

Université de Montréal

**Les modèles vectoriels et multiplicatifs avec erreurs
non-négatives de séries chronologiques**

par

Emilie Moutran

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en Discipline

mai 2014

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

**Les modèles vectoriels et multiplicatifs avec erreurs
non-négatives de séries chronologiques**

présenté par

Emilie Moutran

a été évalué par un jury composé des personnes suivantes :

Nom du président du jury

(président-rapporteur)

Pierre Duchesne

(directeur de recherche)

Nom du membre de jury

(membre du jury)

Mémoire accepté le:

Date d'acceptation

SOMMAIRE

L'objectif du présent mémoire vise à présenter des modèles de séries chronologiques multivariés impliquant des vecteurs aléatoires dont chaque composante est non-négative. Nous considérons les modèles vMEM (modèles vectoriels et multiplicatifs avec erreurs non-négatives) présentés par Cipollini, Engle et Gallo (2006) et Cipollini et Gallo (2010). Ces modèles représentent une généralisation au cas multivarié des modèles MEM introduits par Engle (2002). Ces modèles trouvent notamment des applications avec les séries chronologiques financières. Les modèles vMEM permettent de modéliser des séries chronologiques impliquant des volumes d'actif, des durées, des variances conditionnelles, pour ne citer que ces applications. Il est également possible de faire une modélisation conjointe et d'étudier les dynamiques présentes entre les séries chronologiques formant le système étudié.

Afin de modéliser des séries chronologiques multivariées à composantes non-négatives, plusieurs spécifications du terme d'erreur vectoriel ont été proposées dans la littérature. Une première approche consiste à considérer l'utilisation de vecteurs aléatoires dont la distribution du terme d'erreur est telle que chaque composante est non-négative. Cependant, trouver une distribution multivariée suffisamment souple définie sur le support positif est plutôt difficile, au moins avec les applications citées précédemment. Comme indiqué par Cipollini, Engle et Gallo (2006), un candidat possible est une distribution gamma multivariée, qui impose cependant des restrictions sévères sur les corrélations contemporaines entre les variables. Compte tenu que les possibilités sont limitées, une approche possible est d'utiliser la théorie des copules. Ainsi, selon cette approche, des distributions marginales (ou marges) peuvent être spécifiées, dont les distributions en cause ont des supports non-négatifs, et une fonction de copule permet de tenir compte de la dépendance entre les composantes. Une technique d'estimation possible est la méthode du

maximum de vraisemblance. Une approche alternative est la méthode des moments généralisés (GMM). Cette dernière méthode présente l'avantage d'être semi-paramétrique dans le sens que contrairement à l'approche imposant une loi multivariée, il n'est pas nécessaire de spécifier une distribution multivariée pour le terme d'erreur.

De manière générale, l'estimation des modèles vMEM est compliquée. Les algorithmes existants doivent tenir compte du grand nombre de paramètres et de la nature élaborée de la fonction de vraisemblance. Dans le cas de l'estimation par la méthode GMM, le système à résoudre nécessite également l'utilisation de solveurs pour systèmes non-linéaires. Dans ce mémoire, beaucoup d'énergies ont été consacrées à l'élaboration de code informatique (dans le langage R) pour estimer les différents paramètres du modèle.

Dans le premier chapitre, nous définissons les processus stationnaires, les processus autorégressifs, les processus autorégressifs conditionnellement hétéroscédastiques (ARCH) et les processus ARCH généralisés (GARCH). Nous présentons aussi les modèles de durées ACD et les modèles MEM. Dans le deuxième chapitre, nous présentons la théorie des copules nécessaire pour notre travail, dans le cadre des modèles vectoriels et multiplicatifs avec erreurs non-négatives vMEM. Nous discutons également des méthodes possibles d'estimation. Dans le troisième chapitre, nous discutons les résultats des simulations pour plusieurs méthodes d'estimation. Dans le dernier chapitre, des applications sur des séries financières sont présentées. Le code R est fourni dans une annexe. Une conclusion complète ce mémoire.

SUMMARY

The objective of this master thesis is to present models for multivariate time series involving random vectors where each component is non-negative. We consider vMEM models (vector multiplicative error models) presented by Cipollini, Engle and Gallo (2006) and Cipollini and Gallo (2010). These models represent a generalization to the multivariate case of MEM models introduced by Engle (2002). These models are applied especially with financial time series. The vMEM models can be used to model time series involving asset volumes, durations, conditional variances, among these applications. It is also possible to model the variables jointly and to study the dynamics between these time series forming the system under study.

To model multivariate time series with non-negative components, several specifications for the term vector error have been proposed in the literature. One approach is to consider the use of random vectors where the distribution of the error term is such that each component is non-negative. However, finding a sufficiently flexible multivariate distribution defined on the positive support is rather difficult, at least with the applications mentioned above. As indicated by Cipollini, Engle and Gallo (2006), a possible candidate is the multivariate gamma distribution, which however imposes severe restrictions on the contemporaneous correlations between variables. Since the possibilities are limited, one possible approach is to use the theory of copulas. Thus, according to this approach, margins can be specified, such that the distributions in question have non-negative supports, and a copula function takes into account the dependency between components. A possible estimation technique is the method of maximum likelihood. An alternative approach is the generalized method of moments (GMM). This latter method has the advantage of being semi-parametric in the sense that unlike the approach imposing a multivariate

distribution, it is not necessary to specify a multivariate distribution for the error term.

Generally, estimating vMEM models is complicated. Existing algorithms must take into account the large number of parameters and the elaborate nature of the likelihood function. In the case of the GMM estimation method, solving the system requires the use of solvers for non linear systems. In this project, considerable energy has been devoted to the development of computer code (in the R language) to estimate the different parameters of the model.

In the first chapter, we define the stationary processes, autoregressive processes, the autoregressive conditional heteroskedasticity processes (ARCH) and the generalized ARCH processes (GARCH). We also present duration models ACD and MEM models. In the second chapter, we present the theory of copulas needed for our work, under the vector multiplicative error models vMEM. We also discuss the possible estimation methods. In the third chapter, we discuss the simulation results for several estimation methods. In the last chapter, applications to financial time series are presented. The R code is provided in an Appendix. A conclusion completes this thesis.

TABLE DES MATIÈRES

Sommaire	iii
Summary	v
Liste des figures	ix
Liste des tableaux	x
Remerciements	1
Introduction	2
Chapitre 1. Préliminaires	7
1.1. Séries temporelles	7
1.2. Notion de stationnarité	8
1.3. Modèles ARCH univariés	11
1.3.1. Propriété de stationnarité	12
1.3.2. Modèles ARCH	14
1.4. Modèles GARCH	15
1.4.1. Propriété de stationnarité	16
1.4.2. GARCH multivariés	17
1.5. Modèles ACD	19
1.5.1. Motivation	19
1.5.2. Modèles ACD	19
1.6. Modèles MEM	23
1.6.1. Spécification de ϵ_t	25
1.6.2. Spécification de μ_t	26
1.6.3. Estimation et inférence du modèle MEM	29
Chapitre 2. Modèles vMEM	32

2.1.	Distribution gamma multivariée	32
2.2.	Propriétés fondamentales des copules	34
2.2.1.	Définition de la fonction de répartition conjointe et des lois marginales	36
2.2.2.	Simuler une copule normale	39
2.3.	Présentation du modèle vMEM	40
2.3.1.	Exemples d'utilisation de l'approche vMEM	41
2.3.2.	Définition et formulation	43
2.3.3.	Spécification de μ_t	44
2.3.4.	Spécification de ϵ_t	46
2.3.4.1.	Copule normale de distributions marginales gamma	46
2.4.	Estimation du modèle vMEM	47
2.4.1.	Maximum de vraisemblance	48
2.4.1.1.	Les dérivées de la log vraisemblance concentrée	50
2.4.2.	Méthode des moments généralisés (GMM)	52
2.4.3.	Méthode des erreurs standard associées aux paramètres (WNLS)	
	55	
Chapitre 3.	Simulations	58
3.1.	Déscription de la simulation	58
3.1.1.	Problèmes numériques	62
3.2.	Présentations des résultats	63
Chapitre 4.	Analyses de données réelles	72
4.1.	Analyse et modélisation de chaque série	72
Conclusion	82
Bibliographie	84
Annexe A.	A-i
Annexe B.	B-i

