

# RAPPORT FINAL DU PROJET SCIENTIFIQUE COLLECTIF

Modèle Multi-Fractal pour la prévision en finance

Août 2022 - Avril 2023

Soufiane AJI, Mehdi BENFRAIHA, Ossama FARAJI  
Martin MORANGE, Ziad-Eddine OUMZIL, Khadija SLIM



## REMERCIEMENTS

Avant d'entamer cet ouvrage, nous voudrions exprimer ma gratitude envers l'École polytechnique qui, sous la vigilance du corps administratif et professoral, oeuvre ardemment à procurer aux étudiants une formation à la fois humaine et pratique.

Par la présente, nous souhaitons remercier l'ensemble des acteurs impliqués dans la réalisation de ce projet. Nous tenons tout d'abord à exprimer notre reconnaissance envers notre coordinateur, Monsieur Julien COMBE, pour sa vigilance et son souci constant quant à la bonne marche du projet. En effet, Monsieur COMBE n'a pas ménagé ses efforts en prodiguant des conseils avisés qui nous ont grandement aidé à réussir notre projet.

Nous tenons également à adresser nos remerciements les plus sincères à tous les membres du groupe, dont les motivations, le sérieux et l'implication sans faille ont été les vecteurs de la progression durant toute la durée du projet. Le soutien inconditionnel et la collaboration active ont été les piliers de notre réussite.

Par-dessus-tout, nous souhaitons exprimer notre gratitude envers Monsieur Laurent CALVET. Grâce à son expérience et son expertise, il a su nous apporter des conseils précieux qui ont eu un impact considérable dans notre avancement. Sa contribution a été déterminante et nous lui sommes extrêmement reconnaissants pour tout ce qu'il a pu faire pour nous aider à atteindre nos objectifs.

## TABLE DES MATIÈRES

---

<b>1 Introduction</b>	<b>4</b>
1.1 Motivations . . . . .	4
1.2 Organisation du travail au sein du groupe et difficultés rencontrées . . . . .	4
<b>2 Les options</b>	<b>6</b>
2.1 Explications sommaires : . . . . .	6
2.1.1 Définition : . . . . .	6
2.1.2 Exemple : . . . . .	6
2.1.3 Les déterminants de la valeur d'une option : . . . . .	7
2.2 Pourquoi s'intéresser à la volatilité : premium et volatilité du cours du titre de base . . . . .	7
<b>3 Stratégies d'investissement</b>	<b>9</b>
3.1 Les opérations de base sur titres : . . . . .	9
3.1.1 Stratégie 1 : Achat d'actions ou position acheteuse en actions : . . . . .	9
3.1.2 Stratégie 2 : Vente à découvert d'actions ou position vendeuse en actions : . . . . .	9
3.2 Les spéculations de base à risque limité : . . . . .	9
3.2.1 Stratégie 3 : Achat d'options d'achat : . . . . .	9
3.2.2 Stratégie 4 : Achat d'options de vente : . . . . .	10
3.3 Vente d'options d'achat : . . . . .	10
3.3.1 Stratégie 5 : Vente découverte d'option d'achat : . . . . .	10
3.4 Vente d'options de vente : . . . . .	10
3.4.1 Stratégie 6 : Vente découverte d'option de vente. . . . .	10
3.4.2 Call ouvert et Call ouvert Inversé : . . . . .	11
3.4.3 Put protecteur et Put protecteur inversé : . . . . .	12
<b>4 Présentation de différents modèles pour la volatilité</b>	<b>13</b>
4.1 Le modèle MSM binomial . . . . .	13
4.2 Le modèle GARCH( $p, q$ ) . . . . .	15
4.3 Le modèle RNN . . . . .	15
<b>5 Implémentation, Résultats et comparaisons</b>	<b>17</b>
5.1 Implémentation du modèle MSM-Binomial . . . . .	17
5.2 Comparaison des modèles . . . . .	19
5.2.1 Le choix des estimateurs utilisés pour la comparaison . . . . .	19
5.2.2 Limitations des statistiques utilisés . . . . .	20
5.2.3 Les données de test utilisées . . . . .	20
5.2.4 Comparaisons . . . . .	21
5.3 Notes sur les algorithmes et implémentations . . . . .	24
5.3.1 Basin-Hopping . . . . .	24

<b>6 Skew MSM : un modèle ajusté aux options</b>	<b>25</b>
6.1 Motivation . . . . .	25
6.2 Le modèle Skew MSM . . . . .	26
6.2.1 Une classe de modèle de diffusion conditionnellement affine . . . . .	26
6.2.2 Modèle ajusté au risque . . . . .	27
6.2.3 Le modèle SkewMSM . . . . .	28
6.2.4 SkewMSM ajusté au risque . . . . .	28
6.3 Description des paramètres du modèle . . . . .	29
6.4 Implémentation de Skew MSM . . . . .	30
6.4.1 Notes sur l'implémentation . . . . .	30
6.4.2 Estimation sur les données d'equity . . . . .	30
6.5 Benchmarks . . . . .	31
6.5.1 Spécifications . . . . .	31
6.5.2 Résultats empiriques . . . . .	32
6.5.3 Résultats sur les données d'options . . . . .	32
6.5.4 Volatility smiles . . . . .	34
<b>7 Conclusion</b>	<b>36</b>
<b>8 Annexes</b>	<b>37</b>
8.1 Données et codes informatiques : . . . . .	37
8.2 Pistes d'approfondissement . . . . .	37
8.2.1 Optimisations informatiques pour Skew MSM . . . . .	37
8.2.2 GARCH . . . . .	37
8.2.3 RNN . . . . .	38



# 1 INTRODUCTION

## 1.1 MOTIVATIONS

La volatilité est une mesure de l'ampleur des variations des prix d'un actif financier sur une période donnée. Elle revêt une importance capitale pour les investisseurs, car elle peut influencer les décisions de trading et les stratégies d'investissement. En effet, une volatilité élevée peut offrir des opportunités de profit plus importantes, mais elle peut également augmenter les risques de pertes importantes. Ainsi, la capacité à prédire la volatilité future d'un actif est une compétence essentielle pour les investisseurs, leur permettant de prendre des décisions éclairées et de minimiser les risques. Cela implique l'utilisation de techniques d'analyse de marché avancées, telles que les modèles de volatilité, pour anticiper les fluctuations futures des prix.

C'est dans ce cadre que s'inscrit le sujet de notre PSC, où nous nous sommes focalisés sur la compréhension, l'implémentation et la comparaison de différents modèles de prédiction de la volatilité. Nous avons approfondi l'étude des modèles multifractaux, notamment le modèle MSM binomial, SKEW MSM (pour l'étude des options) et MSMI, que nous avons comparés à d'autres modèles tels que ARCH, GARCH et RNN.

Suite au travail de nos camarades qui ont traité ce sujet l'an dernier et sous les conseils de Monsieur Calvet, nous avons repris les prévisions établies l'an dernier pour nous assurer de l'exactitude de nos programmes avant de nous attaquer aux options, axe sur lequel nous avons choisi de travailler et qui constituera notre principal ajout sur le sujet.

## 1.2 ORGANISATION DU TRAVAIL AU SEIN DU GROUPE ET DIFFICULTÉS RENCONTRÉES

Un des grands enjeux de ce projet a été de cerner correctement les différents aspects de celui-ci et leurs enjeux : que ce soit la compréhension de la théorie qui sous-tend les modèles étudiés, l'implémentation pratique de ces modèles ou encore les différentes analyses économiques que nous avons pu produire, chacune de ces composantes mobilise des compétences très différentes. Nous avons donc décidé de diviser le travail et de créer trois "sous-groupes", chacun s'occupant d'un aspect du projet (mathématique, informatique et économique). Des réunions hebdomadaires nous ont permis de garder une dynamique commune tout en contrôlant les progrès effectués dans la semaine.

Ces réunions hebdomadaires, quoique quelque peu chaotiques au début du projet, ont vité été cadrées, avec un chef de groupe annonçant l'ordre du jour, laissant alors la parole à chaque sous-groupe qui pouvait partager ses résultats avec le reste du groupe. Le chef de groupe fixait alors des objectifs à atteindre pour la semaine suivante, ce qui a par ailleurs permis au projet d'avoir un rythme d'avancement assez soutenu.

Nous avons aussi eu l'occasion de communiquer avec M. Calvet, notre tuteur, à l'aide de réunions Zoom, ou par échange de mails, ce qui nous a permis de nous débloquer lorsque des difficultés théoriques ou pratiques ralentissaient l'avancée du projet.

Cependant, l'organisation du travail a aussi provoqué certaines difficultés. La grande marge de manœuvre qui nous a été laissée a pu être déroutante, nous forçant à nous organiser de A à Z, à fixer nos propres attentes et objectifs sur le projet. L'emploi du temps dense à certaines périodes de l'année a ainsi pu contribuer à ralentir la vitesse de progression du projet, même si notre investissement est resté à peu près constant tout au cours de l'année.

## 2 LES OPTIONS

### 2.1 EXPLICATIONS SOMMAIRES :

#### 2.1.1 • DÉFINITION :

Un contrat d'option sur actions confère à son porteur un droit : celui d'acheter ou de vendre un nombre déterminé de ces actions à un prix fixé et désigné dans le contrat comme prix d'exercice. Ce droit peut-être exercer que durant la vie de l'option, c'est-à-dire jusqu'à la date d'exercice (ou encore date d'échéance).

Si l'option accorde à celui qui la détient une possibilité d'achat, il s'agit d'un « *call* ».

Si l'option accorde à celui qui la détient une possibilité de vente, il s'agit d'un « *put* ».

Un contrat d'option se caractérise par les éléments suivants :

- Le premium : Il est payé par l'acheteur au vendeur car le premier a le droit et non l'obligation de prendre une décision à laquelle le vendeur devra se soumettre ;
- La nature de l'option : option d'achat ou option de vente ;
- Le titre de base ou titre support du contrat d'option : devises, matières premières, actions... .

#### 2.1.2 • EXEMPLE :

Sur ce call EUF/CHF à prix d'exercice 1.0800 : En prenant en compte le **premium** l'acheteur est gagnant et a tout intérêt à exercer l'option d'achat (call) si la rapport EUR/CHF dépasse les 1.0900. Si le cours ne dépasse pas le prix d'exercice, le mieux à faire pour ce dernier est de ne pas exercer l'option d'achat. Ces 2 options sont représentées par les zones de pertes et zones de gains.

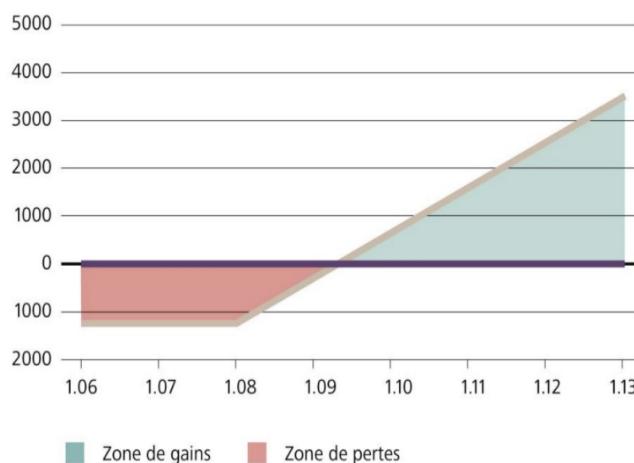


FIGURE 1 – Option long call EUR/CHF, prix d'exercice 1.0800

Le raisonnement est analogue lorsqu'il s'agit d'options de vente (put) : il suffit d'inverser zone de gains et zones de pertes. Plus le cours baissera plus l'option sera rentable à celui qui l'a acheté.

### 2.1.3 • LES DÉTERMINANTS DE LA VALEUR D'UNE OPTION :

- Le premium** d'une option dépend principalement de 6 paramètres :
- du cours du titre de base : plus précisément de son **niveau** et de sa **volatilité**.
  - des éléments caractéristiques du contrat : **prix d'exercice** et **date d'exercice**.
  - des **taux d'intérêt**.
  - les **dividendes** susceptibles d'être distribuées par l'action pendant la durée de vie de l'option.
- Dans l'optique de ce PSC, la partie qui nous intéressera et que nous devrons évaluer est **la volatilité**.

## 2.2 POURQUOI S'INTÉRESSER À LA VOLATILITÉ : PREMIUM ET VOLATILITÉ DU COURS DU TITRE DE BASE

---

La théorie financière et l'observation empirique de la distribution des taux de rentabilité d'une action indiquent que si l'on veut faire des gains élevés en achetant des actions, il faut accepter de faire éventuellement face à des pertes élevées.

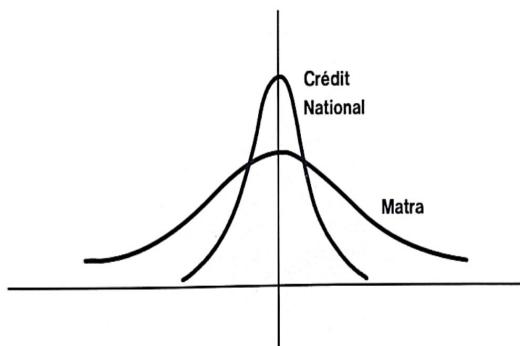


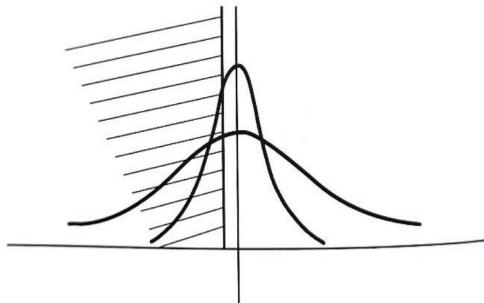
FIGURE 2 – Distribution des taux de rentabilité hebdomadaire de Matra et Crédit National

Sur la figure ci-dessus, si l'investisseur souhaite une rentabilité élevée il achètera Matra. Ce faisant, il pourra se retrouver avec des pertes plus élevées que s'il avait acheté le titre Crédit National.

On dit que **Matra est plus volatile que Crédit National**

Alors que les gains et pertes sont symétriques dans le cas des actions, ils sont asymétriques dans le cas des options. Si le produit gagne en valeur, l'acheteur des options va gagner beaucoup. Au contraire, si une baisse survient, il perd le même montant quel que soit l'amplitude de cette baisse, faible ou forte.

