtenus après une étude sur des portefeuilles linéaires et non-linéaires.

Le reverse stress test, un outil indispensable pour une gestion optimale des risques

Un outil complémentaire au stress testing

Bien que robustes et offrant une bonne couverture des risques extrêmes, les stress tests continuent de poser un problème d'exhaustivité dans l'estimation des risques à l'échelle de l'entreprise. Un stress test part d'un scénario construit *a priori*, c'est-à-dire à partir d'hypothèses faites par l'établissement, qui va ensuite chercher à en mesurer l'impact. Cet impact ne sera pas forcément matériel à son échelle, et donc échouera à identifier les scénarios conduisant à des pertes mettant en péril l'établissement.

Les tests de résistance inversés, quant à eux, sont basés sur un montant de perte défini à l'avance, donc nécessairement pertinent. A ce montant sera associé le scénario le plus plausible, via des méthodes quantitatives.

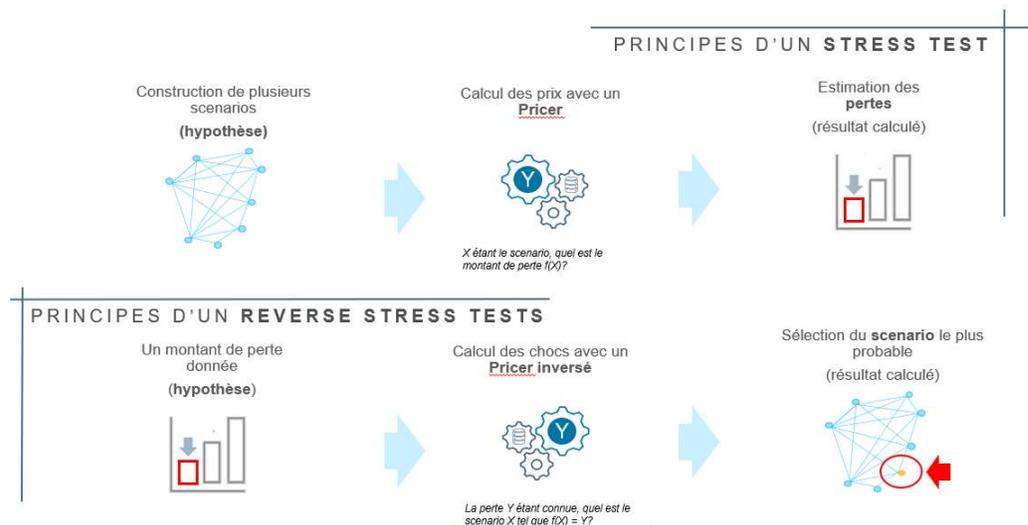


Fig. 1 : Stress Test et Reverse Stress Test, principes de fonctionnement, source : MPG Partners

Les Reverse Stress Tests ont donc tout leur intérêt en mettant en lumière des scénarios non identifiés *a priori*.

Les régulateurs recommandent le reverse stress testing dans le dispositif global de stress test

Les premières mentions de la nécessité de mettre en place le reverse stress testing par les institutions financières font suite à la crise financière et bancaire de 2008. Les principales instances de régulation ont progressivement publié des recommandations et lignes directrices mettant en exergue le caractère fondamental des tests de résistance inversés. Ils présentent la particularité d'identifier des scénarios de risques

beaucoup plus sévères que ceux définis *a priori* dans les scénarios de stress tests traditionnels.

Plus récemment, le « **Final Report on Guidelines on institutions' stress testing** » de l'EBA (réf. EBA/GL/2018/04) consacre les reverse stress tests comme un **outil indispensable pour le pilotage des risques, sans pour autant indiquer comment ils doivent être réalisés**. Bien qu'elles aient aujourd'hui pris conscience des avantages du reverse stress testing, les banques éprouvent néanmoins des difficultés quant à son implémentation, en particulier sur les aspects méthodologiques.

Le reverse stress test : un exercice difficile

Une démarche en cinq étapes...