LISTE DES FIGURES

- B.1 Graphiques représentant les séries des trois variables relatives à S&P500 : Le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas..... B-ii
- B.2 Graphiques représentant les séries des trois variables relatives à S&P400 : Le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas.....B-iii
- B.3 Graphiques représentant les séries des trois variables relatives à FTSE100 : Le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas.....B-iv

LISTE DES TABLEAUX

3.1	Nombre de cas sur 1000 où il y a eu convergence pour les sept modèles de 1000 réalisations par rapport au nombre d'observations	63
3.2	Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 1 défini par (3.1.1) pour $T = 1000$ et 5000	65
3.3	Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 2 défini par (3.1.2) pour $T = 1000$ et 5000	66
3.4	Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 3 défini par (3.1.3) pour $T = 1000$ et 5000	67
3.5	Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 4 défini par (3.1.4) pour $T = 1000$ et 5000	68
3.6	Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 5 défini par (3.1.5) pour $T = 1000$ et 5000	69
3.7	Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 6 défini par (3.1.6) pour $T = 1000$ et 5000	70
3.8	Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 7 défini par (3.1.7) pour $T = 1000$ et 5000	71
4.1	Estimation des paramètres du modèle $vMEM(1,1)$ en utilisant la méthode semi-paramétrique GMM.	74

4.2	Estimation des paramètres du modèle vMEM(1,1) en utilisant la méthode paramétrique maximum de vraisemblance.	75
4.3	Test de Ljung-Box multivarié des indices boursiers étudiés par la méthode semi-paramétrique GMM.	76
4.4	Test de Ljung-Box multivarié des indices boursiers étudiés par la méthode paramétrique maximum de vraisemblance.	77
4.5	Estimation des paramètres du modèle vMEM(2,1) en utilisant la méthode semi-paramétrique GMM et la méthode paramétrique maximum de vraisemblance de l'indice boursier S&P500.	78
4.6	Estimation des paramètres du modèle vMEM(2,1) en utilisant la méthode semi-paramétrique GMM et la méthode paramétrique maximum de vraisemblance de l'indice boursier S&P400.	79
4.7	Estimation des paramètres du modèle vMEM(2,1) en utilisant la méthode semi-paramétrique GMM et la méthode paramétrique maximum de vraisemblance de l'indice boursier FTSE100.	80
4.8	Test de Ljung-Box multivarié du modèle vMEM(2,1) des indices boursiers étudiés par la méthode semi-paramétrique.	80
4.9	Test de Ljung-Box multivarié du modèle vMEM(2,1) des indices boursiers étudiés par la méthode paramétrique.	81

REMERCIEMENTS

Je tiens à remercier de tout cœur mon directeur de recherche, Monsieur Pierre Duchesne, pour l'orientation, les précieux conseils, le soutien financier, les encouragements, les motivations, la patience et la confiance durant ce travail.

Je tiens à remercier ma famille, spécialement mon mari Roy pour son soutien moral, son encouragement continu et tout ce qu'il a fait pour moi.

Je tiens à remercier également tous nos enseignants durant les années des études, aussi Guillaume Provencher pour le soutien informatique qu'il a présenté durant ce projet.

Je tiens à remercier toute personne qui a participé de près ou de loin à l'exécution de ce travail.

Enfin, je tiens à remercier également les membres du jury pour l'intérêt qu'ils ont accordé à mon travail.

INTRODUCTION

L'étude du comportement des marchés financiers est souvent basée sur l'analyse de la dynamique des processus dont les valeurs sont non-négatives. Des exemples sont les volumes totaux de transactions au cours d'une période donnée, les volatilités réalisées, la différence entre le prix le plus haut et le prix le plus bas observés au cours d'un intervalle de temps (en anglais *range*), les rendements absolus, les durées financières et le nombre de transactions. Voir par exemple l'ouvrage de Bauwens et Giot (2001).

Les classes de modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) ou leurs généralisations (GARCH) sont souvent utilisées afin de modéliser des variances conditionnelles. Ces variances conditionnelles sont en fait des mesures de volatilité dans la littérature financière. Voir Tsay (2010) pour les applications des séries chronologiques en finance. L'ouvrage de Francq et Zakoïan (2009) contient une étude approfondie des modèles GARCH. Les modèles ARCH et GARCH ont été généralisés afin de modéliser des durées. Ces modèles sont utiles car ils permettent de modéliser le temps entre les transactions boursières, qui est en général une variable aléatoire lorsque les transactions sont enregistrées en temps réel. Voir Engle et Russell (1998) qui étudient les modèles de durées conditionnellement autorégressives (ACD). Les caractéristiques communes des modèles ARCH, GARCH et ACD ont suggéré la classe plus large des modèles multiplicatifs avec erreurs non-négatives (MEM). Ces modèles ont été introduits par Engle (2002). Ils sont utiles pour la modélisation de séries financières non-négatives caractérisées par des phénomènes de regroupement (en anglais *clustering*) et de persistance.

Le processus décrivant la dynamique des variables peut être spécifié comme le produit d'un facteur d'échelle conditionnellement déterministe qui évolue selon une équation de type GARCH et un terme d'erreur jouant le rôle de l'innovation qui est souvent supposé un bruit blanc fort constitué de variables

aléatoires indépendamment distribuées et de moyennes unitaires. D'autres modèles ont été proposés dans la littérature afin de modéliser les phénomènes décrits précédemment. On peut consulter Manganelli (2002), Chou (2005) et Engle et Gallo (2006).

Les modèles MEM permettent de modéliser une série chronologique univariée, ce qui veut dire qu'ils sont utiles lorsque l'on dispose d'une réalisation finie d'une variable dans le temps. Cependant, dans les applications, plusieurs variables sont observées dans le temps et il peut être pertinent de procéder à une modélisation simultanée. Parmi les applications évidentes, l'étude des relations de causalité entre les variables peut s'avérer d'intérêt. De plus, si certaines variables sont reliées, elles peuvent être utilisées afin de proposer des prévisions plus précises. De manière générale, la modélisation de plusieurs séries chronologiques univariées amène l'étude des séries chronologiques multivariées. Afin de simplifier l'estimation, il peut être tentant de procéder à l'estimation équation par équation, chaque équation étant estimée séparément. L'estimation équation par équation proposée par Manganelli (2002) est simple et donne des estimateurs convergents sous l'hypothèse de l'indépendance des termes d'innovations. Cependant, les résultats empiriques récents suggèrent une interdépendance dynamique entre plusieurs indicateurs. Ceci implique qu'une approche vectorielle peut fournir de meilleures inférences statistiques et des interprétations plus fiables de leurs dynamiques. Voir Cipollini, Engle et Gallo (2012).

Les modèles vectoriels et multiplicatifs avec erreurs non-négatives (vMEM) permettent de modéliser des vecteurs aléatoires avec composantes qui sont toutes non-négatives. Ils ont été introduits par Engle (2002) et développés par Engle et Gallo (2006). Ces modèles représentent une généralisation au cas multivarié des modèles MEM introduits par Engle (2002). Pour ces modèles, les variables d'intérêt sont représentées comme le produit élément par élément d'un vecteur d'espérances conditionnelles par un vecteur de moyenne unitaire des innovations. Le vecteur aléatoire jouant le rôle de l'innovation admet une structure de covariance générale. Cette particularité permet de modéliser la dépendance entre les séries chronologiques.

Deux approches ont été proposées afin d'estimer les modèles vMEM. Une première approche, l'approche paramétrique complète, consiste à compléter le modèle avec la spécification de la fonction de densité conjointe pour le vecteur

des termes d'erreurs. Cependant, trouver une distribution multivariée suffisamment souple définie sur le support positif est une tâche plutôt difficile, du moins pour les applications issues de l'économétrie et de la finance. Comme indiqué par Cipollini, Engle et Gallo (2006), une possibilité est une distribution gamma multivariée. Toutefois, elle impose des restrictions sévères sur les corrélations contemporaines entre les variables. Compte tenu de ces restrictions, une autre approche a été proposée dans la littérature. En effet, puisque les distributions multivariées existantes avec composantes non-négatives ne sont pas suffisamment flexibles, l'utilisation de la théorie des copules semble tout à fait indiquée. Ainsi, selon cette approche, les distributions marginales peuvent être spécifiées, dont les supports associés sont non-négatifs, et une fonction de copule permet de tenir compte de la dépendance entre les composantes. L'estimation peut se faire en utilisant la théorie du maximum de vraisemblance. Les copules elliptiques (par exemple, la copule normale ou la copule de Student) sont attrayantes parce qu'elles peuvent être utilisées également pour un nombre d'observations modérément grand. De même, elles peuvent accommoder la dépendance dans la queue de la distribution. Cependant, elles ont un comportement elliptiquement symétrique qui peut représenter une limite dans certains contextes. Voir McNeil, Frey et Embrechts (2005). Au contraire, certaines copules archimédiennes peuvent accommoder les profils de dépendance asymétriques mais semblent moins utilisables lorsque la dimension tend à augmenter (voir McNeil, Frey et Embrechts (2005)). Il faut cependant noter que les copules ne sont pas toujours capables de modéliser de manière adéquate l'association entre les composantes du terme d'erreur. Quelques détails sur certaines spécifications possibles de copules liées à la famille de copules elliptiques sont discutés dans Cipollini, Engle et Gallo (2012).