Dans ces deux cas, il abandonne seulement son droit et perd le premium :



L'option d'achat sur le titre le plus volatile (Matra) est donc la plus intéressante. À probabilité de résultat nul (sans prise en compte de la valeur du premium) égale, l'option Matra offre pour les mêmes probabilités d'occurrence des profits nettement supérieurs à ceux de l'option sur Crédit National : L'opérateur paiera plus chère la distribution Matra que la distribution Crédit National, quelle que soit son aversion au risque.

## 3

# STRATÉGIES D'INVESTISSEMENT

Afin de mieux apprécier l'apport des options, nous commencerons par présenter les opérations de bases : achat et ventes d'actions.

### 3.1 LES OPÉRATIONS DE BASE SUR TITRES :

#### 3.1.1 • STRATÉGIE 1 : ACHAT D'ACTIONS OU POSITION ACHETEUSE EN ACTIONS :

Cette opération revient à acheter un nombre  $X$  d'actions en espérant pouvoir les revendre plus chères qu'il ne les a achetées dans certains. Cela revient à anticiper une hausse du cours du titre. Par euro investi, il apparaît un profit de 1 euros supplémentaire chaque fois que le cours du titre augmente de 1 euro. Par symétrie il apparaît une perte de 1 euro chaque fois que le cours du titre diminue de 1 euro.

#### 3.1.2 • STRATÉGIE 2 : VENTE À DÉCOUVERT D'ACTIONS OU POSITION VENDEUSE EN ACTIONS :

Cette opération revient à emprunter à un tiers un nombre  $X$  d'actions revendues immédiatement. Il devra les racheter dans un certain temps pour les rendre à celui qui les a prêtées. Il espère donc pouvoir les racheter moins chères qu'il ne les a initialement vendues. Ce faisant, l'investisseur en vendant des actions à découvert anticipe une baisse du cours de l'action.

Ce dernier perd donc un euro par euro de hausse du titre et gagne un euros par euro de baisse du titre.

### 3.2 LES SPÉCULATIONS DE BASE À RISQUE LIMITÉ :

L'un des attraits des options négociables est de fournir des opportunités de spéculation qui, en cas de retournement, n'exposent leurs initiateurs qu'à des risques de pertes limitée.

#### 3.2.1 • STRATÉGIE 3 : ACHAT D'OPTIONS D'ACHAT :

L'investisseur achète une option d'achat LVMH septembre/1000 pour 100 euros (prix du premium). Le cours de l'action s'établissant également à la même date pour 1000 euros.

Ce faisant l'investisseur escompte une hausse du titre. Le résultat à l'échéance, pour l'acheteur sera égal au maximum de zéro et de la différence entre le cours de l'action à l'échéance et le prix d'exercice, dont il faut retrancher le montant du premium payé.

- Si le cours de l'action dépasse 1000 euros, les prédictions de l'investisseurs se sont avérées correctes et ce dernier aura tout intérêt à exercer son option et l'investisseur verra sa richesse augmenter de 1 euro pour chaque augmentation de 1 euro de la valeur du titre. Si par exemple le cours atteint les 1500 euros, ses gains seront de  $1500 - 1000 = 500$  duquel il faut retrancher le prix du premium soit des gains nets de  $500 - 100 = 400$  euros.



- Si le cours de l'action ne dépasse pas les 1000 euros et chute, l'action ne sera pas exercée et la perte de l'investisseur se limite donc au prix du premium, soit 100 euros.

### 3.2.2 • STRATÉGIE 4 : ACHAT D'OPTIONS DE VENTE :

Le raisonnement est analogue : L'investisseur escompte une baisse du titre. A l'échéance le résultat de l'opération sera égal au maximum de zéro et la différence entre le prix d'exercice et le cours de l'action, dont il faut retrancher le montant du premium payé. L'investisseur voit sa richesse augmenter de 1 F pour chaque diminution de 1 F de la valeur du titre de base.

## 3.3 VENTE D'OPTIONS D'ACHAT :

Il existe 2 types de vente d'options d'achat : la vente d'option d'achat découverte et la vente d'option d'achat couverte. Nous aborderons la vente couverte dans la partie suivante.

### 3.3.1 • STRATÉGIE 5 : VENTE DÉCOUVERTE D'OPTION D'ACHAT :

L'investisseur vend une option d'achat LVMH septembre/1000 pour 100 euros, alors qu'il n'a pas les titres en portefeuille. Il s'agit d'une vente effectuée dans le but de réaliser un bénéfice dans un marché neutre, légèrement à la baisse ou légèrement à la hausse.

Si les anticipations de faibles variations du cours du titre support se sont effectivement matérialisées et que le cours de LVMH à l'échéance du contrat d'option est proche du prix d'exercice, l'investisseur aura encaissé le premium, soit 100 euros sans qu'il ai à payer quoi que ce soit en contrepartie. La stratégie semble à première vue d'un rapport infini, puisqu'elle engendre des gains sans contrepartie apparente. En fait, il y a une contrepartie : toute vente d'option constituant une dette potentielle, les autorités du marché exigent un dépôt de couverture qui garantit la bonne fin des opérations quelle que soit l'évolution subséquente du cours de l'action.

## 3.4 VENTE D'OPTIONS DE VENTE :

Il existe 2 types de vente d'options de vente : la vente d'option de vente découverte et la vente d'option de vente couverte. Nous aborderons la vente couverte dans la partie suivante.

### 3.4.1 • STRATÉGIE 6 : VENTE DÉCOUVERTE D'OPTION DE VENTE.

Comme pour les ventes de calls, les ventes découvertes d'options de vente présentent de sérieux risque. Cette position correspond à une anticipation de faible hausse ou du maintien du cours de LVMH à son niveau actuel pendant la durée de vie de l'option. Il n'y a donc AUCUN bénéfice à la hausse et une perte CROISSANTE à la baisse.

Si les anticipations de l'investisseur se sont matérialisées, il aura encaissé un premium avec pour seule contrepartie l'immobilisation du dépôt de couverture pendant la durée de l'opération. Les stratégies simples d'achat et de vente « secs » d'une option parient sur l'évolution à la hausse ou à la baisse du cours de l'actif sous-jacent. Cependant, elles peuvent être associées à l'achat ou à la vente de l'actif sous-jacent.

Toutes ces stratégies sont uniquement possibles de part le fait que les gains et pertes sont symétriques

dans le cas des actions mais asymétriques dans le cas des options.

Le modèle de graphique suivant sera utilisé pour illustrer ces différentes stratégies.

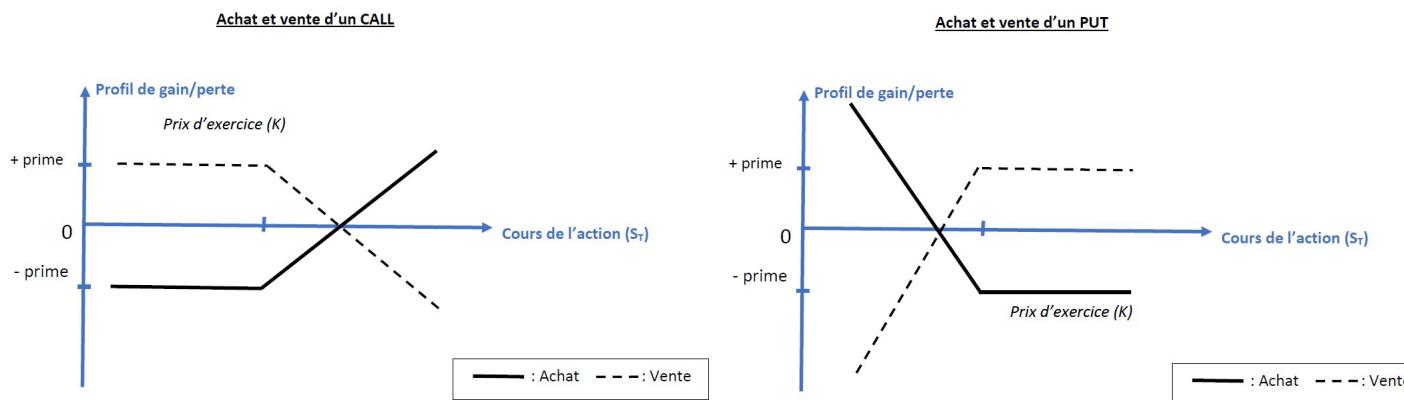


FIGURE 3 – Achat et vente d'un CALL/PUT

### 3.4.2 • CALL OUVERT ET CALL OUVERT INVERSÉ :

On parle alors de call ouvert lorsque la stratégie implique une position acheteuse en actions et vendeuse en call. Le risque pour le vendeur du call est que le marché remonte. Si tel est le cas, il sera protégé par sa position en actions.

Dans le cas d'une vente simple de call, le vendeur subit autant de perte que le marché ne remonte (à partir du prix d'exercice et en soustrayant le prix premium). Cependant, avec une position acheteuse en actions prise au même moment que la position vendeuse en call, les gains engendrés par celle-ci lorsque le cours de l'action remonte sont égaux aux pertes engendrées par la vente du call.

De ce fait, cette stratégie fixe une limite de gain et protège le vendeur du call en cas d'augmentation trop importante du marché.

Cependant, si le marché redescend, cette couverture perd tout son intérêt car même si la vente du call n'engendre pas de perte, la vente d'action elle en engendre autant que le marché ne baisse.

Dans ce qui suit :

Le trait gris représente le/la gain/perte engendré par l'option achetée/vendue.

Le trait noir représente somme gain/perte option et action.

Le trait en pointillé, quant à lui, représente le/la gain/perte engendré par l'action achetée/vendue.

$K$  étant le prix d'exercice.

$S_T$  : Cours de l'action.

La stratégie inverse implique une position acheteuse en calls et vendeuse en actions. Cette stratégie inverse symétrique s'appelle la stratégie du call couvert inversée (Cf les deux graphes ci-dessous).

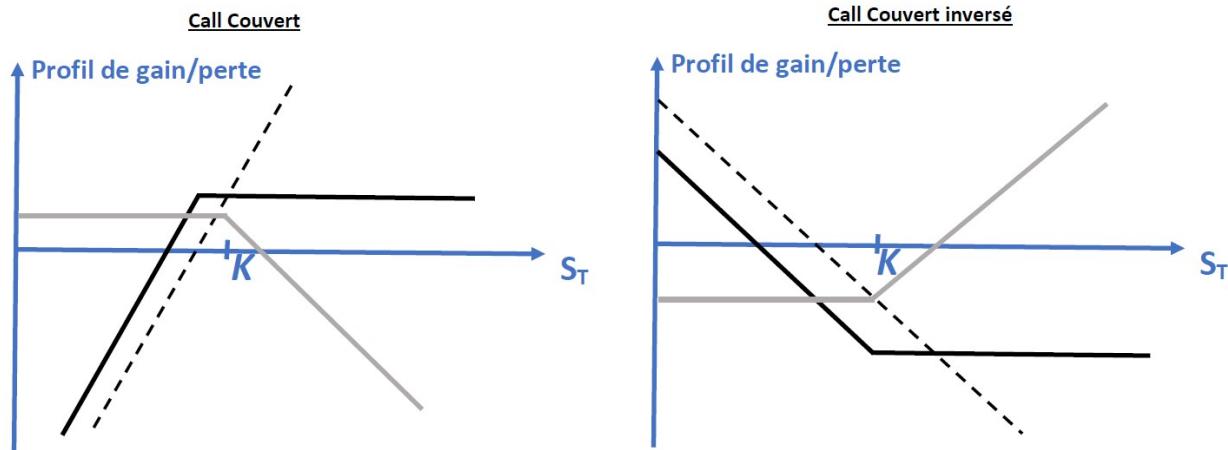


FIGURE 4 – Différence entre Call Couvert et Call Couvert inversé

### 3.4.3 • PUT PROTECTEUR ET PUT PROTECTEUR INVERSÉ :

On parle de put protecteur lorsque la stratégie implique une position acheteuse en put et acheteuse en actions. Le risque pour le vendeur en action est que le marché baisse. Si tel est le cas, il sera protégé par sa position en put.

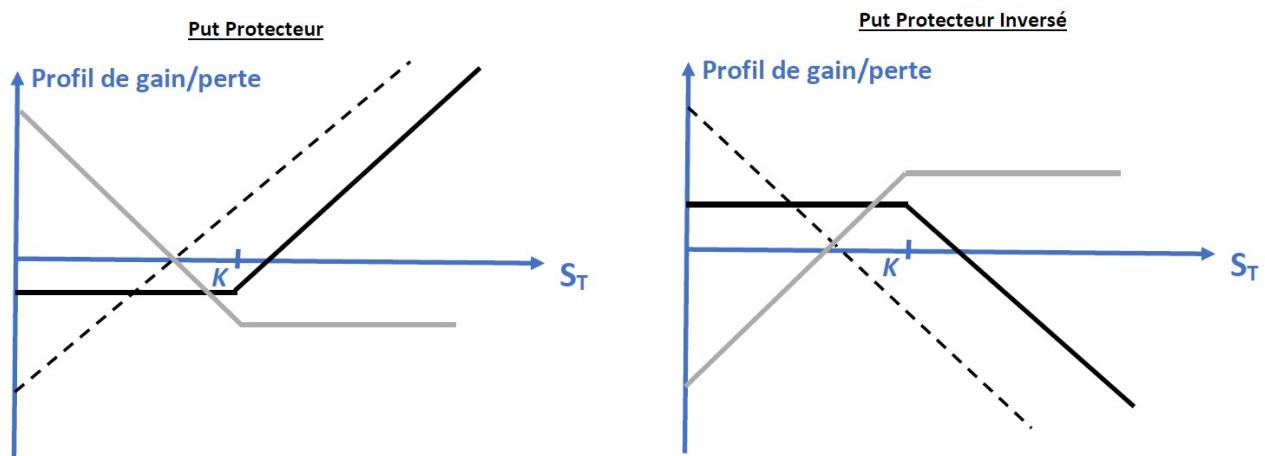


FIGURE 5 – Différence entre Put Protecteur et Put Protecteur inversé

La stratégie inverse implique une position vendeuse en put et vendeuse en actions. Cette stratégie inverse symétrique s'appelle la stratégie du call protecteur inversé.