1. **Calibration des pertes** : L'objectif fixé (considérer une faillite de la banque, la perte d'une année de de résultat, un seuil de pertes arbitraire, etc) doit permettre de définir précisément l'hypothèse de perte
2. **Périmètre et risques captés** : Le choix du périmètre permet de définir les risques ciblés par l'exercice (crédit, marché, opérationnel, liquidité). Un nombre important de facteurs de risques peut également entraîner des problèmes de calcul (problèmes de dimension...).
3. **Méthodologie** : Les formules fermées ne sont pas toujours inversibles. L'inversion d'une fonction de prix par Monte Carlo n'est pas possible dans la plupart des cas. Enfin, en partant d'un niveau de perte donnée, il existe une infinité de scénarios possibles. La sélection du plus pertinent et plausible doit passer par une méthodologie robuste.
4. **Analyse des résultats** : Il faut être en mesure de comprendre, expliquer et certifier le scénario obtenu, et de s'assurer de sa pertinence pour le périmètre considéré.
5. **Intégration, suivi et gestion des scénarios de risques identifiés** : il est important de se poser la question des potentielles vulnérabilités mises en avant, et les raisons pour lesquelles elles n'ont pas toujours été identifiées jusque-là.

... qui présente de nombreuses difficultés, notamment méthodologiques

De manière pratique, la mise en place d'un test de résistance inversé repose sur un nombre important d'hypothèses simplificatrices qui altèrent la qualité des résultats.

Quelques hypothèses simplificatrices :

- L'utilisation de données historiques est la première des simplifications. Elle constitue un argument fort pour justifier du réalisme des hypothèses ;
- Sur certains périmètres, une approche par scénario est complexe à mettre en œuvre. Une analyse par les sensibilités apparaît alors comme une solution ;

- Lorsque le nombre de facteurs de risque à considérer est trop grand, une diminution de leur nombre est nécessaire (ex : via une Analyse en Composante Principale) ;
- Compte tenu de la lourdeur des process, les calculs sont souvent réalisés à une fréquence peu élevée.

En simplifiant l'exercice, les établissements perdent certaines informations essentielles à savoir :

- Certains effets (effets croisés, corrélations), importants en période de stress, ne sont pas captés ;
- L'analyse des sensibilités de 1er ordre, en ne prenant pas en compte la convexité, ne permet pas d'estimer tous les risques matériels

La méthodologie optimale est celle qui minimise le recours aux hypothèses simplificatrices, tout en intégrant les contraintes opérationnelles. C'est l'objectif de notre solution.

CoDPy, une solution basée sur le Machine Learning

CoDPy présente l'avantage d'adresser l'ensemble des problèmes rencontrés

Parmi les travaux à réaliser, deux problématiques majeures doivent être adressées :

1. **Comment inverser les fonctions de calcul du P&L ?**
2. **Comment établir une méthodologie robuste permettant le choix du scénario le plus plausible ?**

Le Machine Learning et sa méthodologie itérative dans la recherche de fonctions de décision et de solutions optimales, permet d'adresser ces problèmes. Les approches de Machine Learning habituelles obligent à faire un compromis entre **interprétabilité des résultats et pouvoir prédictif**, ce qui les rend moins pertinentes pour le calcul et l'évaluation des risques dans le contexte des banques et autres institutions financières.

Notre solution de Machine Learning CoDPy développée en Python (cf. [4] en annexe) fournit des résultats **interprétables** et **reproductibles** même avec une profondeur de données historiques limitée. Elle permet d'inverser la fonction prix et de sélectionner le scénario le plus plausible à l'origine de l'hypothèse de pertes. Elle répond donc aux problématiques identifiées et permet la réalisation d'un reverse stress test selon la démarche classique de Machine Learning :