La spécification complète de la distribution de terme d'erreur peut être moins intéressante si l'objectif principal de l'analyse porte sur la dynamique de la moyenne conditionnelle. Ainsi, une approche alternative semi-paramétrique peut s'avérer prometteuse. La méthode des moments généralisés, notée par GMM, est dans cette catégorie. Cette méthode présente l'avantage d'être semi-paramétrique dans le sens que contrairement à l'approche imposant une loi multivariée, il n'est pas nécessaire de spécifier une distribution multivariée pour le terme d'erreur. Voir Cipollini, Engle et Gallo (2012).

Dans ce mémoire, après avoir introduit ces modèles, nous voulons comparer les deux approches dominantes, paramétrique et semi-paramétrique, des

modèles vectoriels et multiplicatifs avec erreurs non-négatives. Nous appliquons trois méthodes d'estimation. Pour l'approche paramétrique complète, nous considérons la méthode du maximum de vraisemblance (plus précisément la méthode du maximum de vraisemblance concentrée), tandis que pour l'approche semi-paramétrique nous présentons la méthode des moments généralisés (GMM) et la méthode des erreurs standard associées aux paramètres (WNLS). Nous simulons des modèles vectoriels et multiplicatifs à erreurs non-négatives et nous appliquons les méthodes d'estimation indiquées ci-dessus. Dans nos expériences de simulation, plusieurs modèles sont considérés, pour plusieurs tailles d'échantillons. Le but est de voir si le nombre d'observations a une influence sur le biais des estimateurs et ainsi de comparer les deux approches, paramétrique et semi-paramétrique, au niveau du biais, de l'erreur quadratique moyenne, de la variance échantillonnale et de la moyenne des estimateurs de variance. Nous illustrons l'application des modèles vMEM à ces données et nous analysons les résultats de simulations. Finalement, nous effectuons une analyse avec des données financières portant sur trois indices boursiers S&P500, S&P400 et FTSE100. Pour chaque indice boursier, nous considérons une série multivariée formée par trois variables : le rendement absolu, la volatilité réalisée et la différence entre le prix le plus élevé et le prix le plus bas. Nous modélisons dans un premier temps trois séries multivariées pour les trois indices boursiers à l'aide des modèles vectoriels et multiplicatifs avec erreurs non-négatives et nous calculons les statistiques de Ljung-Box multivariées pour étudier l'indépendance des résidus estimés. Puis dans un deuxième temps, nous exposons les paramètres estimés par les méthodes paramétriques et semi-paramétriques des modèles vectoriels et multiplicatifs avec erreurs non-négatives mais avec des ordres de modèles différents de la première analyse, ainsi que l'application du test Ljung-Box multivarié sur les erreurs estimées de la nouvelle analyse. Comme pour les modèles autorégressifs moyennes mobiles ARMA, des ordres de modèles doivent être considérés. Pour les modèles vectoriels et multiplicatifs avec erreurs non-négatives, nous définissons les modèles vMEM(1,1) et vMEM(2,1). Avec les tests diagnostics mentionnés précédemment, nous constatons qu'une spécification vMEM(2,1) semble supérieure à un modèle vMEM(1,1).

Ce mémoire est constitué de quatre chapitres. Dans le premier chapitre, nous introduisons certaines définitions importantes sur les séries chronologiques. Les notions sur la stationnarité ainsi que les modèles des processus

autorégressifs conditionnellement hétéroscédastiques (ARCH) et des processus ARCH généralisés (GARCH) sont aussi présentés dans le premier chapitre. Ensuite, nous décrivons les processus ACD avec quelques propriétés ainsi que les modèles multiplicatifs à erreurs non-négatives. Nous présentons les modèles, les conditions de positivité et de stationnarité et les méthodes d'estimation. Au chapitre 2, nous définissons les copules ainsi que la théorie nécessaire pour la spécification des termes d'erreurs. Ensuite, nous présentons les modèles vectoriels et multiplicatifs à erreurs non-négatives. Nous donnons les caractéristiques principales de ces modèles ainsi que quelques exemples sur l'application de ces modèles. Enfin, nous décrivons différentes méthodes d'estimation possibles dont la méthode du maximum de vraisemblance concentrée. Nous utilisons l'algorithme d'optimisation BHHH proposé par Berndt, Hall, Hall et Hausman (1974), la méthode des moments généralisés et la méthode des erreurs standard associées aux paramètres. Au chapitre 3, nous décrivons les différentes étapes de nos simulations afin d'illustrer nos résultats. Les méthodes utilisées pour générer les copules, pour simuler un modèle vectoriel et multiplicatif à erreurs non-négatives et pour estimer les paramètres de différentes modèles sont décrites à cet endroit. Ainsi, les problèmes de programmation rencontrés durant ce travail sont décrits dans cette section. La composante empirique a été programmée en R et le code utilisé est fourni en annexe. Les résultats de simulations sont aussi analysés dans ce chapitre. Finalement, l'application avec des données financières est discutée dans le chapitre 4. Après une brève description sur l'expérience, la taille et la période des données, plusieurs modèles sont ajustés et la qualité de l'ajustement est considérée. Une conclusion termine ce travail.

Chapitre 1

PRÉLIMINAIRES

Dans ce chapitre, nous présentons en premier lieu, la définition des processus stochastiques suivie par la définition d'une série temporelle multivariée, les notions de stationnarité, le processus autorégressif (AR), le processus autorégressif à moyenne mobile (ARMA) et la différence de martingales. En second lieu, nous exposons les modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) et leurs généralisations (GARCH) qui modélisent les variances conditionnelles. Ensuite, nous présentons les modèles de durées conditionnellement autorégressives (ACD) qui modélisent les durées comme le temps entre des transactions boursières. Enfin, nous exposons les modèles multiplicatifs avec erreurs non-négatives (MEM) qui possèdent des caractéristiques communes à celles des modèles ARCH, GARCH et ACD. Cette section présente les définitions essentielles pour la compréhension générale de notre problématique. Plus de détails se trouvent dans Gouriéroux (1992, 1995) et Engle et Russell (1998).

1.1. SÉRIES TEMPORELLES

Un processus stochastique est une suite de variables aléatoires définies sur un même espace Ω , appelé espace des états de la nature. On note le processus stochastique $\mathbf{Y} = \{\mathbf{Y}_t; t \in \mathbb{Z}\}$, où \mathbb{Z} est l'ensemble des entiers relatifs, telle que $\mathbf{Y}_t = (Y_t(1), \dots, Y_t(K))^T$ est de dimension K .

Il est courant dans un contexte de séries chronologiques de supposer que le processus stochastique est défini pour tous t . Voir Shumway et Stoffer (2006, p. 28). Une série temporelle multivariée est une réalisation finie de taille T d'un processus stochastique multivarié $\{\mathbf{Y}_t; t \in \mathbb{Z}\}$, tel que le vecteur aléatoire \mathbf{Y}_t est de dimension K , avec $\mathbf{Y}_t \in \mathbb{R}^K$.