## 4 PRÉSENTATION DE DIFFÉRENTS MODÈLES POUR LA VOLATILITÉ

### 4.1 LE MODÈLE MSM BINOMIAL

Le modèle MSM (Markov Switching Multifractal) est un modèle utilisé pour mesurer la volatilité sur les marchés financiers. La volatilité représente l'amplitude des variations de prix pour un actif. Dans ce modèle, on suppose que les variations de prix suivent une distribution normale centrée, dont la variance dépend uniquement d'une variable appelée  $M_t$ , qui est une chaîne de Markov. Le rendement  $r_t$  est calculé à partir de cette distribution et du prix de l'actif à un moment donné. Dans le cas du modèle MSM, on suppose que ce rendement suit une loi gaussienne centrée dont la variance ne dépend que de  $M_t$ , où  $(M_t)$  est une chaîne de Markov :

$$r_t = \sigma(M_t) \varepsilon_t$$

avec  $r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$ ,  $P_t$  dénotant le prix de l'actif à l'instant  $t$ .

L'état de Markov  $M_t$  est un vecteur à  $\bar{k}$  dimensions, où  $\bar{k}$  est une variable exogène, donné par :

$$M_{k,t} = \begin{cases} \text{tiré selon une distribution } M \text{ avec une probabilité } \gamma_k \\ M_{k,t-1} \text{ avec une probabilité } 1 - \gamma_k \end{cases}$$

avec  $\gamma_k = 1 - (1 - \gamma_1)^{b^{k-1}}$ ,  $\gamma_1 \in (0; 1)$ ,  $b \geq 1$

Dans le cadre du modèle MSM binomial, on prend pour  $M$  une distribution binomiale, valant  $m_0$  ou  $2 - m_0$  avec une probabilité de  $\frac{1}{2}$  pour les deux valeurs.

Enfin,  $\sigma(M_k)$  est donné par :

$$\sigma(M_k) = \bar{\sigma} \sqrt{\prod_{k=1}^{\bar{k}} M_{k,t}}$$

où  $\bar{\sigma}$  est un réel quelconque.

Ainsi, La volatilité  $\sigma$  est donnée par une formule qui dépend de  $M_k$  et d'un réel  $\bar{\sigma}$ . En somme, il n'y a que quatre paramètres nécessaires pour décrire ce modèle :  $m_0$ ,  $\bar{\sigma}$ ,  $\gamma_1$  et  $b$  :

$$\Psi = (m_0, \bar{\sigma}, \gamma_1, b)$$

Voici une explication détaillée des quatre paramètres du modèle MSM :

- $m_0$  : Représente la valeur de base de la chaîne de Markov  $M_t$ . Dans le modèle MSM binomial,  $M$  suit une distribution binomiale valant soit  $m_0$ , soit  $2 - m_0$ , avec une probabilité de  $\frac{1}{2}$  pour chaque valeur.
- $\bar{\sigma}$  : Une mesure de la variabilité de la distribution normale centrée utilisée pour calculer les rendements. En d'autres termes,  $\sigma$  mesure la volatilité du marché financier que le modèle tente de représenter. Plus  $\sigma$  est élevé, plus la volatilité du marché est grande.

- $\gamma_1$  : Le troisième paramètre est  $\gamma_1$ , qui représente la probabilité de transition d'un état à un autre de la chaîne de Markov. Dans le modèle MSM, la probabilité de transition d'un état à l'autre varie en fonction de la valeur de  $k$ , qui est un entier compris entre 1 et  $\bar{k}$  (une variable exogène). En effet, la probabilité de transition de l'état 1 à l'état  $k$  est définie par la formule  $\gamma_k = 1 - (1 - \gamma_1)^{b^{k-1}}$ , où  $b$  est un entier supérieur ou égal à 1.
- $b$  : Le dernier paramètre est  $b$ , qui est un autre paramètre lié à la chaîne de Markov. Comme expliqué précédemment,  $b$  est un entier supérieur ou égal à 1 qui affecte la probabilité de transition d'un état à un autre de la chaîne de Markov. Plus la valeur de  $b$  est grande, plus la probabilité de transition d'un état à un autre diminue rapidement.

En somme, ces quatre paramètres  $(m_0, \bar{\sigma}, \gamma_1, b)$  permettent de décrire les caractéristiques clés de notre modèle et de modéliser la volatilité sur les marchés financiers. Pour cela, il faut déterminer les quatre paramètres importants :  $m_0, \bar{\sigma}, \gamma_1, b$ .

La méthode courante pour trouver ces paramètres est de maximiser la log-vraisemblance à l'aide d'une méthode de Newton, telle que BFGS, BHHH, Broyden, DFP ou SR1. Cependant, cette méthode nécessite de calculer l'inverse de la matrice hessienne de la log-vraisemblance, ce qui peut être coûteux en temps de calcul.

La log-vraisemblance est une fonction qui mesure la probabilité de trouver les paramètres optimaux pour le modèle. Elle est calculée en comparant les données réelles des marchés financiers aux résultats théoriques prédits par le modèle. En maximisant la log-vraisemblance, on obtient les paramètres optimaux qui rendent le modèle le plus cohérent avec les données réelles.

La méthode de Newton est une méthode d'optimisation qui permet de trouver des paramètres en résolvant des équations non-linéaires. Elle consiste à faire une première estimation des paramètres, puis à calculer la dérivée de la log-vraisemblance par rapport à chaque paramètre. À partir de ces dérivées, on construit la matrice hessienne, qui est une matrice carrée de secondes dérivées de la log-vraisemblance. Enfin, on utilise cette matrice pour trouver les valeurs optimales des paramètres.

Cependant, le calcul de l'inverse de la matrice hessienne peut être très coûteux en temps de calcul, car il implique d'effectuer des opérations sur une matrice de grande taille. C'est pourquoi, malgré son efficacité, la méthode de Newton peut être difficile à utiliser pour des modèles complexes tels que le modèle MSM.

Nous considérons tout d'abord la log-vraisemblance  $l$ , fonction de  $\mathbb{R}$  dans  $\mathbb{R}$ . Cherchons les zéros de  $l'$  :

Soit  $x$  tel que  $l'(x) = 0$ , un développement de Taylor au voisinage de  $x$  nous permet d'écrire :

$$l'(x) = l(y) + l''(y)(x - y) + o(\|x - y\|)$$

Et ainsi :

$$x = y - (l''(y))^{-1} l(y)$$

On constate donc qu'il suffit de résoudre avec  $u_0$  donné l'équation suivante :

$$u_{n+1} = u_n - (l''(u_n))^{-1} l(u_n)$$

Les méthodes quasi-Newton citées en haut donnent des approximations de la matrice hessienne, le requis pour cette résolution.

## 4.2 LE MODÈLE GARCH( $p, q$ )

Les modèles GARCH, qui sont également utilisés pour modéliser la volatilité des marchés financiers, sont efficaces pour fournir des prévisions à court terme sur la dynamique de la volatilité. Ils sont également relativement simples à utiliser et ne nécessitent que quelques paramètres pour les ajuster. Cependant, ils ont tendance à sous-estimer la volatilité sur de longues périodes, ce qui peut entraîner des erreurs de prévision importantes.

En comparaison, le modèle MSM est plus complexe et nécessite plus de paramètres pour l'ajuster, mais il peut mieux s'adapter à des périodes de volatilité élevée ou basse. De plus, il peut fournir des informations sur la distribution de la volatilité, ce qui peut être utile pour la gestion des risques. Cependant, sa complexité peut également rendre la méthode de Newton coûteuse en termes de temps de calcul.

Dans le cadre du modèle GARCH( $p, q$ ), on se donne  $(\varepsilon_t)$  une suite i.i.d. de variables gaussiennes centrées réduites,  $\mu, \alpha, \beta > 0$ . La volatilité  $\sigma_t^2$  est alors donnée par :

$$\sigma_t^2 = \mu + \sum_{i=0}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=0}^q \beta_i \sigma_{t-i}^2$$

Le log-return  $r_t$  est alors donné par :

$$r_t = \sigma_t \varepsilon_t$$

Pour l'estimation des paramètres, on peut utiliser la librairie python ARCH.

## 4.3 LE MODÈLE RNN

Le modèle RNN est un modèle de réseau de neurones qui se base sur l'apprentissage profond.

Le modèle se définit comme suit :

On a pour donnée une famille de vecteurs de dimension  $d_x$  qu'on note  $x_1, x_2, \dots, x_n$ .

A partir de cette famille, on génère une deuxième famille de vecteurs de dimension  $d_y$  qu'on notera  $y_1, y_2, \dots, y_n$  :

$$y_t = f_y(W_{hy}h_t + b_y)$$

Avec  $W_{hy}, b_y$  des matrices qui ont été entraînés avant d'être utilisés pour la prédiction.

La famille de matrice  $h_1, h_2, \dots, h_n$  est une famille d'états cachés de dimension  $d_h$  que le modèle génère et utilise.

Nous arrivons ainsi à l'étape finale qui est le calcul du rendement  $r_t$  à partir des rendements précédents.

On a donc comme donnée un vecteur  $r$  de dimension  $f$  contenant les rendements antérieurs  $r_{t-f}, r_{t-f+1}, \dots, r_{t-1}$ .



Le rendement se déduit donc ainsi :

$$r_t = W_1 y_f + b_1$$

Avec  $W_1$  et  $b_1$  également des données entraînées.

## 5 IMPLÉMENTATION, RÉSULTATS ET COMPARAISONS

### 5.1 IMPLÉMENTATION DU MODÈLE MSM-BINOMIAL

La réalisation du modèle MSM-binomial a été codée en Python, et est constituée d'une partie de simulation des paramètres  $\Psi = (m_0, \bar{\sigma}, b, \gamma_k)$  et une partie d'estimation des paramètres en utilisant ces mêmes données.

Nous avons simulé MSM-binomial sur les données du rendement de l'action de Microsoft jusqu'au 10/21/2022, en choisissant  $k = 4$  et les hyper-paramètres de l'algorithme d'optimisation (Basin-Hopping), niter = 10, temperature = 1.0 et stepsize = 1.0, et nous avons obtenu les estimations des paramètres. Nous avons simulé les données avec les paramètres estimés en comparant avec les données originales :

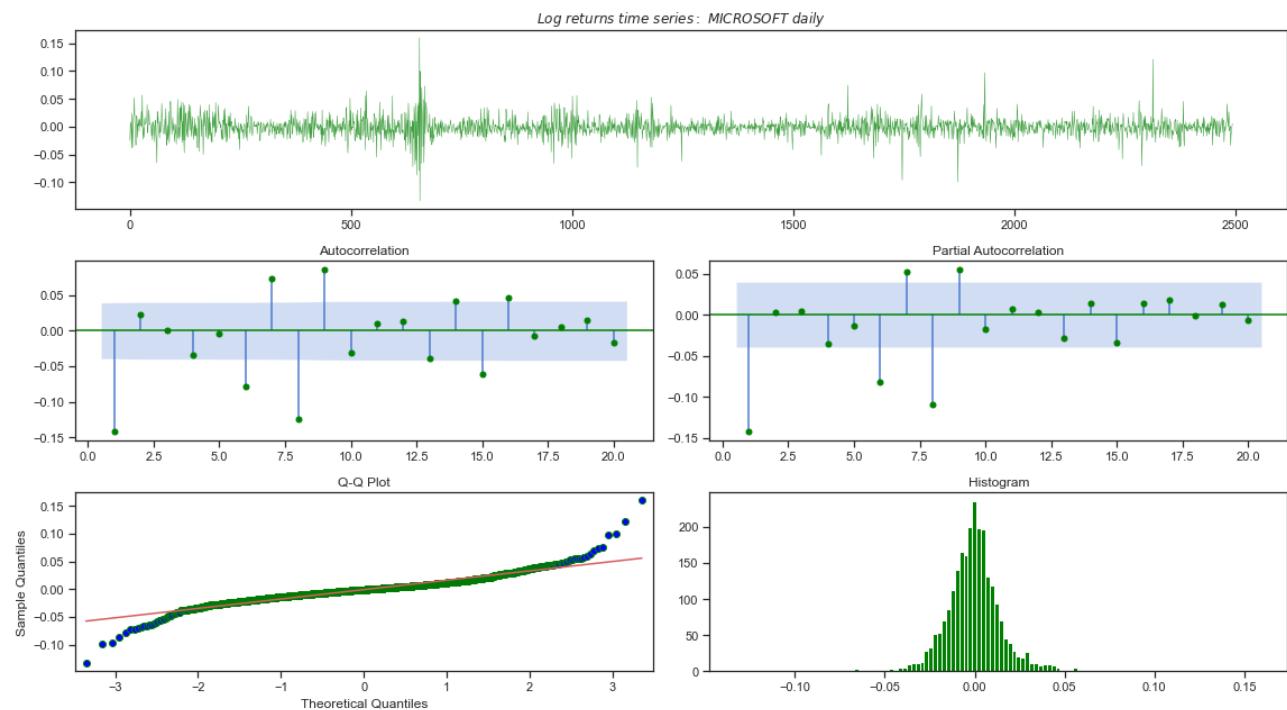


FIGURE 6 – Les rendements de l'action de MICROSOFT. Sur la première figure, on a tracé les rendements logarithmiques pour chaque jour. Dans la deuxième figure, on a tracé le quantile-quantile plot : la droite rouge représente les quantiles de la distribution gaussienne qui approxime le mieux notre distribution de rendements logarithmiques, ce qui permet de mesurer l'écart entre notre distribution et celle d'une loi normale. Dans la dernière figure est dessiné l'histogramme des rendements logarithmiques.