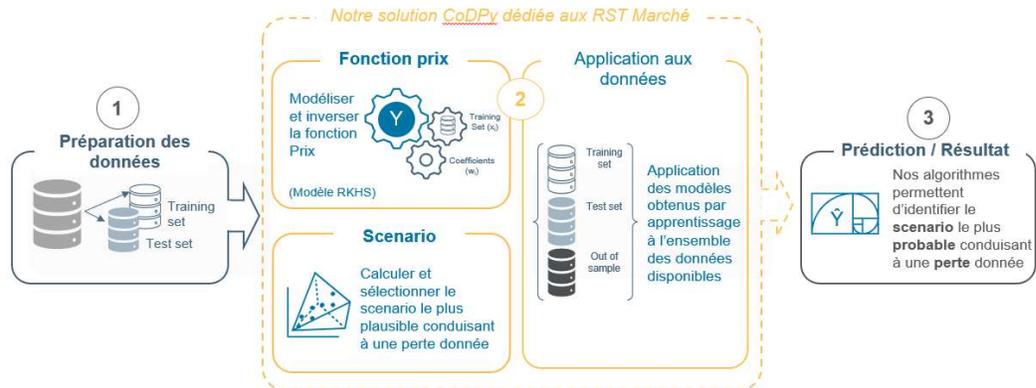


Fig. 2 : Les trois étapes de la méthodologie CoDPy, source : MPG Partners

Notre méthodologie en trois étapes

1. Sourcing et préparation des données

La *phase apprentissage* de notre solution consiste à construire un ensemble d'entraînement. Ce sont ces données qui vont permettre à la machine de déterminer une solution optimale à généraliser à partir de scénarios et P&L historiques. L'ensemble d'apprentissage est constitué de scénarios ($\Delta RF_1, \Delta RF_2, \dots, \Delta RF_{n-1}$) déjà calculés (scénarios de VaR, sVaR et Stress Tests), et de leurs P&L associés, sur les positions en $j-1$, mais aussi sur les positions en $j-2, j-3$, etc.

variable expliquée		variables explicatives				
Scenario	P&L	RiskFactor RF1	RiskFactor RF2	RiskFactor RF3	...	RiskFactor RF _{n-1}
scenario 1	P&L (1)	$\Delta RF_1 (1)$	$\Delta RF_2 (1)$	$\Delta RF_3 (1)$...	$\Delta RF_{n-1} (1)$
scenario 2	P&L (2)	$\Delta RF_1 (2)$	$\Delta RF_2 (2)$	$\Delta RF_3 (2)$...	$\Delta RF_{n-1} (2)$
scenario 3	P&L (3)	$\Delta RF_1 (3)$	$\Delta RF_2 (3)$	$\Delta RF_3 (3)$...	$\Delta RF_{n-1} (3)$
...
scenario m	P&L (m)	$\Delta RF_1 (m)$	$\Delta RF_2 (m)$	$\Delta RF_3 (m)$...	$\Delta RF_{n-1} (m)$

Un scenario k se définit par des mouvements de données de marché, notés $\Delta RF_i(k)$, correspondant au choc appliqué sur le facteur de risque RF_i , dans le scenario k . Chaque scenario k correspond à un unique P&L(k).

Les facteurs de risques (ou variables explicatives) sont, par exemple : indices d'actions, taux d'intérêts, spreads de crédit, etc.... Pour la variable expliquée, nous pouvons par exemple considérer le P&L d'un book de trading.

2. Inversion des fonctions de prix

La démarche consistant à inverser des fonctions de P&L n'est pas usuelle pour les banques. Elles ont en effet pour habitude de calculer les impacts de scénarios sur leurs résultats et/ou bilan et non l'inverse.

Une fois les données d'apprentissage acquises, CoDPy va utiliser son algorithme (RKHS : *Reproducing Kernel Hilbert Space*) pour estimer le P&L à partir de la matrice de Gram d'un noyau reproduisant k . Les valeurs x, y et z représentent respectivement :

x : N-uplets constitués de variations de facteurs risques obtenues historiquement ou générées par chocs,

y : un ensemble de paramètres internes au modèle (le cas le plus fréquent est $y = x$)

z : scénarios implicites par l'algorithme de Machine Learning.