La fonction moyenne est définie par :

$$\mathbf{m}_t = E(\mathbf{Y}_t) = (m_t(1), \dots, m_t(K))^T, t \in \mathbb{Z}.$$

La fonction de covariance entre \mathbf{Y}_s et \mathbf{Y}_t est définie quant à elle par :

$$\Gamma(s, t) = \text{Cov}(\mathbf{Y}_s, \mathbf{Y}_t) = E\{(\mathbf{Y}_s - \mathbf{m}_s)(\mathbf{Y}_t - \mathbf{m}_t)^T\}, s, t \in \mathbb{Z}.$$

La matrice des covariances $\Gamma(s, t)$ est telle que $\Gamma(s, t) = (\gamma_{ij}(s, t))$ avec $i, j \in \{1, \dots, K\}$. Ainsi elle permet de décrire la dépendance entre les observations de différentes séries ($i \neq j$), mais aussi la dépendance entre $Y_s(i)$ et $Y_t(i)$, c'est-à-dire la dépendance pour une même série pour les instants s et t .

Plusieurs processus ont une propriété commune importante. C'est la propriété de stationnarité. Dans la section suivante, nous présentons la définition d'un processus stationnaire.

1.2. NOTION DE STATIONNARITÉ

On distingue deux types de stationnarité, la stationnarité stricte ou forte et la stationnarité au second ordre ou au sens faible. La stationnarité stricte est une propriété portant sur les propriétés d'invariance des distributions. Toutes les définitions fondamentales liées aux concepts de stationnarité se trouvent dans Brockwell et Davis (1991).

Définition 1.2.1. *Le processus $\{\mathbf{Y}_t\}$ est dit strictement ou fortement stationnaire si \forall N -uplet d'indices temporels $t_1 < t_2 < \dots < t_N$, tel que $t_i \in \mathbb{Z}$ et pour tout temps $h \in \mathbb{Z}$ avec $t_{i+h} \in \mathbb{Z}, \forall i, i = 1, \dots, N$, le vecteur aléatoire $(\mathbf{Y}_{t_1+h}^T, \dots, \mathbf{Y}_{t_N+h}^T)^T$ a la même loi de probabilité que celle de $(\mathbf{Y}_{t_1}^T, \dots, \mathbf{Y}_{t_N}^T)^T$.*

La stationnarité au sens strict peut être compliquée à vérifier en pratique, mais utile dans l'élaboration des résultats théoriques. Un concept moins exigeant et qui porte que sur les deux premiers moments est la stationnarité au second ordre.

Définition 1.2.2. *Un processus $\{\mathbf{Y}_t\}$ est dit stationnaire au second ordre, ou stationnaire au sens faible, si les trois conditions suivantes sont satisfaites :*

- $\forall t \in \mathbb{Z}, E(\mathbf{Y}_t^2(i)) < \infty, i \in \{1, \dots, K\}$;

- $\forall t \in \mathbb{Z}, \mathbf{m}_t = \mathbf{m}$, indépendant de t ;
- $\forall (s, t) \in \mathbb{Z}^2, \Gamma(s, t) = \Gamma(t - s)$.

On notera dans la Définition 1.2.2 un léger abus de notation où la fonction Γ est utilisée également dans le cas stationnaire.

Selon la Définition 1.2.2, la moyenne ne dépend pas du temps et les covariances ne dépendent que de la différence entre les indices temporels des deux vecteurs aléatoires.

Un processus stationnaire de base dans l'analyse des séries chronologiques est le bruit blanc. La définition de ce processus est la suivante :

Définition 1.2.3. *Le processus $\{\epsilon_t\}$ est un bruit blanc si :*

- $E(\epsilon_t) = \mathbf{0}$,
- $Var(\epsilon_t) = \Sigma_\epsilon$,
- $Cov(\epsilon_t, \epsilon_s) = \mathbf{0}$; pour $t \neq s$.

Selon la Définition 1.2.3, le bruit blanc est une suite de vecteurs aléatoires centrés en zéro, de même matrice des variances et covariances et non-corrélés entre eux.

Si les vecteurs non-corrélés qui constituent le processus bruit blanc sont indépendants, alors nous sommes en présence d'un bruit blanc *fort*. Lorsque les vecteurs aléatoires sont présumés seulement non-corrélés, alors nous avons un bruit blanc *faible*.

Nous achevons cette section par la présentation du processus autorégressif d'ordre p et nous introduisons les processus possédant la propriété de différence de martingales.

Définition 1.2.4. *Un processus stochastique $\{Y_t\}$ est un processus autorégressif d'ordre p , noté $AR(p)$, s'il satisfait l'équation aux différences stochastique suivante :*

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t, \quad \forall t \in \mathbb{Z}, \quad (1.2.1)$$

où les paramètres autorégressifs $(\phi_1, \phi_2, \dots, \phi_p)$ sont des réels avec ϕ_p non-nul. Le processus $\{\epsilon_t\}$ est un bruit blanc.

Nous remarquons que Y_t dépend des variables explicatives $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$. Donc les variables explicatives sont des valeurs retardées de la variable dépendante Y_t .

Le modèle AR(p) peut être généralisé à la classe des modèles ARMA étudiés en profondeur dans Brockwell et Davis (1991). Cependant, nous n'avons besoin que des modèles autorégressifs, ce qui explique que l'on s'en tient à cette définition.

Le modèle (1.2.1) admet aussi la représentation suivante :

$$\Phi(B)Y_t = \epsilon_t, \quad (1.2.2)$$

où $\Phi(B)$ est le polynôme autorégressif défini par $\Phi(Z) = 1 - \Phi_1 Z - \Phi_2 Z^2 - \dots - \Phi_p Z^p$. La quantité B est appelé l'opérateur retard tel que $B^0 = I$ et $B^p Y_t = Y_{t-p}$, $p \geq 0$.

Le processus défini par (1.2.1) est stationnaire si et seulement si les racines de $\Phi(z) = 0$ sont à l'extérieur du cercle unitaire. Une preuve de ce résultat se trouve dans Brockwell et Davis (1991).

Il existe également une classe de processus très importante qui est un bruit blanc particulier. Il s'agit du processus différence de martingales.

Définition 1.2.5. *Un processus stochastique $\{Y_t, t \in \mathbb{Z}\}$ est une différence de martingales si et seulement si*

$$E(Y_t | \mathcal{F}_{t-1}) = 0,$$

où \mathcal{F}_t désigne l'information passée disponible jusqu'à l'instant t inclusivement. On va noter également $E(Y_t / \underline{Y}_{t-1}) = 0$.

Dans les deux sections qui suivent, nous présentons une classe de modèles qui s'inspirent des modèles AR et ARMA mais qui permettent la modélisation des variances conditionnelles et donc de la volatilité.

1.3. MODÈLES ARCH UNIVARIÉS

Les séries monétaires et financières sont caractérisées par des périodes de fortes perturbations suivies de périodes de faibles perturbations. On parle ici de phénomènes de volatilité survenant par grappes (en anglais *clustering*). Ce phénomène, que nous appelons aussi l'hétéroscédasticité conditionnelle, est fréquent dans les données boursières et les taux de change.

Or, pour ce type de séries, les modèles linéaires standards utilisés souvent comme les processus autorégressifs à moyennes mobiles, sont limités parce qu'ils ne prennent pas en considération les phénomènes de variabilité de la volatilité. Par exemple, pour les modèles ARMA avec bruit blanc fort, l'espérance conditionnelle varie au cours du temps tandis que la variance conditionnelle ne change pas. Or, en finance, la variabilité dans le temps des variances est un fait bien établi. On peut consulter Andersen, Bollerslev, Christoffersen et Diebold (2005) pour des travaux sur la modélisation et la prévision de la volatilité.

Face aux déficiences des représentations ARMA(p,q) pour les problèmes financiers et monétaires, une classe de modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) a été proposée par Engle (1982). Les modèles ARCH sont utilisés couramment dans la modélisation de séries temporelles financières qui comportent des volatilités variables. Les modèles ARCH permettent de décrire de telles situations. Nous trouvons ici la définition du modèle ARCH(1).

Définition 1.3.1. *Le processus $\{Y_t\}$ est un processus ARCH(1) si*

$$Y_t = \epsilon_t \sqrt{(\alpha_0 + \alpha_1 Y_{t-1}^2)},$$

où $\{\epsilon_t\}$ est un bruit blanc gaussien, et donc $\epsilon_t \sim \mathcal{N}(0, 1)$.