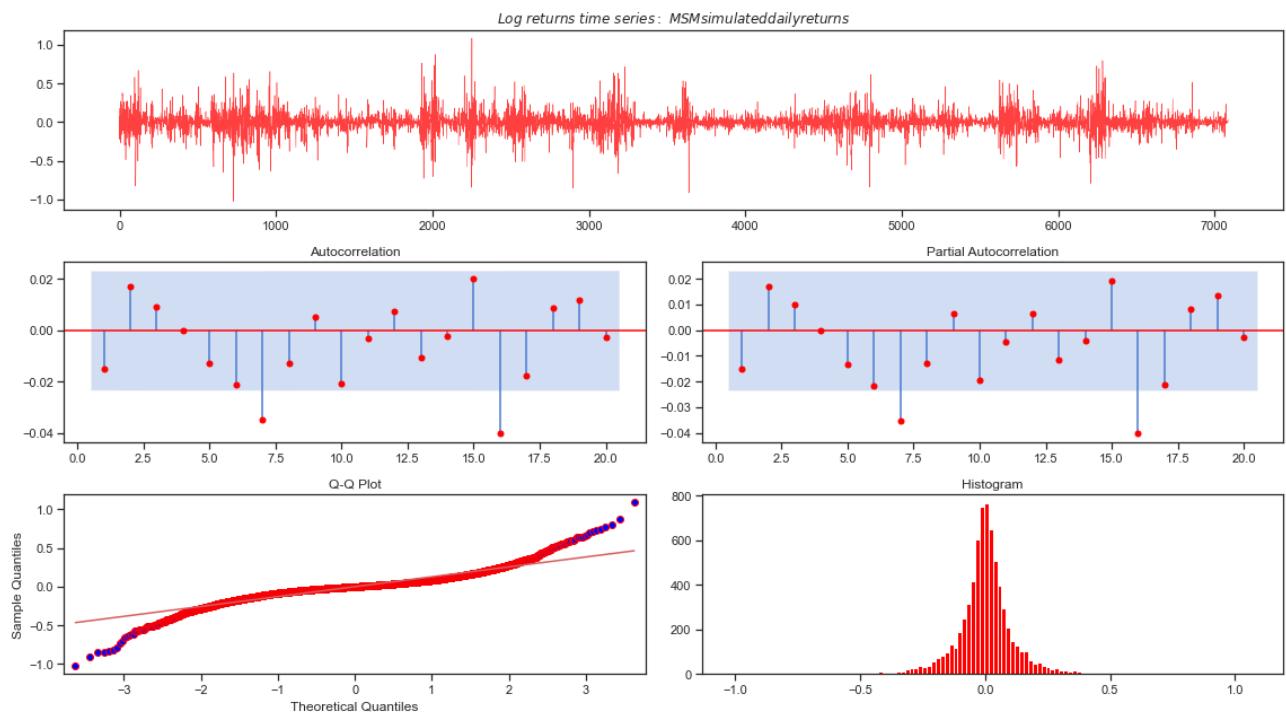
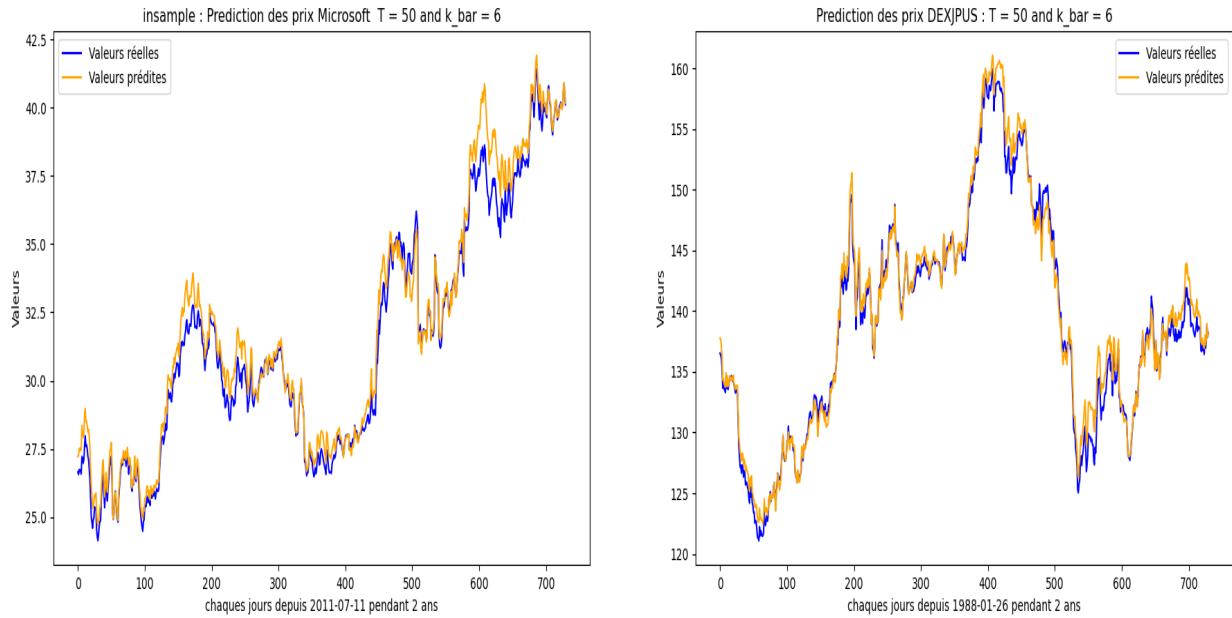


FIGURE 7 – Les données simulées. On peut observer que la distribution des rendements logarithmiques simulés est très proche de la distribution approchée réelle des rendements logarithmiques de MICROSOFT (cf Fig.1).



## 5.2 COMPARAISON DES MODÈLES

### 5.2.1 • LE CHOIX DES ESTIMATEURS UTILISÉS POUR LA COMPARAISON

Pour évaluer la qualité des simulations de nos modèles, nous avons besoin de mesurer l'écart entre les valeurs simulées et les valeurs réelles du marché. Pour cela, nous avons choisi trois statistiques d'erreur différentes pour évaluer la performance de nos modèles. Ces statistiques ont chacune leurs avantages et leurs limites. Nous avons donc décidé d'utiliser plusieurs statistiques pour avoir une évaluation plus complète. Dans la suite de notre travail, nous noterons les valeurs réelles du marché par  $Y$ , qui peut être les log-rendements ou les prix, et les valeurs simulées par  $X$ . Pour les modèles MSM, nous avons estimé les paramètres en utilisant des fonctions implémentées. Pour le modèle GARCH, nous avons utilisé la bibliothèque ARCH-PY pour estimer les paramètres. Pour le modèle RNN, nous avons utilisé la bibliothèque Keras du module TensorFlow, en spécifiant les paramètres "*optimizer = SGD*" (pour la descente de gradient stochastique) et "*loss = mean\_squared\_error*". Pour calculer les statistiques d'erreur, nous avons utilisé le module "sklearn.metrics" qui offre des implémentations de qualité de ces estimateurs.



### 5.2.1.1 Coefficient de détermination : $R^2$

$R^2$  est le coefficient de détermination, qui est une mesure pour qualifier la qualité d'une régression. Nous utiliserons  $R^2$  comme un critère de comparaison. Étant donné un ensemble de prédictions  $\{\hat{y}_1, \dots, \hat{y}_n\}$  et les valeurs réelles correspondantes  $\{y_1, \dots, y_n\}$ , nous pouvons calculer la valeur de  $R^2$  en utilisant la formule suivante :

$$R^2 = 1 - \frac{\sum_{i=1}^{i=n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{i=n} (y_i - \bar{y})^2}$$

### 5.2.1.2 Le risque quadratique moyen : MSE

Le risque quadratique moyen donne des informations sur la variance de notre estimateur. Dans nos calculs, nous avons défini

$$MSE = \sum_{i=1}^n \frac{(y_i - x_i)^2}{n}$$

Même si l'interprétation de MSE est plus délicate, cette statistique pénalise plus fortement les grandes erreurs. Nous avons donc jugé pertinent de la considérer.

### 5.2.1.3 L'erreur absolue moyenne : MAE

Si l'on souhaite comparer l'échantillon de prédiction X aux valeurs réelles Y, on définit

$$MAE = \sum_{i=1}^n \frac{|y_i - x_i|}{n}$$

L'intérêt réside en ce que le MAE est 'scalé' par rapport aux valeurs des données. Il s'agit d'une comparaison absolue. Il est donc plus évident de comprendre la valeur de sortie.

## 5.2.2 • LIMITATIONS DES STATISTIQUES UTILISÉS

Les deux métriques MAE et MSE calculent des distances, mais ne prennent pas en compte la direction. À tout moment, nous aurions pu utiliser un deuxième ensemble d'estimateurs pour pénaliser l'overfitting (resp. underfitting). Toutefois, l'étude des résultats ne nous a pas incitée à ajouter de tels critères.

## 5.2.3 • LES DONNÉES DE TEST UTILISÉES

Au début du projet, nous avons utilisé les données fournies par Nasdaq Data Link pour sélectionner et adapter les modèles avec lesquels nous allions travailler. Ces données étaient fiables et comprenaient des informations sur tous les indices financiers disponibles sur le site. Cependant, nous avons rencontré certaines limitations avec ces données, notamment le fait qu'un utilisateur normal n'a pas accès aux données historiques sur de longues périodes. De plus, nous ne pouvions pas modifier la fréquence des données récupérables sur ce site. C'est pourquoi nous avons cherché une autre source de données qui nous permettrait de confronter les modèles sur des plages de temps différentes et de voir en quoi les différents modèles peuvent être complémentaires.

Pour mener à bien notre projet, nous avons cherché une source de données qui nous permettrait de tester la fiabilité de nos modèles sur des plages de temps différentes. Nous nous sommes donc tournés vers l'API Alphavantage, disponible en ligne et accessible gratuitement après la création d'un compte.

Cette source de données nous permet de récupérer des données quotidiennes sur une période de plusieurs mois au format CSV, ainsi que des données sur des plages de temps plus courtes, avec des délais entre les points variant de 1 minute à 1 heure.

Cependant, nous avons remarqué une légère différence de prix sur les actions, équivalant à 1 dollar, ce qui peut varier en fonction de l'action considérée. Nous avons tout de même considéré ces données comme acceptables pour la plupart des actions et leurs variations. En effet, cette différence de prix est constante sur la durée considérée, ce qui maintient la cohérence de nos modèles. Nous recommandons cependant de ne pas se baser uniquement sur ces données pour effectuer des investissements, mais plutôt de les utiliser en complément d'une source de données plus précise et payante.

En conclusion, les données que nous avons utilisées dans notre projet sont complémentaires. Les données de Nasdaq nous ont permis de tester la fiabilité des différents modèles face au modèle MSM en termes de précision numérique, tandis que les données d'Alphavantage nous ont permis de comparer les différents modèles sur des plages de temps différentes.

#### 5.2.4 • COMPARAISONS

Une fois que les critères de comparaison ont été définis, nous avons procédé à la comparaison de la performance de quatre modèles différents, à savoir le MSM, le MSM avec état initial, le GARCH et le RNN. Pour chaque série de données, nous avons utilisé 60% d'entre elles pour estimer les paramètres du modèle. Une fois que les paramètres ont été estimés, nous avons évalué séparément la performance de chaque modèle sur les données d'entraînement (in sample) et sur les données de validation (out of sample). Nous avons également simulé les données d'entraînement et les données futures qui n'ont pas été utilisées pour l'estimation des paramètres. Nous avons pris en compte le fait qu'être un bon prévisionniste out-of-sample ne garantit pas nécessairement un bon ajustement in-sample, et vice versa. Nous avons analysé ces deux qualités en gardant à l'esprit que c'est la prédiction qui est pertinente pour l'investisseur.

### 5.2.4.1 In Sample

MSM-binomial	horizon	MSE-Insample					R2-Insample					MAE-Insample				
		1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
kbar	fréquence															.
4	1-day	3.663	18.733	33.64	55.821	113.166	0.998	0.993	0.989	0.983	0.967	1.342	3.278	4.407	5.74	8.145
5	1-day	3.072	20.071	34.463	59.678	115.301	0.999	0.993	0.989	0.982	0.964	1.234	3.372	4.472	5.9	
6	1-day	7.808	54.97	94.536	156.5	296.335	0.997	0.983	0.972	0.982	0.902	1.967	5.567	7.374	9.57	
GARCH	MSE-Insample					R2-Insample					MAE-Insample					
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
	fréquence															.
(1,1)	1 day	0.08209	1.01357	3.57182	13.1221	80.3037	-0.0434	-0.1055	-0.1861	-0.3421	-0.8039	0.22144	0.77908	1.45941	2.79454	6.90238
RNN	MSE-Insample					R2-Insample					MAE-Insample					
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
	fréquence															.
	1-day	3.665	49.119				-2.238	-2.238				1.409	5.154			
MSM-initial state	MSE-Insample					R2-Insample					MAE-Insample					
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
	frequency															.
4	1-day	0.893	3.384	6.642	12.543	31.173	-0.313	-0.243	-0.231	-0.235	-0.251	0.446	0.982	1.445	2.103	3.591
5	1-day	0.843	2.952	5.704	11.871	30.764	-0.297	-0.213	-0.208	-0.215	-0.237	0.407	0.893	1.324	2.006	3.494
6	1-day	0.416	1.277	2.558	4.638	11.755	-0.149	-0.0934	-0.0908	-0.0846	-0.0863	0.277	0.572	0.847	1.239	2.11