La machine d'apprentissage f_z se définit comme suit :

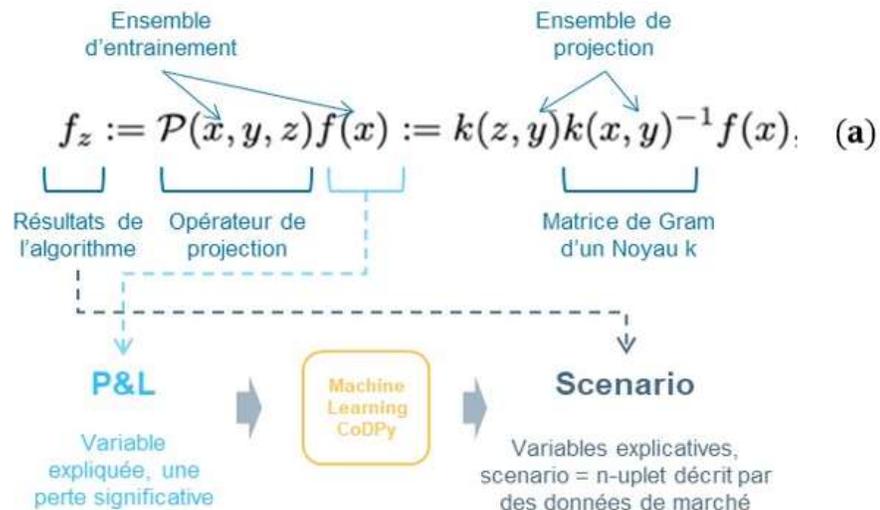


Fig. 3 : La fonction ou machine d'apprentissage RKHS de CoDPy, source : MPG Partners

3. Choix du scénario le plus plausible

En partant d'un montant de perte cible, on débouche sur une infinité de scénarios i.e. de combinaisons des différents facteurs de risques.

Afin de déterminer le scénario le plus « **plausible** », notre solution repose par construction, sur la notion de « **distance** » entre les données d'apprentissage et les scénarios déterminés par notre algorithme de Machine Learning.

Considérant un échantillon $x = (x^1, \dots, x^N)$ indépendantes et identiquement distribuées d'une variable aléatoire X , un ensemble $y = (y^1, \dots, y^M)$ et une mesure de distance quelconques, $d(y, x)$, on définit $p(y, x)$ comme une mesure de la plausibilité d'un scénario de la manière suivante :

$$p(y, x) := \frac{E^X(d(y, \cdot) | d(y, \cdot) \geq d(y, x))}{E^X(d(y, \cdot))} \quad (b)$$

Dans le cadre du reverse stress testing, x sont par exemple les historiques de facteurs de risques constituant notre ensemble d'apprentissage et y le scénario déterminé par l'algorithme de machine Learning. La mesure de distance $d(x, y)$ est quelconque (par exemple la « **distance de Mahalanobis** », **classiquement utilisée en reverse stress testing**).

Ici, nous faisons le choix d'une distance appelée « **erreur de discrédance** » notée d_k (cf. [1] en annexe). Cette mesure a été introduite initialement dans la publication référencée [2] en annexe sous le nom de « discrédance moyenne maximum » et, est naturellement induite par les méthodes à noyaux sur lesquelles se base l'algorithme d'apprentissage.

Définie comme suit, elle présente l'avantage d'être précise, indépendante d'une hypothèse quelconque de distribution (la « distance de Mahalanobis » fait l'hypothèse d'une distribution de type gaussienne) et permet d'évaluer la distance entre deux nuages de points x et y :

$$d_k(x, y)^2 := \frac{1}{N_x^2} \sum_{n=1, m=1}^{N_x, N_x} k(x^n, x^m) + \frac{1}{N_y^2} \sum_{n=1, m=1}^{N_y, N_y} k(y^n, y^m) - \frac{2}{N_x N_y} \sum_{n=1, m=1}^{N_x, N_y} k(x^n, y^m),$$

Où k est un noyau quelconque,

N_x et N_y sont les cardinaux des ensembles x et y .