On note généralement $h_t = \alpha_0 + \alpha_1 Y_{t-1}^2$, et donc $Y_t = \epsilon_t \sqrt{h_t}$. La quantité h_t désigne une variable positive conditionnelle à l'ensemble d'information des valeurs passées de $Y_s, s \leq t-1$. La première version du modèle spécifiait le processus $\{\epsilon_t\}$ comme étant gaussien. Rapidement, des spécifications alternatives ont été considérées, et la distribution de ϵ_t pourrait être de loi autre que celle de la normale. Voir Gouriéroux (1992).

La spécification ARCH(1) montre que le carré des perturbations suit un processus autorégressif d'ordre un. Ce modèle permet une forme autorégressive pour la variance conditionnelle telle que la variance conditionnelle aux informations passées est fonction du temps.

Nous exposons dans la suite la notion de stationnarité de ce processus.

1.3.1. Propriété de stationnarité

Pour que le processus $\{Y_t\}$ soit stationnaire au second ordre, la variance inconditionnelle de $\{Y_t\}$ doit être constante.

Or,

$$E(Y_t^2) = \alpha_0 + \alpha_1 E(Y_{t-1}^2),$$

c'est-à-dire que l'on doit avoir $\alpha_1 < 1$, et donc comme $E(Y_t) = 0$,

$$\text{Var}(Y_t) = \frac{\alpha_0}{1 - \alpha_1}.$$

Ainsi, une condition nécessaire pour que $\{Y_t\}$ soit stationnaire au second ordre est $0 < \alpha_0 / (1 - \alpha_1) < \infty$. Or h_t doit être non-négative. Ainsi, il faut également que $\alpha_0 > 0$ et $\alpha_1 \geq 0$. Ainsi, l'espace paramétrique doit se restreindre à $\alpha_0 > 0$ et $0 \leq \alpha_1 < 1$.

Le processus ARCH(1) est une différence de martingales car :

$$\begin{aligned} E(Y_t | \underline{Y}_{t-1}) &= E(\epsilon_t \sqrt{h_t} | \underline{Y}_{t-1}), \\ &= \sqrt{h_t} E(\epsilon_t | \underline{Y}_{t-1}), \\ &= \sqrt{h_t} E(\epsilon_t), \\ &= 0. \end{aligned}$$

De plus, pour $t > s$, le même genre de raisonnement permet d'établir :

$$\begin{aligned}
E(Y_t Y_s) &= E\{E(Y_t Y_s \mid \underline{Y}_{t-1})\}, \\
&= E\{Y_s \sqrt{h_t} E(\epsilon_t \mid \underline{Y}_{t-1})\}, \\
&= E\{Y_s \sqrt{h_t} E(\epsilon_t)\}, \\
&= E\{0\}, \\
&= 0.
\end{aligned}$$

Ainsi, on constate que $\{Y_t\}$ est une suite de variables aléatoires non-corrélées.

De plus, la variance conditionnelle du processus $\{Y_t\}$ suit un ARCH(1), définie par :

$$\begin{aligned}
\text{Var}(Y_t \mid \underline{Y}_{t-1}) &= E(\epsilon_t^2 h_t \mid \underline{Y}_{t-1}), \\
&= h_t E(\epsilon_t^2), \\
&= h_t, \\
&= \alpha_0 + \alpha_1 Y_{t-1}^2.
\end{aligned}$$

La variance conditionnelle du terme Y_t dépend du terme Y_t au carré décalée d'une période. On note que la variance conditionnelle du terme Y_t doit être toujours positive, alors il faut que $\alpha_1 > 0$.

Nous pouvons établir les moments conditionnels et inconditionnels d'ordre quatre du processus ARCH(1). En fait, le moment conditionnel centré d'ordre quatre du processus $\{Y_t\}$ est égal à :

$$E(Y_t^4 \mid \underline{Y}_{t-1}) = 3\{\alpha_0 + \alpha_1 Y_{t-1}^2\}^2,$$

et le moment inconditionnel centré d'ordre quatre du processus $\{Y_t\}$ vérifie :

$$E(Y_t^4) = 3 \left\{ \alpha_0^2 + \frac{2\alpha_1 \alpha_0^2}{1 - \alpha_1} + \alpha_1^2 E(Y_{t-1}^4) \right\} = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - 3\alpha_1^2)(1 - \alpha_1)}.$$

La condition d'existence du moment inconditionnel d'ordre quatre est $3\alpha_1^2 < 1$. On en déduit alors l'expression de la kurtosis inconditionnelle associée au processus ARCH(1) :

$$\kappa = \frac{E(Y_t^4)}{E(Y_t^2)^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3.$$

Rappelons qu'un coefficient de kurtosis égal à trois correspond au cas de la loi normale. Les queues de la distribution marginale d'un processus ARCH(1) sont donc plus épaisses (distribution leptokurtique) que pour un processus gaussien.

Ainsi, il ressort de ceci que le processus ARCH(1) est constitué de variables aléatoires non-corrélées qui sont cependant dépendantes. On note que

$$Y_t | \underline{Y}_{t-1} \sim \mathcal{N}(0, \text{Var}(Y_t | \underline{Y}_{t-1})).$$

En particulier, il ressort que $\{Y_t\}$ est un bruit blanc faible qui ne peut pas être gaussien (dans ce dernier cas, la non-corrélation et l'indépendance sont équivalentes).

De la même manière qu'un modèle autorégressif AR(1) n'est pas suffisamment flexible, menant à l'introduction de plusieurs retards additionnels dans les modèles ARCH(1), les modèles ARCH ont été généralisés afin de décrire des mécanismes de la volatilité plus élaborés. Les propriétés présentées ici pour les ARCH(1) peuvent se généraliser au cas des processus ARCH(q).

1.3.2. Modèles ARCH

Les modèles ARCH(q) sont des extensions des modèles ARCH(1). Les modèles ARCH(q) font intervenir plusieurs retards,

$$Y_t = \epsilon_t \sqrt{h_t}, \quad (1.3.1)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2. \quad (1.3.2)$$

Souvent, il est supposé que la loi de ϵ_t est normale et donc que $\epsilon_t \sim \mathcal{N}(0, 1)$. Puisque h_t est une variance conditionnelle, on doit avoir $h_t \geq 0$. Des conditions suffisantes sur l'espace paramétrique sont $\alpha_0 > 0$, $\alpha_i \geq 0$ pour $i = 1, \dots, q$.

Le processus $\{Y_t\}$ est stationnaire si $\sum_{i=1}^q \alpha_i < 1$. Alors, la variance du processus existe et est de la forme :

$$\text{Var}(Y_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i}.$$

On peut alors écrire :

$$\begin{aligned} E(Y_t | \underline{Y}_{t-1}) &= 0, \\ E(Y_t^2 | \underline{Y}_{t-1}) &= \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i}^2 = h_t. \end{aligned}$$

On montre aussi que $\text{Cov}(Y_t, Y_s) = 0$, $t \neq s$. Donc on retrouve les deux propriétés vues précédemment, à savoir la propriété de différence de martingales et la propriété de la variabilité de la variance conditionnelle dans le temps.

De la même façon qu'un modèle ARMA n'est pas suffisamment flexible pour considérer la variabilité de la variance conditionnelle, menant à l'introduction des modèles ARCH, les modèles ARCH ont été étendus afin de modéliser les phénomènes de persistance de chocs mais avec moins de paramètres à estimer. Ces modèles, que nous présentons dans la section suivante, s'appellent les modèles autorégressifs conditionnellement hétéroscédastiques généralisés (GARCH).

1.4. MODÈLES GARCH

Pour de nombreuses applications des séries monétaires et financières, afin de tenir compte de la longue mémoire de la volatilité, l'insertion d'un grand nombre de retards q dans l'équation de la variance conditionnelle du modèle ARCH(q) est nécessaire. Cependant, ce nombre important de paramètres peut compliquer l'estimation et être associée à une perte d'efficacité. Compte tenu de ces considérations, une extension importante a été suggérée par Bollerslev (1986), le modèle autorégressif conditionnellement hétéroscédastique généralisé (GARCH). L'extension des modèles ARCH aux modèles GARCH ressemble à l'extension des modèles AR aux modèles ARMA. Ainsi, une composante de type moyenne mobile est ajoutée. Plus précisément, la variance conditionnelle de la variance étudiée est déterminée par le carré des q termes d'erreurs passés et des p variances conditionnelles retardées.

Dans cette section, nous exposons la définition des modèles GARCH et GARCH multivariés, les moments conditionnels et les propriétés de stationnarité.