TABLE 1 – Comparaison des modèles en série des données DEXJPUS in-sample

MSM-binomial	horizon	MSE-Insample					R2-Insample					MAE-Insample				
		1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
kbar	fréquence															.
4	1 min	0.107	0.497	0.904	1.492	2.374	-10.547	-9.663	-8.75	-7.366	-4.75	0.26	0.56	0.756	0.971	1.223
4	5 min	0.868	3.447	6.035	10.563	20.451	-15.946	-12.676	-11.097	-9.526	-7.54	0.709	1.455	1.939	2.572	3.586
4	15 min	0.34	1.086	1.655	2.721	6.071	-3.579	-2.278	1.747	-1.432	-1.258	0.413	0.797	0.999	1.293	1.939
5	1 min	0.129	0.536	0.882	1.369	2.268	-12.619	-10.434	-8.625	-6.74	-4.497	0.282	0.58	0.745	0.93	1.198
5	5 min	0.678	2.837	4.862	8.156	16.067	-12.473	-10.406	-8.856	-7.363	-5.779	0.615	1.319	1.738	2.261	3.174
5	15 min	0.268	0.774	1.193	2.05	4.987	-2.834	-1.623	-1.255	-1.082	-1.036	0.361	0.668	0.845	1.121	1.761
6	1 min	0.177	0.789	1.39	2.149	2.961	-17.264	-15.4	-13.44	-10.699	-5.914	0.333	0.706	0.938	1.167	1.367
6	5 min	1.548	7.07	11.995	20.271	39.846	-28.413	-25.781	-22.042	-18.393	-14.479	0.925	2.081	2.729	3.555	4.996
6	15 min	0.368	0.837	1.187	1.983	4.867	-3.875	-1.762	-1.25	-1.044	-1.013	0.427	0.695	0.841	1.1	1.745
GARCH	MSE-Insample					R2-Insample					MAE-Insample					
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
	fréquence															.
(1,1)	1 min	0.01627	0.08354	0.16898	0.36201	1.04295	-1.5844	-1.6181	-1.6605	-1.7868	-2.0705	0.10146	0.22966	0.32664	0.47794	0.81147
	5 min	0.0177	0.11585	0.29729	0.81129	3.22873	-0.322	-0.4261	-0.5435	-0.7463	-1.194	0.10553	0.27005	0.432	0.715	1.42344
	15 min	0.05228	0.48523	1.0245	2.07432	5.18822	-0.5517	-1.0207	-1.0938	-1.0908	-1.0919	0.18193	0.55335	0.80403	1.14685	1.81657
RNN	MSE-Insample					R2-Insample					MAE-Insample					
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
	fréquence															.
	1 min	0.0532	0.444				-4.323	-4.321				0.158	0.479			
	5 min	0.303	3.011				-4.382	-4.379				0.367	1.085			
	15 min	0.429	3.581				-3.572	-3.569				0.476	1.369			
MSM-initial state	MSE-Insample					R2-Insample					MAE-Insample					
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days
	frequency															.
4	1 min	0.0444	0.213	0.406	0.731	1.485	-4.298	-4.154	-3.919	-3.63	-2.963	0.15	0.336	0.471	0.648	0.953
4	5 min	0.421	1.869	3.561	6.885	16.21	-7.9	-6.982	-6.631	-6.313	-5.841	0.448	1.005	1.419	2.009	3.125
4	15 min	0.205	0.743	1.341	2.488	5.984	-2.163	-1.559	-1.419	-1.311	-1.265	0.205	0.636	0.888	1.237	1.929
5	1 min	0.0168	0.0825	0.159	0.569	0.68	-1.634	-1.598	-1.553	-2.849	-1.394	0.0915	0.207	0.292	0.57	0.63
5	5 min	0.253	1.092	2.105	4.059	9.806	-4.779	-4.09	-3.925	-3.803	-3.671	0.342	0.757	1.079	1.526	2.424
5	15 min	0.197	0.582	1.0344	2.0129	4.971	-2.079	-1.228	-1.093	-1.05	-1.042	0.281	0.564	0.781	1.11	1.762
6	1 min	0.0318	0.155	0.303	0.583	1.241	-3.0819	-3.014	-2.931	-2.85	-2.478	0.125	0.283	0.402	0.572	0.868
6	5 min	0.699	2.88	5.515	10.613	25.618	-13.146	-10.813	-10.362	-9.984	-9.589	0.569	1.234	1.754	2.476	3.914
6	15 min	0.201	0.576	1.013	1.965	4.936	-2.121	-1.212	-1.07	-1.021	-1.006	0.277	0.559	0.772	1.096	1.758

TABLE 2 – Comparaison des modèles en série des données IBM in-sample

### 5.2.4.2 Out Of Sample

MSM-binomial		MSE-outsamp					R2-outsamp					MAE-outsamp						
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days		
kbar	fréquence																	
		4	1-day	0.967	4.937	8.737	15.053	30.404	-1.354	-1.385	-1.216	-1.03	-0.842	0.708	1.726	2.308	3.047	4.342
		5	1-day	0.823	5.318	9.105	15.377	30.299	-1.158	-1.486	-1.286	-1.077	-0.842	0.653	1.779	2.355	3.081	4.334
RNN	fréquence	6	1-day	2.082	14.454	24.976	40.893	79.498	-2.909	-4.088	-3.51	-2.88	-2.217	1.04	2.926	3.898	5.01	7.025
		MSE-outsamp					R2-outsamp					MAE-outsamp						
		1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days		
GARCH	fréquence																	
		4	1-day	1.503	16.776				-2.102	-2.217				0.924	3.145			
		5	1-day															
MSM-initial state	fréquence	6	1-day	0.0218	0.2688	0.93734	3.43616	20.2164	-0.031	-0.0762	-0.132	-0.2472	-0.5772	0.117	0.4102	0.76538	1.46422	3.5420
		MSE-outsamp					R2-outsamp					MAE-outsamp						
		1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days		
kbar	horizon																	
		4	1-day	0.0275	0.145	0.315	0.754	2.613	-0.0387	-0.0406	-0.0451	-0.0523	-0.0723	0.115	0.284	0.429	0.673	1.25 S
		5	1-day	0.0214	0.115	0.266	0.659	2.475	-0.029	-0.0329	-0.0372	-0.0482	-0.0704	0.0991	0.248	0.387	0.624	1.23
MSM-initial state	horizon	6	1-day	0.00973	0.0475	0.106	0.243	0.846	-0.0137	-0.0131	-0.0146	-0.0183	-0.023	0.0676	0.161	0.247	0.381	0.72

TABLE 3 – Comparaison des modèles en série des données DEXJPUS out-of-sample

MSM-binomial		MSE-outsamp					R2-outsamp					MAE-outsamp						
	horizon	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days		
kbar	fréquence																	
		4	1 min	0.113	0.523	0.943	1.58	2.484	-12.335	-11.329	-10.16	-8.392	-5.303	0.267	0.575	0.774	1.001	1.249
		5	1 min	0.87	3.472	6.155	10.417	20.771	-8.576	-6.848	-6.017	-5.112	-4.132	0.703	1.463	1.96	2.553	3.606
4	fréquence	15 min	0.289	0.914	1.431	2.306	5.159	-1.551	-0.98	-0.764	-0.615	-0.539	0.381	0.733	0.928	1.193	1.794	
		1 min	0.135	0.558	0.932	1.434	2.39	-14.704	-12.135	-10.059	-7.711	-5.042	0.289	0.593	0.766	0.95	1.225	
		5 min	0.684	2.87	4.886	8.246	16.28	-6.755	-5.662	-4.779	-3.989	-3.173	0.618	1.328	1.745	2.275	3.201	
5	fréquence	15 min	0.227	0.658	1.002	1.755	4.257	-1.217	-0.706	-0.535	-0.464	-0.442	0.332	0.616	0.775	1.037	1.635	
		1 min	0.185	0.827	1.456	2.229	3.161	-20.133	-18.002	-15.662	-12.155	-10.656	0.341	0.723	0.959	1.19	1.415	
		5 min	1.57	7.105	12.179	20.505	40.699	-15.482	-14.054	-11.943	-10.001	-7.979	0.932	2.088	2.75	3.585	5.073	
6	fréquence	15 min	0.313	0.716	1.011	1.737	4.178	-1.676	-0.766	-0.538	-0.45	-0.431	0.394	0.642	0.777	1.027	1.616	
		MSE-outsamp					R2-outsamp					MAE-outsamp						
		1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days		
GARCH	fréquence																	
		(1,1)	1 min	0.01694	0.08709	0.17909	0.37394	1.07401	-1.8348	-1.884	-1.9406	-2.0374	-2.3216	0.10365	0.23469	0.336346	0.4868	0.82101
		5 min	0.01788	0.11808	0.30054	0.82461	3.25418	-0.1769	-0.233	-0.2946	-0.4069	-0.6513	0.10619	0.27284	0.43517	0.72143	1.43294	
RNN	fréquence	15 min	0.0442	0.4108	0.88082	1.77048	4.33886	-0.2366	-0.4376	-0.4703	-0.4672	-0.47	0.1673	0.50915	0.74587	1.05859	1.65491	
		MSE-outsamp					R2-outsamp					MAE-outsamp						
		1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days		
MSM-initial state	fréquence																	
		4	1 min	0.0506	0.343				-4.389	-4.327				0.15	0.426			
		5	1 min	0.479	1.934				-4.875	-3.353				0.584	1.028			
kbar	horizon	15 min	0.83	5.425				-3.365	-3.419				0.612	1.657				
		MSE-outsamp					R2-outsamp					MAE-outsamp						
		1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days	1-day	5-days	10-days	20-days	50-days		
4	fréquence																	
		4	1 min	0.381	1.681	3.157	6.01	14.144	-3.688	-3.253	-3.078	-2.893	-2.726	0.425	0.943	1.32	1.864	2.921
		5	15 min	0.105	0.44	0.913	1.904	5.023	-0.556	-0.464	-0.485	-0.509	-0.528	0.201	0.48	0.728	1.079	1.769
5	fréquence	1 min	0.0445	0.211	0.393	0.688	1.343	-4.806	-4.541	-4.224	-3.755	-2.843	0.151	0.336	0.466	0.628	0.906	
		5 min	0.233	0.96	1.805	3.42	8.362	-2.245	-1.845	-1.74	-1.659	-1.624	0.327	0.705	0.991	1.39	2.209	
		15 min	0.102	0.368	0.778	1.67	4.278	-0.545	-0.395	-0.414	-0.438	-0.442	0.191	0.444	0.678	1.012	1.632	
6	fréquence	1 min	0.0413	0.199	0.383	0.706	1.546	-4.474	-4.322	-4.15	-3.809	-3.09	0.144	0.323	0.457	0.632	0.963	
		5 min	0.631	2.468	4.547	8.611	20.176	-6.093	-4.748	-4.425	-4.148	-3.968	0.539	1.133	1.571	2.202	3.455	
		15 min	0.1008	0.357	0.746	1.607	4.12	-0.538	-0.38	-0.397	-0.424	-0.43	0.184	0.435	0.661	0.99	1.606	

TABLE 4 – Comparaison des modèles en série des données IBM out-of-sample

### 5.2.4.3 Commentaires

Au cours de l'analyse des données, nous avons constaté que le modèle RNN n'était pas en mesure de prédire les valeurs sur le long terme. Les prédictions se dégradaient rapidement à mesure que l'écart temporel augmentait, et semblaient converger vers un point fixe lorsque la valeur prédictée était utilisée en entrée pour obtenir la prédition suivante. Nous avons donc décidé de nous concentrer sur la performance du modèle RNN à court terme, c'est-à-dire à 1 jour et 5 jours, dans notre comparaison avec

les modèles GARCH, MSM et MSM avec état initial (MSMI). Selon les résultats de la comparaison, en général, le modèle GARCH, MSM et MSMI sont plus performants que le modèle RNN pour toutes les horizons et fréquences. De plus, le MSMI est généralement meilleur que le MSM dans la plupart des cas. Cependant, la comparaison entre le GARCH et le MSMI est plus complexe. Dans les séries de données plus volatiles telles que les rendements d'actions IBM et BA, le modèle GARCH se comporte généralement mieux que le MSMI ; tandis que dans les séries de données moins volatiles telles que DEXJPUS, c'est le modèle MSMI qui prédomine. De plus, la performance du modèle MSMI est meilleure lorsque nous choisissons un horizon plus grand et une fréquence plus basse dans une série de données donnée.

## 5.3 NOTES SUR LES ALGORITHMES ET IMPLÉMENTATIONS

---

### 5.3.1 • BASIN-HOPPING

Lorsqu'on cherche à maximiser la qualité d'un modèle statistique, on peut recourir à l'algorithme de Basin-hopping. Ce dernier permet de trouver les valeurs optimales des paramètres d'un modèle en maximisant la log-vraisemblance, qui est une mesure de la qualité du modèle. En effet, la log-vraisemblance permet de mesurer la probabilité d'observer les données en fonction des paramètres du modèle. Plus cette probabilité est grande, meilleure est la qualité du modèle. Ainsi, l'algorithme de Basin-hopping peut être un outil précieux pour trouver les meilleurs paramètres et donc améliorer la performance du modèle statistique.

Le pseudo-code général de l'algorithme de Basin-hopping pour la maximisation de la log-vraisemblance est le suivant :

- Initialisez les paramètres initiaux du modèle  $\theta$ .
- Boucle :
  1. Utilisez la dynamique moléculaire pour trouver un minimum local de la log-vraisemblance en utilisant les paramètres actuels  $\theta$ .
  2. Générez des nouveaux paramètres  $\theta'$  en utilisant une méthode de recherche aléatoire, telle qu'une perturbation gaussienne ou un saut de Monte Carlo.
  3. Calculez la log-vraisemblance pour les paramètres  $\theta$  et  $\theta'$ .
  4. Si la log-vraisemblance de  $\theta'$  est supérieure à celle de  $\theta$ , mettez à jour  $\theta$  pour être égale à  $\theta'$ . Sinon, acceptez ou rejetez les nouveaux paramètres en utilisant une probabilité de Metropolis-Hastings.
  5. Répétez jusqu'à ce que la convergence soit atteinte ou que le nombre maximal d'itérations soit atteint.

Il est important de noter que les étapes de la dynamique moléculaire et de la recherche aléatoire peuvent varier considérablement en fonction de la spécificité du système étudié et des besoins spécifiques de l'application.

En résumé, l'algorithme de Basin-hopping utilise une combinaison de méthodes de dynamique moléculaire et de recherche aléatoire pour explorer efficacement l'espace des solutions et trouver les paramètres optimaux qui maximisent la log-vraisemblance d'un modèle statistique.

## 6

# SKEW MSM : UN MODÈLE AJUSTÉ AUX OPTIONS

### 6.1 MOTIVATION

Les investisseurs remarquent la dissymétrie lorsqu'ils évaluent la distribution des rendements, car celle-ci, tout comme l'aplatissement, prend en compte les extrêmes de l'ensemble de données plutôt que de se concentrer uniquement sur la moyenne. Les investisseurs à court et moyen terme doivent en particulier regarder les valeurs aberrantes car ils sont moins susceptibles de maintenir une position suffisamment longtemps pour espérer que le rendement s'équilibre autour de sa moyenne.

Le risque d'asymétrie -skewness- est le risque accru de trouver un point de données présentant une forte dissymétrie dans une distribution asymétrique. De nombreux modèles financiers qui tentent de prédire la performance future d'un actif supposent une distribution normale, dans laquelle les mesures de tendance centrale sont égales. Si les données sont asymétriques, ce type de modèle sous-estimera toujours le risque de dissymétrie dans ses prévisions. Plus les données sont asymétriques, moins précis sera ce modèle financier.