Ce choix est motivé par l'expression suivante (cf. [3] en annexe) qui permet de calculer l'erreur commise par la machine d'apprentissage **(a)** :

$$\|f(z) - f_z\| \leq d_k(x, y, z) \|f\|_{H_k}, \text{ où :}$$

$d_k(x, y, z)$ est l'erreur de discrédance

$\|f\|_{H_k}$ est la norme de f , $\|\cdot\|_{H_k}$ étant déduite du noyau k

Cette formule s'applique aux machines d'apprentissage définies par **(a)**, et s'interprète naturellement en statistique comme un indicateur de vraisemblance. Cette observation rend son utilisation dans la formule **(b)** très efficace.

Use cases et résultats numériques

Afin de nous assurer de la justesse de notre méthodologie, nous avons considéré deux cas basiques mais représentant une large part des cas rencontrés.

A. Démarche

Nous avons testé notre méthodologie sur les deux types possibles de portefeuille : **linéaire et non linéaire**. Pour les besoins de nos tests, nous avons considéré des portefeuilles simples, avec peu de sous-jacents. Ces deux types d'exposition au sous-jacent (linéaire et non linéaire) sont représentatives de la plupart des positions dans un book de trading, ou le portefeuille d'un fonds.

B. Hypothèses et objectifs

Hypothèses	Objectifs
<ul style="list-style-type: none"> Les données d'apprentissage que nous utilisons sont constituées d'historiques quotidiens de market data sur les cinq dernières années et des fonctions de calcul de P&L connues, permettant de calculer les P&L 10 jours de 2015 à 2020 L'horizon de temps retenu est de 10 jours, en ligne avec les stress tests marché de plusieurs BFI en France Les fonctions de P&L sont supposées connues 	<ul style="list-style-type: none"> Vérifier la pertinence de la fonction de P&L déduite des données d'apprentissage. Nous la quantifions grâce à la notion « d'erreur relative » Examiner les résultats obtenus en reverse stress tests (scenario le plus plausible) et s'assurer de leur cohérence

C. Use case n°1 : portefeuille linéaire

i. Description du portefeuille et hypothèses

Nous considérons un portefeuille **linéaire sensible à trois facteurs de risque** avec les expositions suivantes :

- Spot Eurostoxx50, noté S1, avec $\text{delta1} = 1 \text{ Meur}$
- Spot EUR/USD, noté S2, avec $\text{delta2} = 1 \text{ M€}$
- Euribor3mois, noté S3, avec $\text{delta3} = -100 \text{ K€}/\text{bp}$

La fonction de P&L associée est donc connue, le P&L entre les dates t et t' s'écrit:

$$(Eq1) \ P\&L(t, t') = \text{delta1} \times \left(\frac{s1(t') - s1(t)}{s1(t)} \right) + \text{delta2} \times \left(\frac{s2(t') - s2(t)}{s2(t)} \right) + \text{delta3} \times (S3(t') - S3(t))$$

ii. Description des données d'apprentissage

L'ensemble d'apprentissage est constitué de quadruplets constitués du P&L et des variations 10 jours de l'Eurostoxx50, l'Euribor 3 mois (absolu) et l'EUR/USD

Variable à expliquer	Variables explicatives		
	P&L 10 jour en EUR	Eurostoxx50	Euribor3 mois
68 217,82	6,28%	-0,1	-0,45%
13 148,85	2,02%	-0,1	-1,70%
-3 355,99	-2,15%	-0,1	0,82%
9 826,27	-1,47%	-0,2	0,45%
19 400,40	-1,33%	-0,3	0,27%
19 402,53	-2,52%	-0,4	0,46%
37 644,06	-1,90%	-0,4	1,66%

iii. Résultats

- **Réplication de la fonction de calcul de P&L**

Nous vérifions ici que le P&L estimé par notre algorithme, pour deux scénarios donnés, est satisfaisant :

Scenarios				Résultats		
Scenario	Choc Eurostoxx50	Choc Euribor3 mois	Choc EUR/USD	P&L déduit (€)	P&L exact (€)	Erreur relative
scenario 1	-10%	+2 bp	+5%	-250 000	-250 000	0%
scenario 2	+20%	-5 bp	-10%	+ 600 000	+ 600 000	0%

Les P&L « exact » (cf. Eq1) et « déduit » de notre algorithme de machine Learning démontrent une erreur relative nulle, **soit une réplication parfaite**.