Définition 1.4.1. *Le processus $\{Y_t\}$ est GARCH (q,p) si :*

$$Y_t = \epsilon_t \sqrt{h_t}, \quad (1.4.1)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}. \quad (1.4.2)$$

Comme pour les modèles ARCH, une hypothèse souvent faite au niveau de la distribution du terme ϵ_t est qu'il est de la loi normale, de sorte que $\epsilon_t \sim \mathcal{N}(0, 1)$. Des conditions suffisantes pour assurer que $h_t \geq 0$ sont $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $i = 1, \dots, q$; $j = 1, \dots, p$.

L'espérance et la variance conditionnelle des modèles GARCH (q,p) sont :

$$E(Y_t | \underline{Y}_{t-1}) = 0, \quad (1.4.3)$$

$$\text{Var}(Y_t | \underline{Y}_{t-1}) = E(Y_t^2 | \underline{Y}_{t-1}) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}. \quad (1.4.4)$$

Le processus $\{Y_t\}$ satisfaisant le modèle GARCH (q,p) est une différence de martingales et admet des composantes non-corrélés de moyenne nulle.

Dans la prochaine section, nous étudions les conditions de stationnarité des processus admettant une structure GARCH (q,p) .

1.4.1. Propriété de stationnarité

Franck et Zakoïan (2009) donnent les conditions nécessaires et suffisantes de la stationnarité au second ordre des modèles GARCH. En fait, le processus GARCH (q,p) est stationnaire au second ordre si la condition suivante est satisfaite :

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1,$$

où il est implicitement admis que $\alpha_0 > 0$ et dans ce cas, la variance inconditionnelle de Y_t s'exprime sous la forme suivante :

$$V(Y_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}.$$

Considérons maintenant, un cas particulier du processus GARCH(q,p), le processus GARCH(1,1) tel que :

$$Y_t = \epsilon_t \sqrt{h_t}, \quad (1.4.5)$$

$$h_t = \alpha_0 + \alpha_1 Y_{t-1}^2 + \beta_1 h_{t-1}, \quad (1.4.6)$$

où $\{\epsilon_t\}$ est un processus gaussien. Sous la condition de stationnarité de second ordre $\alpha_1 + \beta_1 < 1$, la variance inconditionnelle du processus $\{Y_t\}$ est définie et constante dans le temps. La variance inconditionnelle s'exprime sous la forme suivante :

$$V(Y_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Bollerslev (1986) a montré que l'existence du moment d'ordre quatre exigeait $(\alpha_1 + \beta_1)^2 + 2\alpha_1^2 < 1$ et que l'expression du coefficient de kurtosis s'exprimait comme suit :

$$\kappa_u = \frac{E(Y_t^4)}{E(Y_t^2)^2} = \frac{3(1 - (\alpha_1 + \beta_1)^2)}{1 - (\alpha_1 + \beta_1)^2 \alpha_1^2}.$$

Le kurtosis est toujours supérieur à trois et donc il peut aussi être utile sur des données à queues de distribution plus épaisses que celle de la loi normale.

Nous achevons cette section par une présentation des modèles GARCH multivariés.

1.4.2. GARCH multivariés

Le processus $\{Y_t\}$ est un GARCH multivarié, à K composantes, $Y_t = (Y_{1t}, \dots, Y_{Kt})$, s'il s'écrit comme $Y_t = \epsilon_t H_t^{1/2}$. Ici, la notation $H_t^{1/2}$ veut dire la racine carrée de la matrice H_t , obtenue à l'aide d'une décomposition spectrale de H_t , présumée définie positive. La variance conditionnelle $H_t = \text{Var}(Y_t | Y_{t-1})$ est une matrice $K \times K$ qui a pour élément h_{klt} . Le processus d'erreur $\{\epsilon_t\}$ est tel que les vecteurs aléatoires ϵ_t sont indépendants et identiquement distribués de moyenne 0 et de matrice de covariance la matrice identité I .

L'espérance et la variance conditionnelle des modèles GARCH multivariés

sont :

$$E(\mathbf{Y}_t | \underline{\mathbf{Y}}_{t-1}) = \mathbf{0}, \quad (1.4.7)$$

$$\text{Var}(\mathbf{Y}_t | \underline{\mathbf{Y}}_{t-1}) = \mathbf{H}_t, \quad (1.4.8)$$

où

$$h_{klt} = C_{kl} + \sum_{i=1}^q \left(\sum_{k',l'} \alpha_{k'lk'l'i} Y_{k',t-i} Y_{l',t-i} \right) + \sum_{j=1}^p \left(\sum_{k',l'} \beta_{k'lk'l'j} h_{k',l',t-j} \right). \quad (1.4.9)$$

Les matrices \mathbf{H}_{t-i} , \mathbf{Y}_{t-i} , \mathbf{Y}'_{t-i} sont symétriques telles que :

$$C_{kl} = C_{lk},$$

$$\alpha_{k'lk'l'i} = \alpha_{l'kk'l'i},$$

$$\alpha_{k'lk'l'i} = \alpha_{k'l'l'k'i},$$

$$\beta_{k'lk'l'j} = \beta_{l'kk'l'j},$$

$$\beta_{k'lk'l'j} = \beta_{k'l'l'k'j}.$$

La forme matricielle de la variance conditionnelle est :

$$\text{Vech}(\mathbf{H}_t) = \text{Vech}(\mathbf{C}) + \sum_{i=1}^q \mathbf{A}_i \text{Vech}(\mathbf{Y}_{t-i} \mathbf{Y}'_{t-i}) + \sum_{j=1}^p \mathbf{B}_j \text{Vech}(\mathbf{H}_{t-j}). \quad (1.4.10)$$

Dans l'expression (1.4.10), $\text{Vech}(\mathbf{H})$ désigne l'opérateur consistant à présenter sous forme vectorielle la portion inférieure de la matrice symétrique \mathbf{H} . Les matrices de paramètres \mathbf{A}_i et \mathbf{B}_j contiennent les paramètres α et β .

Les conditions suffisantes pour la positivité sont que les matrices \mathbf{C} , \mathbf{A}_i ; $i = 1, \dots, q$ et \mathbf{B}_j ; $j = 1, \dots, p$ soient semi-définies positives. Voir Gouriéroux (1992).

Les modèles ARCH et GARCH ont été généralisés afin de proposer des modèles permettant de décrire le temps entre des transactions boursières. Dans la section suivante, nous présentons les modèles de durées conditionnellement autorégressives (ACD) et ses propriétés fondamentales.

1.5. MODÈLES ACD

1.5.1. Motivation

Le développement rapide des systèmes électroniques et l'informatisation ont facilité la création et la diffusion des données à hautes fréquences dans les marchés financiers. Les chercheurs ont pu recueillir des données sur des intervalles de temps réguliers. Dans ce contexte, les données sont habituellement disponibles à des intervalles de plus en plus fins. Avec les modèles ARMA, les applications spectaculaires modélisaient des données annuelles ou trimestrielles, et souvent le temps le plus court entre les données était au plus en termes de mois. Avec les modèles ARCH et GARCH, il n'est pas rare de voir des applications où les données sont observées sur une échelle journalière ou même intra-journalière. Cependant, les données sont toujours régulièrement espacées. En fait, traditionnellement, les séries chronologiques ont été analysées pour des variables régulièrement espacées. Cependant, Diamond et Verrecchia (1987) ont pris en considération le temps comme une source d'information, en supposant que l'absence de transaction signifie une mauvaise nouvelle. Easley et O'Hara (1992) ont considéré que le temps dépend du processus de prix car l'absence de transaction suggère l'absence d'arrivée d'informations nouvelles sur le marché. D'où l'importance d'étudier le facteur temps.

En pratique, les données de transaction arrivent à des intervalles de temps irréguliers. Ainsi, le temps entre deux transactions boursières est en fait une variable aléatoire. Ainsi, la prochaine transaction pourrait survenir après quelques secondes ou quelques minutes, à titre d'exemple. Donc, les techniques économétriques standards qui présument que l'intervalle de temps est constant, ne sont pas strictement valides. Il ressort de ceci qu'il n'est pas possible d'utiliser les modèles ARMA ou GARCH sur des données irrégulièrement espacées. Engle et Russel (1998) ont introduit le modèle de durées conditionnellement autorégressives pour analyser des données irrégulièrement espacées.