**Skewness :** Une mesure statistique de l'asymétrie d'une distribution de données par rapport à sa moyenne.

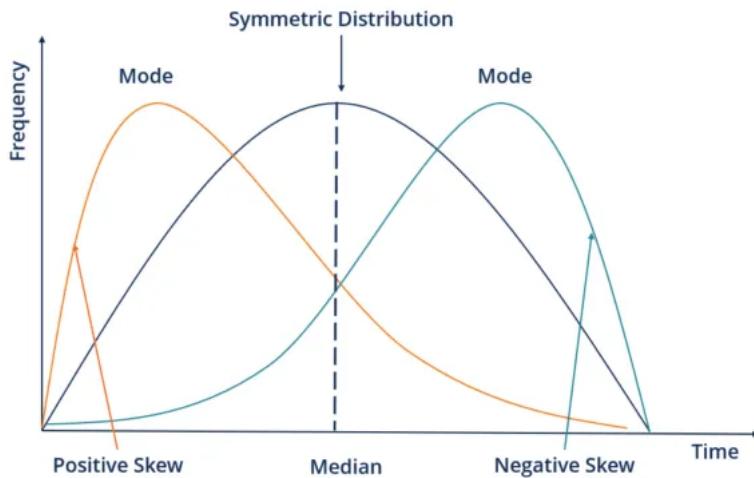


FIGURE 8 – Illustration de la skewness (source : CFI)

**NB.** La dissymétrie indique où se produisent les valeurs aberrantes, mais elle ne donne pas d'indication sur le nombre de valeurs aberrantes.

Le marché boursier global est souvent considéré comme ayant une distribution négativement dissymétrique, ce qui signifie que le marché a plus souvent des rendements légèrement positifs que de fortes pertes négatives. Cependant, des études ont montré que les capitaux propres d'une entreprise individuelle peuvent avoir tendance à être dissymétriques vers la gauche.

Un exemple courant de dissymétrie est la distribution des revenus des ménages aux États-Unis, car les individus ont moins de chances de gagner des revenus annuels très élevés.

## 6.2 LE MODÈLE SKEW MSM

Le modèle SkewMSM (Skewed Markov-Switching Multifractal) est une extension du modèle MSM (Markov-Switching Multifractal) qui permet de prendre en compte la présence de skews (asymétries) dans la distribution de la volatilité. Ce modèle a été proposé pour modéliser la volatilité des marchés financiers, où les rendements peuvent être fortement asymétriques.

### 6.2.1 • UNE CLASSE DE MODÈLE DE DIFFUSION CONDITIONNELLEMENT AFFINE

On se donne  $(M_t)$  une chaîne de Markov à temps continu sur un ensemble d'état  $E = \{m^1, \dots, m^d\}$ ,  $Q = (Q_{i,j})$  sa matrice de transition où  $q_{i,j} = \mathbb{P}(M_{t+dt} = m^j | M_t = m^i)$ . On note  $N_{i,j,t}$  le nombre de transitions entre l'état  $m^i$  et l'état  $m^j$  entre 0 et  $t$  si  $i \neq j$ ,  $N_{i,j,t} = 0$  sinon.

Soit  $S_t$  le prix d'un actif sans risque à l'instant  $t$ ,  $r_t$  son taux d'intérêt composé, et  $d_t$  son rendement en dividendes. Le rendement excédentaire infinitésimal est alors donné par :

$$\frac{dS_t}{S_t} + (d_t - r_t) dt$$

On se propose alors de donner un modèle pour les log-return  $\log(S_t)$  :

$$ds_t = \mu_t dt + \sqrt{v_t} dW_{1,t} + dJ_t - J_{M_t}^- dt$$

$$dv_t = \kappa (\theta_{M_t} - v_t) dt + \sigma \sqrt{v_t} \left( \rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right)$$

avec  $\kappa, \sigma \geq 0, \rho \in [-1, 1]$  et  $\theta_{M_t} \in \mathbb{R}_+^*$  une fonction déterministe de  $M_t$ .

$W_t = (W_{1,t}, W_{2,t})$  un processus de Wiener à deux dimensions :

- $W_0 = (0, 0)$
- $W_{t+u} - W_u$  est indépendant de  $W_s$  pour tout  $s \leq t$
- $W_{t+u} - W_u$  suit une loi normale de variance  $u$
- $W_t$  est continue en  $t$

Dans le cadre de ce modèle,  $v_t$  est un processus de variance stochastique,  $dJ_t - J_{M_t}^- dt$  est un processus de sauts compensés, introduisant des discontinuités dans le modèle.

On donne ainsi :

$$J_t = \beta \sum_{1 \leq i, j \leq d} J_{i,j} N_{i,j,t}$$

$$J_{M_t}^- = \mathbb{E} \left( e^{dJ_t - 1} | M_{t-dt} \right) \simeq \mathbb{E} (dJ_t | M_{t-dt})$$

où  $J_{i,j} \in \mathbb{R}$  et  $J_{i,i} = 0$ . Cela montre que  $J_t$  modélise des "sauts" dans la valeur des log-return, tandis que  $J_{M_t}$  est un terme de compensation.

Enfin,  $\mu_t$  est donné par :

$$\mu_t = r_t - d_t + \left( \alpha - \frac{1}{2} \right) v_t \text{ avec } \alpha \text{ constant}$$

de telle sorte à ce que :

$$\mathbb{E}_{\mathbb{P}} \left( \frac{dS_t}{S_t} + (d_t - r_t) | M_t, v_t \right) = \mathbb{E}_{\mathbb{P}} (ds_t + (d_t - r_t) | M_t, v_t) = \alpha v_t$$

### 6.2.2 • MODÈLE AJUSTÉ AU RISQUE

Intéressons nous dorénavant à la mesure de probabilité neutre à l'égard du risque, notée  $\mathbb{Q}$  par la suite. Celle-ci correspond à une actualisation de la mesure physique  $\mathbb{P}$  (la "vraie" probabilité qu'un événement se produise), au regard du risque. Cela consiste dans les faits à augmenter la probabilité de mauvais résultats du marché, et à baisser celle de bons résultats. Cette mesure décrit en fait le prix effectivement observé d'un actif sur le marché considéré.

La mesure  $Q$  a la dérivée de Radon-Nikodym suivante :

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \xi_t^{(1)} \xi_t^{(2)}$$

c'est-à-dire que :

$$\forall A, \mathbb{Q}(A) = \int_A \xi_t^{(1)} \xi_t^{(2)} d\mathbb{P}$$

Dans ce cadre, les probabilités de transition de  $(M_t)$  sont changées comme suit :

$$q_{i,j}^* = q_{i,j} \frac{\Lambda_j}{\Lambda_i} \text{ pour tout } i \neq j$$

$$q_{i,i}^* = - \sum_{k \neq i} q_{i,k}^*$$

avec  $\Lambda_i$  un coefficient strictement positif, appelé coefficient de prime de risque . Si l'état  $j$  et l'état  $i$  vérifient  $\Lambda_j > \Lambda_i$ , la probabilité de transition du "bon" état  $i$  vers le "mauvais" état  $j$  est plus grande sous  $\mathbb{Q}$  que sous  $\mathbb{P}$  comme annoncé. On note  $Q^*$  la matrice de transition sous  $\mathbb{Q}$ .

Sous  $\mathbb{Q}$ , on a alors :

$$ds_t = (r_t - d_t - \frac{V_t}{2}) dt + \sqrt{V_t} dW_{1,t}^* + dJ_t - J_{M_t}^* dt$$

$$dV_t = \kappa^*(\theta_{M_t}^* - V_t) dt + \sigma \sqrt{V_t} (\rho dW_{1,t}^* + \sqrt{1-\rho^2} dW_{2,t}^*)$$

avec  $\kappa^*$  constant,  $\theta_{M_t}^* = \kappa \frac{\theta_{M_t}}{\kappa^*}$ ,  $W_t^*$  un processus de Wiener à deux dimensions sous  $\mathbb{Q}$ , et  $J_{M_t}^* = \mathbb{E}^{\mathbb{Q}}(dJ_t | M_{t-dt})$

### 6.2.3 • LE MODÈLE SKEWMSM

Le modèle SkewMSM est en fait un choix correct de paramètres dans la classe de modèle précédemment définie. Cela revient à supposer que la chaîne de Markov  $M_t = (M_{1,t}, \dots, M_{\bar{k},t} \in \mathbb{R}_+^{\bar{k}})$  suit un processus MSM.

On choisit de plus :

$$\theta_{M_t} = \bar{\sigma} \prod_{k=1}^{\bar{k}} M_{k,t}$$

avec  $\bar{\sigma}$  comme définie précédemment.

On pose enfin :

$$J_{i,j} = \beta \sum_{k=1}^{\bar{k}} \frac{1}{\gamma_k} \text{signe}(m_k^j - m_k^i)$$

où  $m_k^j$  désigne le  $k$ -ième composant de l'état  $m^j$ .

Pour l'estimation de la vraisemblance, nous allons calculer un estimateur de cette dernière en utilisant la méthode *bootstrap filter algorithm*.

1. On commence par créer pour tout  $k \in 1, \dots, K$  une particule  $\chi_t^k = (M_{t-1}, M_t, v_t)$  que l'on propagera par la suite.

2. A partir de chacune de ces particules, on regénère une deuxième que l'on notera :  $\tilde{\chi}_{t+1}^k$ .

On commence par tirer  $M_{t-1}$

Ensuite, on simule une loi gaussienne  $\epsilon_{2,t}$  :

En combinant les résultats nous obtenons :

$$v_{t+1} = v_t + \kappa(\theta_{M_t} - v_t)dt + \sigma\rho(y_t - \mu_t\Delta t - (dJ_t - \tilde{J}_{M_{t-1}}\Delta t)) + \sigma\sqrt{1 - \rho^2}\epsilon_{2,t}$$

Ainsi, la connaissance de  $M_t, M_{t-1}$  et  $v_t$  nous permet de générer la nouvelle particule.

3. Nous arrivons ainsi à la phase de *resampling* : une pondération  $\mathbb{P}(r_{t+1}|\tilde{\chi}_{t+1}^k)$  sera associée à chaque particule.

A partir de ces pondérations nous pourrons tirer  $K$  nouvelles particules et déduire l'estimateur de vraisemblance :

$$\hat{L}(r_1, \dots, r_{t+1}) = \hat{L}(r_1, \dots, r_t) \left( \frac{1}{K} \sum_{k=1}^K \mathbb{P}(r_{t+1} | \tilde{\chi}_{t+1}^k) \right)$$

### 6.2.4 • SKEWMSM AJUSTÉ AU RISQUE

Pour utiliser la mesure ajustée au risque, il reste à choisir  $\Lambda_i$  :

$$\Lambda_i = \prod_{k=1}^{\bar{k}} (b^*)^{\bar{k}-k} \mathbb{1}_{(m_k^i = 2-m_0)} + \frac{1}{b^{\bar{k}-k}} \mathbb{1}_{(m_k^i = m_0)}$$

avec  $b^* \geq 1$ .

Nous avons donc présenté le modèle skewMSM. Cependant, l'article propose également une variante ajustée au risque. Ainsi, il s'agit de rajouter un "risk premium" pour la transition  $M_t^k = 2 - m_0 \rightarrow M_{t+1}^k = m_0$ . Ce "risk premium" est donné par

$$rp = (b^*)^{2(\bar{k}-k)}$$

c'est à dire que la transition  $M_t^k = 2 - m_0 \rightarrow M_{t+1}^k = m_0$  a pour probabilité  $\frac{\lambda_k}{2} rp$  et la transition inverse  $M_t^k = m_0 \rightarrow M_{t+1}^k = 2 - m_0$  a pour probabilité  $\frac{\lambda_k}{2rp}$

Finalement, le modèle SkewMSM est entièrement déterminé par les variables suivantes :

$$\Psi_{\mathbb{P}} = (\gamma_1, b, m_0, \beta, \kappa, \bar{\sigma}, \sigma, \rho, \alpha)$$

Le modèle ajusté au risque dépend lui uniquement de deux paramètres additionnels :

$$\Psi_{\mathbb{Q}} = (\kappa^*, b^*)$$

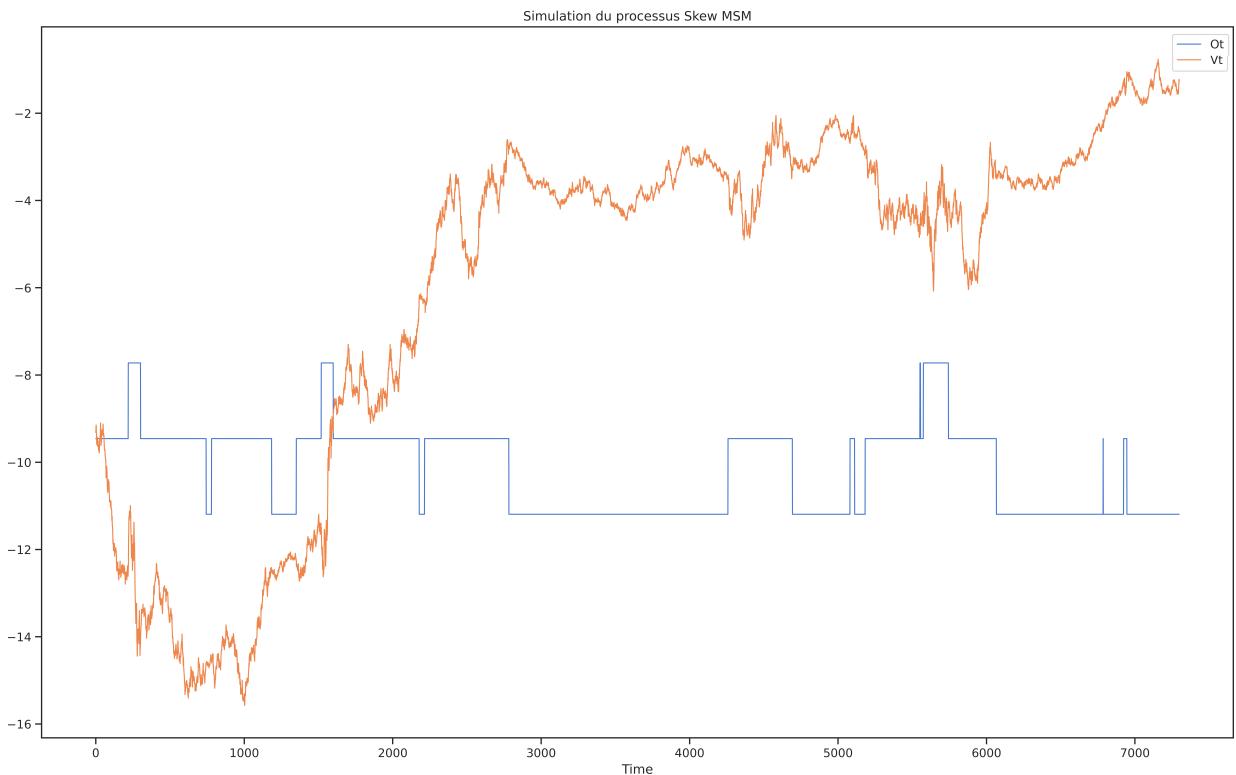


FIGURE 9 – Simulation du processus Skew MSM. En bleu le processus  $\theta_{M_t}$  et en orange la variance  $v_t$

### 6.3 DESCRIPTION DES PARAMÈTRES DU MODÈLE

Dans cette section, nous allons décrire ce que modélise chacun des paramètres précédemment étudiés dans le modèle Skew-MSM.