Le résultat peut être facilement généralisé pour les portefeuilles linéaires, sous réserve d'un nombre de scénarios disponibles suffisant pour la construction de l'ensemble d'apprentissage.

- **Test de résistance inversé : sélection du scénario**

Nous observons ensuite les scénarios fournis par la machine, pour deux montants de pertes donnés

Scenarios		Résultats					
Scenario	Hypothèse de P&L (€)	Choc Eurostoxx50	Choc Euribor3 mois	Choc EUR/USD	Probab	P&L réel (€)	Erreur relative
scenario 1	-10 000	-0,60%	- 0,06 bp	0,98%	26,51%	-10 000	0%
scenario 2	-1 000 000	6,54%	10,51 bp	1,48%	1,06%	-1 000 000	0%

Les probabilités associées à chaque scénario correspondent aux chances de réalisation des montants de pertes cibles. Nous constatons par exemple que nous avons une probabilité de réalisation de **1,06% pour un montant de perte de 1M€**. Le scénario associé est le **scénario le plus plausible** selon notre algorithme.

D. Use case n°2 : portefeuille optionnel

i. Description du portefeuille et Hypothèses

L'idée ici est de tester la capacité de notre algorithme à traiter le cas d'un portefeuille comportant de la **convexité**.

Nous considérons un portefeuille composé d'un call européen sur un sous-jacent Eurostoxx50 avec les caractéristiques suivantes :

underlying	maturity	risk free rate	div. rate	strike	vol.	not.
EUROSTOXX50	01/01/2023	0.0	0.01	3000	0.1	100000

La fonction de P&L s'appuie sur la formule de Black-Scholes. Nous supposons tous les paramètres constants à l'exception du sous-jacent.

ii. Description des données d'apprentissage

L'ensemble d'apprentissage est constitué comme suit :

- Les variables explicatives sont les niveaux historiques de l'Eurostoxx50 ;
- La variable expliquée est le P&L associé (variation de valeur des prix d'options calculés par la formule de Black-Scholes).

iii. Résultats

▪ Réplication de la fonction de calcul de P&L

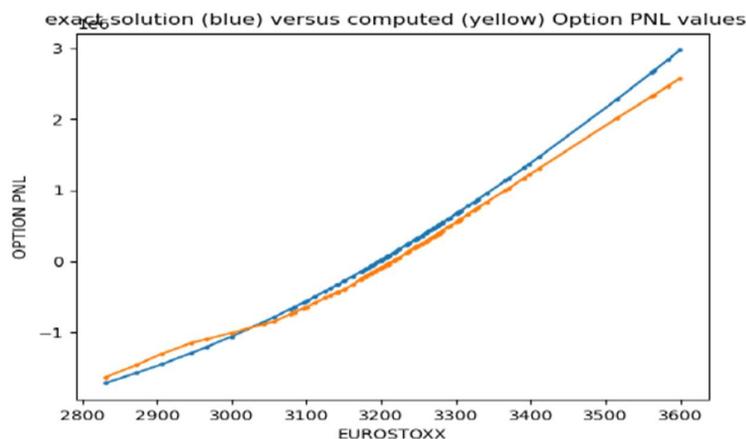
Les résultats obtenus pour deux scénarios de chocs sont les suivants :

Scenarios		Résultats		
scenario	choc Eurostoxx50	P&L déduit (EUR)	P&L exact (EUR)	Erreur relative
scenario1	-10%	-126 551	-116 775	8,4%
scenario2	20%	1 061 630	1 213 667	-12,5%

Les P&L « exact » et « déduit » calculés respectivement par la formule de Black-Scholes et notre algorithme, démontrent une erreur relative moyenne de l'ordre de 10% pour les deux scénarios. **Cet écart reste relativement faible** même pour des chocs particulièrement violents sur l'Eurostoxx50.

Le graphique ci-dessous montre comment évoluent les P&L « exact » (Black Scholes, en bleu) et « déduit » (CoDPy, en jaune) en fonction de l'Eurostoxx50. On observe des comportements très proches, cohérents avec la position considérée (achat d'un call).