Par la suite, nous exposons les modèles ACD et ses propriétés.

1.5.2. Modèles ACD

Soit $x_i = t_i - t_{i-1}$, la durée qui correspond au temps entre deux transactions consécutives i et $i - 1$. Soit ψ_i l'espérance conditionnelle de la i ème durée sachant son passé, telle que $\psi_i = E(x_i | x_{i-1}, \dots, x_1) = \psi_i(x_{i-1}, \dots, x_1; \theta)$. La

définition suivante donne plus précisément ce qu'est un modèle ACD.

Définition 1.5.1. *Le modèle ACD est défini de la façon suivante :*

$$x_i = \psi_i \epsilon_i.$$

La variable aléatoire ψ_i est donc l'espérance conditionnelle de la durée sachant son passé. Les variables aléatoires ϵ_i sont des variables aléatoires positives, indépendantes et identiquement distribuées avec pour densité $p(\epsilon)$ de sorte que $E(\epsilon_i) = 1$, $\text{Var}(\epsilon_i) = \sigma^2$ et $E(x_i | \underline{x_{i-1}}) = \psi_i$.

Un peu à la manière des modèles ARMA, il est possible de considérer un modèle ACD avec ordres m et q , noté ACD(m, q), qui repose sur la paramétrisation linéaire suivante :

$$\begin{aligned} x_i &= \psi_i \epsilon_i, \\ \psi_i &= w + \sum_{j=1}^m \alpha_j x_{i-j} + \sum_{j=1}^q \beta_j \psi_{i-j}. \end{aligned}$$

Les conditions suffisantes pour assurer la positivité sont $w > 0$, $\alpha_j \geq 0$, $j = 1, \dots, m$ et $\beta_j \geq 0$, $j = 1, \dots, q$.

La moyenne inconditionnelle μ s'écrit :

$$E(x_i) = \mu = \frac{w}{1 - \sum (\alpha_j + \beta_j)}.$$

La moyenne inconditionnelle existe lorsque toutes les racines du polynôme caractéristique associées sont toutes extérieures au cercle unité.

Engle et Russel (1998) ont supposé que la modélisation de la dépendance des durées a une structure assez similaire à une formulation GARCH et on constate que le modèle ACD(1,1) et GARCH(1,1) ont des propriétés similaires. Tout comme le GARCH(1,1) est souvent un bon point de départ, l'ACD(1,1) apparaît comme un point de départ naturel. La connexion avec les modèles GARCH est cependant encore plus profonde. Les théories qui établissent les propriétés de quasi-maximum vraisemblance de GARCH même en présence des racines

unitaires peuvent être reportées sur les modèles de ACD avec erreurs exponentielles. La structure du modèle ACD avec erreurs exponentielles, $\psi_{0,i} = E(x_i | \mathcal{F}_{i-1}) = \omega_0 + \alpha_0 x_i + \beta_0 \psi_{0,i-1}$, ressemble à un GARCH avec innovations normales. En plus, les résultats de la fonction quasi-maximum vraisemblance, définie par :

$$L(\theta) = - \sum_{i=1}^{N(T)} \left\{ \log(\psi_i) + \frac{x_i}{\psi_i} \right\},$$

telle que

$$\psi_i = \begin{cases} \omega + \alpha x_i + \beta \psi_{i-1}, & i > 1, \\ \omega / (1 - \beta), & \text{pour } i = 1, \end{cases}$$

se déduisent des résultats pour les GARCH. Voir Engle et Russel (1998) qui obtiennent un résultat qui est un corollaire d'un résultat de Lee et Hansen (1994).

En choisissant différentes spécifications pour la durée conditionnelle ψ_i et différentes densités $p(\epsilon)$, nous pouvons obtenir plusieurs formes de modèles décrivant la dynamique des durées.

Concernant la densité $p(\epsilon)$, quatre densités ont été utilisées fréquemment pour modéliser la moyenne conditionnelle des durées. Il s'agit des distributions gamma généralisée, Weibull, exponentielle et Burr. Voir par exemple Engle et Russel (1994) et l'ouvrage de Tsay(2010).

Pour une distribution gamma généralisée, la fonction de densité est la suivante :

$$f(\epsilon | \gamma_1, \gamma_2) = \begin{cases} \frac{\gamma_1 \epsilon^{\gamma_1 \gamma_2^{-1}} \exp \left[- \left(\frac{\epsilon}{\gamma_3} \right)^{\gamma_1} \right]}{\gamma_3^{\gamma_1 \gamma_2} \Gamma(\gamma_2)}, & \text{si } \epsilon > 0. \\ 0, & \text{sinon,} \end{cases}$$

avec

$$\gamma_3 = \frac{\Gamma(\gamma_2)}{\Gamma\left(\gamma_2 + \frac{1}{\gamma_1}\right)}, \text{ et } \gamma_1 > 0, \gamma_2 > 0,$$

où $\Gamma(\cdot)$ est la fonction gamma définie de la façon suivante pour $k > 0$:

$$\Gamma(k) = \int_0^{\infty} x^{k-1} \exp(-x) dx.$$

Pour avoir la densité de la distribution Weibull, il faut que γ_2 soit égal à un. De même, pour avoir la densité de la distribution exponentielle, il faut que les paramètres γ_1 et γ_2 soient égaux à un. Engle et Russell (1994) ont utilisé les densités des distributions exponentielles et Weibull résultant des modèles de durées nommés ACD exponentiel (EACD) et Weibull ACD (WACD). Dans le cas exponentiel EACD(1,1), le modèle s'écrit :

$$\psi_i = \omega + \alpha x_{i-1} + \beta \psi_{i-1},$$

avec $\alpha, \beta \geq 0, \omega > 0, \forall i$ pour $(i = 1, \dots, N)$. La variance conditionnelle de ce modèle est ψ_i^2 . La variance inconditionnelle (voir Engle et Russel (1998)) s'écrit :

$$\sigma^2 = \mu^2 \left(\frac{1 - \beta^2 - 2\alpha\beta}{1 - \beta^2 - 2\alpha\beta - 2\alpha^2} \right).$$

Pour une preuve détaillée du résultat on peut consulter Engle et Russell (1994). Lorsque le coefficient α est positif, l'écart-type inconditionnel dépassera la moyenne, présentant un "excès de dispersion" comme documenté dans les ensembles de données de durées. Voir Engle et Russel (1994, 1998).

En analysant des données empiriques, Engle et Russell (1994) ont rejeté la spécification exponentielle car la fonction de hasard d'une distribution exponentielle est constante dans le temps alors que pour le modèle WACD, cette fonction est monotone. Il semblait donc préférable d'adopter une spécification plus générale.

Bauwens et Giot (2000) ont introduit une version logarithmique du modèle ACD qu'ils ont appelé log-ACD et qui est tel que la fonction de durée conditionnelle est l'expression suivante :

$$\psi_i = \exp(w + \alpha \epsilon_{i-1} + \beta \psi_{i-1}).$$

En effet, l'avantage de cette version est qu'elle nous assure la positivité de la durée conditionnelle sans avoir de restrictions sur les paramètres du modèle ACD. La condition de la stationnarité du processus est $|\beta| < 1$. Voir Bauwens et Giot (2000).

La modélisation des séries temporelles financières, en particulier les rendements boursiers et les différences entre les prix quotidiens les plus hauts et les prix quotidiens les plus bas (en anglais *range*), peut être abordée avec des modèles multiplicatifs avec erreurs non-négatives (MEM). Ces modèles imbriquent plusieurs modèles de séries chronologiques bien connus tels que les modèles ACD et les GARCH. Dans la section suivante, nous exposons, en détails, les modèles multiplicatifs avec erreurs non-négatives.

1.6. MODÈLES MEM

Dans l'analyse de séries temporelles financières, le problème est souvent de modéliser des processus impliquant des variables aléatoires qui sont non-négatives. Cela se produit lorsque nous considérons les variables telles que les volumes, les échanges commerciaux, les durées, la volatilité réalisée, la différence entre le prix le plus haut et le prix le plus bas quotidien. Voir Brownlee, Cipollini et Gallo (2011). Traditionnellement, on retrouve deux approches dans un contexte de variables aléatoires non-négatives. D'une part, les méthodes standard sont utilisées, en ignorant simplement cette caractéristique de non-négativité. Une autre approche semble de transformer les variables, en appliquant, par exemple, systématiquement des transformations logarithmiques. Les inconvénients de ces approches ont été discutés par Engle (2002, pp. 428-429). Comme solution, le modèle multiplicatif avec erreurs non-négatives est proposé. Le modèle MEM est formulé de telle sorte que la non-négativité du processus est explicitement prise en compte dans la formulation du modèle. Nous allons discuter ces modèles en détails dans cette section.