- $\gamma_{\bar{k}} \in [0; 1]$  est la fréquence à laquelle  $M_{\bar{k},t}$  est recalculé.  $M_{\bar{k},t}$  étant l'état avec la fréquence de recalcul la plus importante, les états  $M_{j,t}$  ont des fréquences inférieures et proportionnelles à  $\frac{\gamma_{\bar{k}}}{\bar{k}}$ .
- $\frac{1}{b} \in [0; 1]$  est le coefficient de proportionnalité entre  $\gamma_{\bar{k}}$  et  $\gamma_{\bar{k}-1}$ , puis entre  $\gamma_{\bar{k}-1}$  et  $\gamma_{\bar{k}-2}$ , etc. -  $m_0 \in ]1; 2[$  mesure l'impact d'un changement d'état de  $M_{\bar{k},t}$  sur la volatilité.
- $\beta$  décrit l'amplitude des sauts dans la valeur des log-return.  $\beta > 0$  indique une brusque augmentation des prix, suivi d'une phase de volatilité élevée, tandis que  $\beta < 0$  indique une diminution des prix (suivie elle aussi d'une phase de volatilité élevée).
- $\kappa \in ]0; \frac{1}{dt}[$  indique la vitesse à laquelle la variance  $v_t$  s'ajuste sur la valeur  $\theta_{M_t}$ , à la fois lors d'un changement d'état de  $M_t$  et lors d'une variation de  $v_t$ .
- $\sigma$  détermine la variance du couplage entre  $s_{t-1}$  et  $v_t$ .
- $\rho$  détermine le couplage entre  $s_{t-1}$  et  $v_t$  via la formule :

$$\sigma\rho = \text{Cov}\left(W_{1,t}, \sigma \left(\rho W_{1,t} + \sqrt{1-\rho^2}W_{2,t}\right)\right)$$

- $\bar{\sigma}$  correspond à la variance moyenne du terme  $\sqrt{v_t}dW_{1,t}$ .
- $\alpha$  correspond à l'excess-return, c'est-à-dire le gain espéré quand la volatilité augmente.
- $b^*$  mesure combien une transition d'un état à faible volatilité vers un état à forte volatilité est plus probable que la transition inverse.

## 6.4 IMPLÉMENTATION DE SKEW MSM

---

### 6.4.1 • NOTES SUR L'IMPLÉMENTATION

Nous avons eu l'opportunité de pouvoir accéder aux implémentations du modèle de M. Calvet. Bien que son code -en langage MATLAB- ait fourni des résultats précis conformes à ceux présentés dans l'article, notre travail principal a consisté à le réimplémenter en langage Python. Cette démarche nous a permis de mieux comprendre le modèle et ses affinités, tout en utilisant un langage de programmation plus familier.

Toutefois, notre implémentation reste loin d'être parfaite : minimas locaux, valeurs aberrantes et échec de convergence sont apparus lors des simulations. Nous avons réussi à reprendre une partie des résultats de l'article de M. Calvet grâce à notre propre algorithme, tandis que pour le reste, nous avons utilisé directement le code fourni par M. Calvet.

### 6.4.2 • ESTIMATION SUR LES DONNÉES D'EQUITY

**Tableau 5** : Estimation sur les données d'equity.

$\bar{k}$	$\lambda_{\bar{k}}$	$b$	$m_0$	$\beta$	$\kappa$	$\sigma$	$\rho$	$\bar{\theta}$	$\alpha$	$\ln L$
1	6.9E - 3	-	1.04	-2.4E - 4	3.0	0.26	-0.56	0.020	1.01	44,01
2	8.8E - 3	5.87	1.16	-3.1E - 5	3.8	0.25	-0.57	0.020	1.00	44,047
3	2.0E - 2	5.50	1.27	-1.3E - 4	5.0	0.29	-0.58	0.022	1.15	44,058
4	9.2E - 3	2.63	1.30	-9.7E - 5	9.7	0.27	-0.60	0.025	1.41	44,092
5	9.9E - 3	2.49	1.38	-4.5E - 5	16.6	0.27	-0.70	0.020	1.32	44,115
6	1.1E - 2	2.42	1.40	-2.9E - 5	19.7	0.28	-0.64	0.019	1.39	44,120

Ce tableau rend compte des résultats obtenus pour le modèle SkewMSM en utilisant des log excess returns S&P 500 du 6 mars 1957 jusqu’au 28 septembre 2007. Chaque ligne correspond à un nombre de composants de fréquences  $\bar{k}$ . La première ligne correspond au modèle à une composante ( $\bar{k} = 1$ ), dans lequel la volatilité oscille entre deux valeurs ( $m_0$  ou  $2 - m_0$ ). et où la taille des sauts est déterministe et est unique. L’estimation de  $\lambda_{\bar{k}}$  correspond à un changement pour la volatilité à long terme qui arrive en moyenne  $240 \times 0.0069 = 1.7$  fois tous les ans. L’estimation de  $m_0$  est proche de 1, ce qui implique que les deux états de  $\theta_{M_t}$  sont proches les uns des autres.

On remarque de plus, que plus  $\bar{k}$  est grand, plus le taux de croissance  $b$  diminue, ce qui indique que les fréquences deviennent de plus en plus proches les unes des autres. À l’inverse,  $\lambda_{\bar{k}}$  et  $m_0$  augmentent, ce qui montre que la composante "diffusive" joue un rôle de plus en plus important dans le modèle. En résumé, plus  $\bar{k}$  est grand, plus les sauts sont petits mais fréquents.

Les paramètres  $\kappa$ ,  $\sigma$  et  $\rho$ , qui déterminent le comportement du terme  $V_t$  exhibent aussi des patrons que l’on peut analyser. Plus  $\bar{k}$  est grand, plus  $V_t$  a tendance à se comporter comme un modèle MSM binomial classique et la vitesse d’ajustement  $\kappa$  est plus rapide. Pour  $\bar{k}$  petit, on retrouve des valeurs proches de celles trouvées dans des modèles de volatilité stochastiques classiques.

## 6.5 BENCHMARKS

---

Les prix des options contiennent un ensemble riche d’informations sur la distribution conditionnelle du processus du prix du sous-jacent, ce qui devrait être utile pour filtrer les états latents. Suivant cette intuition, nous estimons Skew MSM, SV, SVJ0 et SVJ1 à la fois sur les rendements et les options. À notre connaissance, l’article de Mr. Calvet fût le premier à mener une estimation ML et un filtrage utilisant à la fois des options et des rendements pour l’un de ces modèles, y compris les modèles de référence.

Nous fondons notre analyse empirique sur des données d’options call de l’indice S&P 500 mensuel sur la période de janvier 1998 à décembre 2012. Par conséquent, l’échantillon comprend la récente crise financière. L’estimation sur les données d’options est difficile à réaliser par le calcul computationnel. Nous sélectionnons des contrats d’options à partir du premier mercredi de chaque mois de la période d’échantillonnage avec des échéances entre 0,2 et 1 an et des niveaux de moneyness entre 0,9 et 1,1. Nous obtenons ainsi une série de 117 surfaces d’options (voir 1.5.3). Les surfaces d’options ont tendance à être très persistantes, de sorte que le grand écart de temps entre chaque surface d’option dans notre échantillon permet d’estimer les paramètres de prime de risque sur une grande variété de surfaces d’IV.

### 6.5.1 • SPÉCIFICATIONS

Dans l’analyse empirique de cette section, nous comparons Skew MSM à des modèles jump-diffusion en temps continu standards. Nous fournissons maintenant les spécifications de ces modèles de référence.

Les modèles de jump-diffusion sont une famille de processus en temps continu flexibles et largement utilisée dans la finance. Nous considérons des spécifications de la forme

$$\frac{dS_t}{S_t} = \mu(V_t) dt + \sqrt{V_t} dW_{1,t} + Z_t$$

$$dV_t = \kappa(\theta - V_t) dt + \sigma V_t^b \left( \rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right),$$

où  $W_{1,t}$  et  $W_{2,t}$  sont des processus de Wiener indépendants, et  $Z_t$  est un pure-jump process.

On considère en particulier :

- SV :  $Z_t = 0$  et  $b = 1/2$ .
- SVJ0 :  $Z_t$  est un processus de Poisson d'intensité  $\lambda$  et  $b = 1/2$ .
- SVJ1 :  $Z_t$  est un processus de Poisson d'intensité  $\lambda V_t$  et  $b = 1/2$ .

### 6.5.2 • RÉSULTATS EMPIRIQUES

**Tableau 6** Estimation ML In-sample pour les données d'options.

	$\kappa^*$	$b^*$	$Var(e_i)$	ln L	BIC	p-valeur
Skew MSM(4)	4.50	1.00	0.284	6900	-13,743	
Skew MSM(3)	2.70	1.05	0.286	6722	-13,388	9.65 E-2
Skew MSM(2)	1.62	1.05	0.296	5978	-11,900	1.50 E-8
Skew MSM(1)	1.38		0.306	5315	-10,578	1.41 E-11
	$\kappa^*$	$\lambda^*$	$Var(e_i)$	ln L		
SV	2.77		0.363	3966	-7903	0.00E - 0
SVJ0	0.70	0.62	0.304	5587	-11,121	1.54E - 12
SVJ1	1.27	65.1	0.308	5396	-10,739	1.61E - 14

Le tableau précédent présente les paramètres de prime de risque et d'erreur d'observation d'options pour Skew MSM avec  $\bar{k} \in [1; 4]$ . SVJ0 et SVJ1, sont estimés à l'aide d'une méthode de ML en deux étapes sur un panel de données d'options mensuelles sur la période janvier 1998-septembre 2007. Le critère d'information bayésien est donné par  $BIC = -2 \ln L + NP \ln T$ . La dernière colonne donne les valeurs  $p$  d'un test de domination de Skew MSM par le modèle correspondant en utilisant le critère BIC ceci par la méthodologie de Vuong (1989). Une faible valeur  $p$  indique que le modèle correspondant serait rejeté en faveur de Skew MSM(4).

### 6.5.3 • RÉSULTATS SUR LES DONNÉES D'OPTIONS

Nous avons sélectionné un ensemble de 117 surfaces d'options, définies comme étant le premier mercredi de chaque mois entre janvier 1998 et septembre 2007 pour notre analyse In-Sample. Nous avons catégorisé les données en trois groupes de maturité et six groupes de moneyness définis comme suit : (1) :  $0.9 \leq S/K \leq 0.95$ , (2) :  $0.95 < S/K \leq 0.975$ , (3) :  $0.975 < S/K \leq 1.0$ , (4) :  $1.0 < S/K \leq 1.025$ , (5) :  $1.025 < S/K \leq 1.05$ , (6) :  $1.05 < S/K \leq 1.1$ .

**Tableau 7 :** L'erreur quadratique moyenne (RMSE) des log-prices dess options call regroupée par la Moneyness (différence entre le prix d'exercice d'une option et le prix actuel du sous-jacent) et la maturité.

	A. In-sample (Jan. 1998-Sept. 2007)							B. Out-of-sample (Oct. 2007-Dec. 2012)						
	(1)	(2)	(3)	(4)	(5)	(6)	(1) – (6)	(1)	(2)	(3)	(4)	(5)	(6)	(1) – (6)
0.2 < $\tau$ < 0.4 Skew MSM( $\bar{k} = 1$ )	0.102	0.076	0.083	0.078	0.075	0.060	0.081	0.163	0.067	0.094	0.102	0.090	0.079	0.113
Skew MSM( $\bar{k} = 2$ )	0.080	0.075	0.086	0.082	0.077	0.062	0.077	0.130	0.068	0.102	0.110	0.097	0.084	0.104
Skew MSM( $\bar{k} = 3$ )	0.068	0.063	0.070	0.070	0.067	0.054	0.065	0.121	0.066	0.097	0.105	0.096	0.083	0.099
Skew MSM( $\bar{k} = 4$ )	0.059	0.065	0.073	0.073	0.070	0.058	0.065	0.101	0.069	0.102	0.111	0.101	0.090	0.097
SV	0.139	0.069	0.070	0.064	0.061	0.048	0.087	0.168	0.070	0.086	0.090	0.077	0.067	0.111
SVJ0	0.091	0.083	0.084	0.076	0.071	0.057	0.078	0.138	0.065	0.094	0.101	0.090	0.078	0.103
SVJ1	0.094	0.075	0.079	0.074	0.070	0.055	0.077	0.145	0.066	0.091	0.099	0.087	0.076	0.105
0.4 < $\tau$ < 0.6 Skew MSM( $\bar{k} = 1$ )	0.087	0.077	0.086	0.079	0.077	0.065	0.079	0.093	0.117	0.126	0.125	0.122	0.105	0.112
Skew MSM( $\bar{k} = 2$ )	0.072	0.067	0.080	0.075	0.074	0.063	0.071	0.064	0.082	0.103	0.106	0.111	0.099	0.092
Skew MSM( $\bar{k} = 3$ )	0.055	0.051	0.060	0.059	0.060	0.052	0.056	0.058	0.073	0.096	0.100	0.101	0.095	0.086
Skew MSM( $\bar{k} = 4$ )	0.051	0.045	0.060	0.059	0.063	0.055	0.055	0.068	0.054	0.078	0.090	0.091	0.088	0.078
SV	0.141	0.085	0.080	0.073	0.069	0.059	0.096	0.115	0.129	0.135	0.129	0.125	0.106	0.122
SVJ0	0.084	0.068	0.075	0.069	0.068	0.058	0.072	0.073	0.096	0.109	0.111	0.110	0.096	0.097
SVJ1	0.092	0.071	0.076	0.070	0.068	0.058	0.075	0.083	0.105	0.115	0.115	0.112	0.098	0.102
0.6 < $\tau$ < 1 Skew MSM( $\bar{k} = 1$ )	0.121	0.080	0.079	0.080	0.073	0.068	0.090	0.133	0.149	0.146	0.151	0.142	0.121	0.138
Skew MSM( $\bar{k} = 2$ )	0.094	0.059	0.063	0.068	0.064	0.062	0.073	0.082	0.102	0.101	0.115	0.117	0.106	0.101
Skew MSM( $\bar{k} = 3$ )	0.075	0.052	0.055	0.059	0.057	0.056	0.062	0.072	0.088	0.099	0.110	0.108	0.101	0.094
Skew MSM( $\bar{k} = 4$ )	0.060	0.044	0.051	0.059	0.060	0.058	0.056	0.059	0.055	0.070	0.088	0.088	0.086	0.074
SV	0.175	0.110	0.098	0.092	0.078	0.071	0.119	0.186	0.190	0.172	0.176	0.163	0.139	0.172
SVJ0	0.122	0.068	0.064	0.065	0.060	0.057	0.083	0.094	0.111	0.113	0.122	0.117	0.101	0.107
SVJ1	0.132	0.077	0.070	0.069	0.062	0.058	0.089	0.116	0.131	0.127	0.135	0.127	0.110	0.122
All maturities Skew MSM( $\bar{k} = 1$ )	0.108	0.078	0.082	0.079	0.074	0.064	0.085	0.146	0.102	0.115	0.120	0.113	0.100	0.121
Skew MSM( $\bar{k} = 2$ )	0.085	0.068	0.076	0.076	0.072	0.062	0.074	0.110	0.080	0.102	0.111	0.105	0.094	0.102
Skew MSM( $\bar{k} = 3$ )	0.069	0.057	0.063	0.064	0.062	0.054	0.062	0.102	0.073	0.097	0.105	0.101	0.092	0.096
Skew MSM( $\bar{k} = 4$ )	0.058	0.054	0.063	0.065	0.065	0.058	0.060	0.087	0.063	0.092	0.103	0.096	0.088	0.088
SV	0.155	0.090	0.084	0.078	0.070	0.060	0.103	0.167	0.120	0.122	0.123	0.116	0.103	0.132
SVJ0	0.104	0.075	0.075	0.071	0.066	0.057	0.079	0.119	0.084	0.101	0.108	0.102	0.090	0.103
SVJ1	0.112	0.075	0.075	0.071	0.066	0.057	0.082	0.130	0.093	0.105	0.111	0.104	0.093	0.110