Comparaison des P&L réels (Black Scholes) et déduits de la machine, en fonction de l'Eurostoxx 50



Source : MPG Partners, CoDPy

▪ **Test de résistance inversé : sélection du scénario**

Scenarios		Résultats			
scenario	P&L (EUR)	choc Eurostoxx50	Probab	P&L Black & Scholes (EUR)	Erreur relative
scenario1	-10 000	-0.77%	41.7%	-11 478	-12,9%
scenario2	-1 000 000	-33.2%	0.05%	-1 309 076	-23,6%

Le scenario de choc de l'Eurostoxx50 déterminé par CoDPy est de **-0,77% pour un montant de perte de -10K€ avec une probabilité de 41,7%**.

Pour une perte de 1M€, nous obtenons un choc de -33,2% sur l'Eurostoxx50, et une probabilité associée quasi nulle, ce qui est cohérent avec le fait qu'un tel choc n'a jamais été observé sur une période de 10 jours.

Conclusion

Jusqu'à présent, le reverse stress testing, impératif réglementaire, obligeait les établissements financiers à faire des hypothèses simplificatrices qui venaient altérer significativement sa qualité.

Désormais, grâce au Machine Learning, nous sommes capables de nous affranchir de ces contraintes pour réaliser un exercice exhaustif (périmètre, facteurs de risque) qui prend en compte la non-linéarité des positions.

Notre solution, appelée CoDPy, qui utilise les méthodes à noyaux (RKHS : *Reproducing Kernel Hilbert Space*) donne des résultats très satisfaisants, pour l'inversion des fonctions de P&L et la caractérisation de la plausibilité des scenarios obtenus.

En vue des prochains exercices règlementaires, il serait intéressant d'étendre l'étude sur des portefeuilles non-linéaires (optionnels) complexes et capitaliser sur le caractère **interprétable** et **reproductible** de la méthode.

Annexe

Ci-dessous, quelques directives et recommandations publiées par les organes de régulations qui mentionnent la nécessité de mettre en place des scénarios de Reverse Stress tests :

Source	Reference	Régulation
SYSC	20.2.1	"... A firm must reverse stress test its business plan; that is, it must carry out stress tests and scenario analyses that test its business plan to failure...."
SYSC	20.2.2	"...The reverse stress test must be conducted on a solo basis as well as on a consolidated basis in relation [to] the UK consolidation group or the non-EEA sub-group...."
CRR	Page 219 art.368 para.1(g)	"...The institution shall frequently conduct a rigorous program of stress testing, including reverse stress tests, which encompasses any internal model used for [the] purposes of this chapter and the results of these stress tests shall be reviewed by senior management and reflected in the policies and limits it sets...."
EBA GL 32	Section 3.4 para.65	"...Reverse stress testing is seen as one of the risk management tools usefully complementing the "usual" stress testing, which examines outcomes of predetermined scenarios. Reverse stress testing is a useful tool in risk management...."
EBA/GL/2018/04	4. Institutions' stress testing, para 15	"...Institutions should also include reverse stress testing and reverse stress testing scenarios in their stress testing programs....."

Références

- [1] P.G. LeFloch, J.M. Mercier, The Transport-based Mesh-free Method: A Short Review, *Wilmott*, Volume 2020, Issue 109, Pages: 52,57 , September 2020, <https://doi.org/10.1002/wilm.10875>
- [2] Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, Alexander Smola, *The Journal of Machine Learning Research* Volume 133/1/2012 pp 723–773
- [3] Philippe G. LeFloch, Jean-Marc Mercier, Mesh-free error integration in arbitrary dimensions: A numerical study of discrepancy functions, *Computer Methods in Applied Mechanics and Engineering*, Volume 369, 2020, 113245, ISSN 0045-7825, <https://doi.org/10.1016/j.cma.2020.113245>.
- [4] P.G. LeFloch, J.M. Mercier, S. Miryusupov, "CodPy: a Python library for machine learning, mathematical finance, and statistics", internal codpy user manual, to be submitted.