Le panel de gauche du tableau ci-dessus illustre le risque quadratique moyen (MSE) dans le modèle SkewMSM et dans les modèles de comparaison SV, SVJ0 et SVJ1, pour les données utilisées. Le MSE le plus faible est obtenu pour  $\bar{k} = 4$ , avec un MSE de 0.060, ce qui marque une amélioration de 25% par rapport au SVJ0, malgré le fait que SkewMSM dépende d'un paramètre en plus par rapport à SVJ0.

Le panel de droite montre quant à lui le MSE pour les prévisions des différents modèles. On remarque là encore que SkewMSM est plus performant que tous les autres modèles, avec une diminution de près de 15% du MSE par rapport au deuxième modèle le plus performant. De plus, les prévisions de SkewMSM sont plus précises pour des options à moyen ou long terme, là où ses performances sont plus contrastées pour des options à courte maturité. Des améliorations du modèle peuvent donc probablement être réalisées pour mieux prédire le comportement de ce type d'options.

#### 6.5.4 • VOLATILITY SMILES

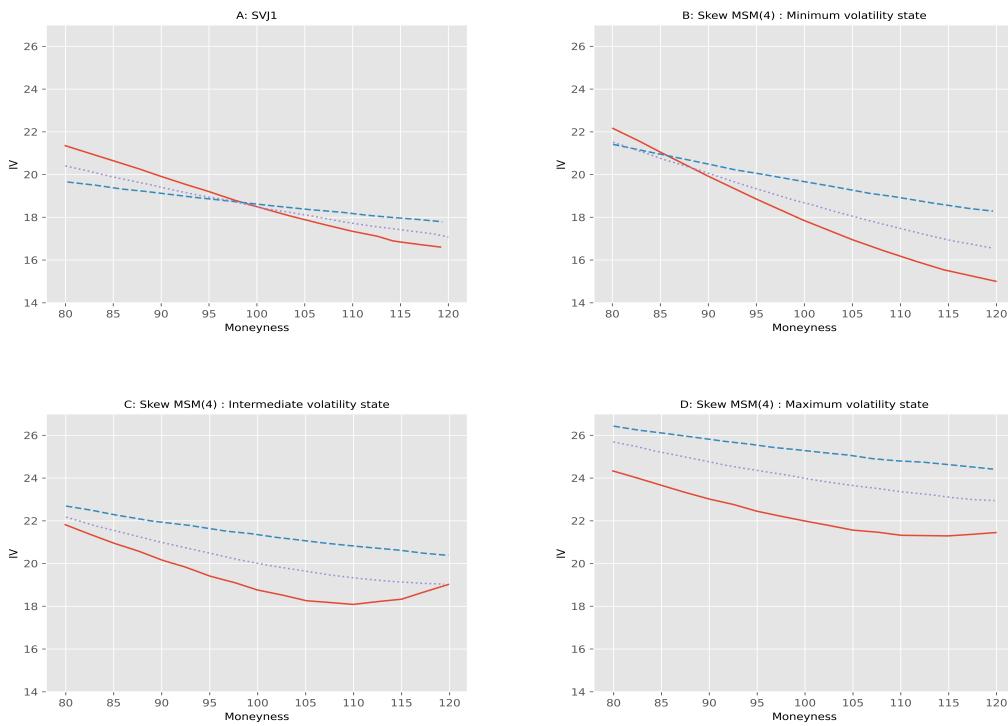


FIGURE 10 – Surface de volatilité implicite. Les graphiques montrent les surfaces de volatilité implicite pour le modèle Skew MSM(4) (panneaux B-D) et SVJ1 (panneau A). Pour tous les graphiques, nous fixons la volatilité spot  $V_t$  à un niveau fixe. Nous utilisons les valeurs de paramètres estimées dans le Tableau 2. Nous traçons les courbes de volatilité implicite pour des maturités de trois mois (ligne solide), six mois (ligne en pointillé), un an (ligne point).

Pour mieux comprendre les performances et la flexibilité de Skew MSM, nous illustrons dans la Figure 11 les Volatility smiles générés par Skew MSM(4) et SVJ1 sous les paramètres du tableau 6. La variance spot  $V_t$  est fixée à un niveau constant. Nous traçons les Volatility smiles pour Skew MSM(4) est dans l'état de volatilité maximale  $M_t = (m_0, m_0, m_0, m_0)$  (Panel B), un état de volatilité intermédiaire  $M_t = (2 - m_0, m_0, m_0, m_0)$  (Panel C) et l'état de volatilité minimale  $M_t = (2 - m_0, 2 - m_0, 2 - m_0, 2 - m_0)$  (Panel D). Respectivement. Dans le Panel A, nous traçons les Volatility smiles produits par SVJ1. Pour  $V_t$  spot constant, Skew MSM peut générer des niveaux de volatilité implicite à différentes maturités en fonction de l'état de  $M_t$ . En revanche, SVJ1 ne peut générer que des Volatility smiles très similaires pour une  $V_t$  spot donnée.

## 7 CONCLUSION

Notre Projet Scientifique Collectif a représenté une opportunité considérable pour notre groupe de mener une étude collective sur un sujet complexe et inédit à nos yeux : la modélisation de la volatilité. Nous avons approché la question selon diverses méthodes et en utilisant différents modèles, confrontés à des impasses et abandonnant certaines approches, mais finalement, nous avons réussi à les apprêhender et à concevoir des programmes informatiques capables de simuler ces modèles et de les comparer adéquatement.

Dans un premier temps, nous avons rédigé ces programmes de manière intuitive avant de les soumettre à des tests et de les améliorer empiriquement. Le modèle MSM s'est avéré être un choix judicieux pour représenter les phénomènes régissant les lois du marché, grâce à sa capacité à prédire l'avenir sur un grand nombre de périodes et à saisir les phénomènes à basse fréquence, tels que les variations sur plusieurs années. Bien qu'il ne soit pas parfait, ce modèle offre une approximation plus que satisfaisante d'une réalité complexe et imprévisible.

En second lieu, le modèle Skew MSM, bien que plus complexe, n'a pas été un exercice de facilité et nous aurions souhaité l'approfondir davantage si le temps nous l'avait permis. Notre tâche était de reprendre les conclusions de l'article de M. Calvet. Nous avons également constaté que la recherche sur le comportement de la volatilité reste un domaine en constante évolution et qu'il y a encore de nombreux travaux à réaliser.

Finalement, nous sommes fiers d'avoir eu la chance d'être encadrés par l'un des pionniers de la recherche scientifique dans ce domaine qui nous tient à cœur, et nous espérons que notre travail pourra éclairer les futures recherches dans ce domaine.

## 8 ANNEXES

---

### 8.1 DONNÉES ET CODES INFORMATIQUES :

---

L'intégralité des codes écrits au cours de notre PSC est à l'adresse : Lien Dropbox  
Les données utilisées pour générer tous les graphes sont incluses dans ce document.

### 8.2 PISTES D'APPROFONDISSEMENT

---

#### 8.2.1 • OPTIMISATIONS INFORMATIQUES POUR SKEW MSM

Comme mentionné précédemment, notre implémentation naïve de Skew MSM a entraîné des temps de calcul très insatisfaisants. Cela était dû au fait que nous avons initialement privilégié un code facilement compréhensible plutôt que des performances optimales. Nous sommes conscients que l'optimisation de la gestion des boucles pour le calcul du RMSE pourrait considérablement améliorer les temps de calcul.

En outre, notre approche pour tracer les Volatility Smiles est encore limitée car nous n'avons pas encore pu obtenir une expression dite "closed-form" pour la volatilité implicite. Nous avons donc dû travailler avec des outils de calcul numérique pour pallier ce problème. Nous prévoyons d'explorer d'autres approches pour améliorer la précision de notre estimation de la volatilité implicite, telles que l'utilisation d'algorithmes d'optimisation ou de techniques d'interpolation plus avancées.

#### 8.2.2 • GARCH

Une approche consiste à modifier la spécification de la fonction d'erreur pour mieux capturer les caractéristiques des données financières. Par exemple, certains chercheurs ont proposé d'utiliser des fonctions d'erreur asymétriques, telles que la fonction d'erreur de Student-t asymétrique généralisée (AST-GARCH), pour mieux modéliser les queues lourdes et les sauts dans les rendements financiers.

Une autre approche est d'utiliser des modèles hybrides qui combinent le GARCH avec d'autres modèles tels que les réseaux de neurones ou les processus de saut. Les modèles hybrides ont été montrés pour améliorer considérablement la performance des modèles GARCH dans certaines situations.

Enfin, l'utilisation de données supplémentaires telles que les données de volume, les données de sentiment des médias sociaux et les données macroéconomiques peuvent aider à améliorer la performance des modèles GARCH. L'incorporation de ces informations supplémentaires peut aider à mieux modéliser les changements de volatilité et à prédire les rendements financiers avec une meilleure précision.

En somme, le modèle GARCH continue d'être un domaine de recherche actif en économétrie financière. Les améliorations apportées à la spécification de la fonction d'erreur, l'utilisation de modèles hybrides et l'incorporation de données supplémentaires sont des pistes prometteuses pour améliorer la performance du modèle GARCH.



### 8.2.3 • RNN

Bien que les RNN puissent avoir des performances inférieures à d'autres modèles dans certains cas, il existe de nombreuses pistes d'approfondissement pour améliorer leur performance.

Tout d'abord, une approche consiste à utiliser des architectures plus avancées telles que les LSTM (Long Short-Term Memory) ou les GRU (Gated Recurrent Unit) qui ont été développées pour résoudre le problème de la disparition du gradient dans les RNN. Ces architectures ont été très efficaces dans de nombreux problèmes de traitement de séquences.

Une autre approche est de combiner les RNN avec d'autres modèles tels que les réseaux de neurones convolutifs (CNN) ou les réseaux de neurones transformer (Transformer), pour exploiter les avantages de chaque modèle dans le traitement de séquences.

Enfin, les techniques d'apprentissage en profondeur telles que le transfert de connaissances, le renforcement et l'apprentissage non supervisé peuvent également aider à améliorer les performances des RNN.

En somme, les RNN ont connu un progrès immense ces dernières années et continuent d'être une piste de recherche active dans de nombreux domaines tels que la reconnaissance de la parole, la traduction automatique et la génération de texte. Les améliorations constantes apportées aux architectures des RNN et aux techniques d'apprentissage en profondeur devraient permettre d'obtenir de meilleures performances et de relever de nouveaux défis dans l'avenir, en particulier le domaine du forecasting.

## RÉFÉRENCES

- Bates, D. S. (2012). U.s. stock market crash risk, 1926–2010. *Journal of Financial Economics*, 105(2) :229–259.
- Calvet, L. and Fisher, A. (2001). Forecasting multifractal volatility. *Journal of Econometrics*, 105(1) :27–58. Forecasting and empirical methods in finance and macroeconomics.
- Calvet, L. E., Fearnley, M., Fisher, A. J., and Leippold, M. (2015). What is beneath the surface ? option pricing with multifrequency latent states. *Journal of Econometrics*, 187(2) :498–511. Econometric Analysis of Financial Derivatives.
- Calvet, L. E. and Fisher, A. (2008). *Multifractal Volatility : Theory, forecasting, and pricing*. Academic Press.
- Clewlow, L. and Strickland, C. (2008). *Implementing derivatives models*. Wiley.
- Gordon, N., Salmond, D., and Smith, A. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings F Radar and Signal Processing*, 140(2) :107.
- Hull, J. C. (2018). *Options, futures, and other derivatives*. Pearson.
- Kou, S. (2007). Chapter 2 jump-diffusion models for asset pricing in financial engineering. In Birge, J. R. and Linetsky, V., editors, *Financial Engineering*, volume 15 of *Handbooks in Operations Research and Management Science*, pages 73–116. Elsevier.
- Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8) :1086–1101.