Ces modèles ont été introduits par Engle (2002) et ensuite ils ont été développés par Engle et Gallo (2006). Dans cette section, nous présentons la définition du modèle MEM, les conditions de stationnarité et de la non-négativité ainsi qu'une méthode d'estimation. Voir aussi Engle et Gallo (2006) pour plus de

détails.

Soit $\{x_t\}$ un processus non-négatif défini sur $[0, +\infty)$, $t \in \mathbb{N}$, et soit \mathcal{F}_{t-1} l'ensemble des informations disponibles pour le processus jusqu'au temps $t - 1$ inclusivement.

Définition 1.6.1. *Le modèle MEM de $\{x_t\}$ introduit par Engle (2002) est donné par :*

$$x_t = \mu_t \epsilon_t, \quad (1.6.1)$$

où conditionnellement à \mathcal{F}_{t-1} , le processus $\{\mu_t\}$ est \mathcal{F}_t -mesurable et le processus stochastique $\{\epsilon_t\}$ est composé de variables aléatoires indépendantes et identiquement distribuées (iid).

L'espérance conditionnelle μ_t est supposée fonction d'un vecteur de paramètres θ , selon la structure suivante :

$$\mu_t = \mu(\mathcal{F}_{t-1}, \theta).$$

Dans le modèle (1.6.1), la variable aléatoire ϵ_t joue le rôle de l'innovation, et c'est donc une variable aléatoire ayant une densité à support non-négatif, tel que conditionnellement à \mathcal{F}_{t-1} ,

$$E(\epsilon_t) = 1, \quad (1.6.2)$$

$$\text{Var}(\epsilon_t) = \sigma^2, \quad (1.6.3)$$

et telle que la loi de ϵ_t est :

$$\epsilon_t \sim D(1, \sigma^2). \quad (1.6.4)$$

Le paramètre σ^2 est habituellement inconnu.

De ce qui précède, nous avons :

$$E(x_t | \mathcal{F}_{t-1}) = \mu_t, \quad (1.6.5)$$

$$\text{Var}(x_t | \mathcal{F}_{t-1}) = \sigma^2 \mu_t^2. \quad (1.6.6)$$

La structure du modèle MEM est générale et inclut plusieurs modèles comme cas particuliers. Par exemple, le modèle MEM généralise le modèle GARCH discuté dans la section 1.4, en posant x_t comme étant le carré de log du rendement tel que μ_t est la variance conditionnelle. Nous pouvons le définir comme :

$$y_t^2 = \mu_t \epsilon_t^2,$$

ou encore

$$y_t = \sqrt{\mu_t} \epsilon_t,$$

avec

$$\epsilon_t | \mathcal{F}_{t-1} \sim \text{iid}(0, 1).$$

Un autre cas particulier pour le modèle MEM est le modèle ACD discuté dans la section 1.5.2, où $\{x_t\}$ représente le processus des durées entre les transactions financières.

Pour compléter le modèle, nous devons spécifier la distribution de la fonction de densité pour l'innovation ϵ_t et spécifier aussi la structure de l'équation μ_t . Ces considérations sont abordées dans la section suivante.

1.6.1. Spécification de ϵ_t

La distribution du terme d'erreur ϵ_t peut être en principe n'importe quelle distribution à support positif, satisfaisant cependant les équations (1.6.2) et (1.6.3). Nous pouvons citer les distributions gamma, Weibull, gamma inversé ou log-normale.

Bauwens et Giot (2000) ont considéré une distribution Weibull dans l'application d'un modèle admettant la structure $\text{ACD}(\Gamma(1 + \phi)^{-1}, \phi)$. Voir Brownlee, Cipollini et Gallo (2011). Engle et Gallo (2006) ont choisi la distribution gamma pour ϵ_t . Nous présentons en détails le cas où la distribution est de type gamma pour ϵ_t .

Considérons

$$\epsilon_t \sim \text{gamma}(\phi, \phi),$$

avec $E(\epsilon_t) = 1$ et $\text{Var}(\epsilon_t) = \frac{1}{\phi}$. Nous pouvons déduire que, conditionnellement à \mathcal{F}_{t-1} ,

$$x_t | \mathcal{F}_{t-1} \sim \text{gamma} \left(\phi, \frac{\phi}{\mu_t} \right).$$

Il existe des techniques d'estimation qui reposent uniquement sur la spécification des deux premiers moments. Si la distribution de ϵ_t n'est pas spécifiée mais seulement les deux moments, nous disons qu'une spécification semi-paramétrique du modèle est appliquée.

1.6.2. Spécification de μ_t

Comme dans Engle (2002) et Engle et Gallo (2006), nous considérons le modèle GARCH le plus simple, dont les ordres sont (1,1) dans la spécification de l'équation μ_t :

$$\mu_t = \omega + \alpha x_{t-1} + \beta \mu_{t-1}. \quad (1.6.7)$$

Pour l'ordre (p,q) la spécification de μ_t est donnée par :

$$\mu_t = \omega + \sum_{j=1}^p \alpha_j x_{t-j} + \sum_{j=1}^q \beta_j \mu_{t-j}, \quad (1.6.8)$$

avec $\omega > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$, afin d'assurer la positivité de μ_t .

Ce modèle est approprié si x_t est une variable à valeurs positives comme les volumes totaux ou les durées financières. En pratique, nous ne disposons pas que des durées de transactions, mais également des rendements. Des modèles tentent d'incorporer des spécifications tenant compte du signe des rendements financiers en fonction du temps. Considérons la série chronologique des rendements associés, que l'on note dans la suite $\{r_t\}$.

Une spécification générale est donnée par :

$$\mu_t = \omega^* + \alpha \left\{ x_{t-1}^{\frac{1}{2}} \text{signe}(r_{t-1}) + \delta^* \right\}^2 + \gamma x_{t-1} I(r_{t-1} < 0) + \beta \mu_{t-1}. \quad (1.6.9)$$

Dans la spécification (1.6.9), la fonction $\text{signe}(x)$ retourne le signe de la variable x . De plus, $I(A)$ représente la variable indicatrice de l'ensemble A . Les paramètres δ^* et γ ont des interprétations dans la littérature financière liées à des considérations d'asymétrie. Le modèle (1.6.9) a été proposé par Cipollini, Engle et Gallo (2006) et tient compte de modifications des équations GARCH par Ding, Granger et Engle (1993) et Glosten, Jaganathan et Runkle (1993).

Puisque μ_t est une durée conditionnelle, on doit avoir $\mu_t \geq 0, \forall t$. Dans la formulation (1.6.9), on note que des conditions suffisantes sont $\omega^* > 0, \alpha \geq 0, \gamma \geq 0$ et $\beta \geq 0$. Le paramètre δ^* étant dans le carré de (1.6.9), on constate qu'il peut être positif ou négatif.

Cependant, le passage du cas univarié au cas multivarié est facilité en considérant la formulation alternative suivante qui repose sur la paramétrisation :

$$\mu_t = \omega + \alpha x_{t-1} + \gamma x_{t-1}^{(-)} + \delta x_{t-1}^{(s)} + \beta \mu_{t-1}, \quad (1.6.10)$$

où $x_t^{(s)} = x_t^{\frac{1}{2}} \text{signe}(r_t)$, $x_t^{(-)} = x_t I(r_t < 0)$, $\omega = \omega^* + \alpha \{\delta^*\}^2$ et $\delta = 2\alpha \delta^*$.

En fait, l'équation de la moyenne conditionnelle en (1.6.10) est obtenue simplement en développant le carré dans (1.6.9). On peut formuler de manière encore plus générale le modèle comme suit :

$$\mu_t = \omega + \mathbf{x}_{t-1}^{*\top} \boldsymbol{\alpha}^* + \beta \mu_{t-1}, \quad (1.6.11)$$

où $\mathbf{x}_t^* = (x_t; x_t^{(-)}; x_t^{(s)})^\top$ et $\boldsymbol{\alpha}^* = (\alpha; \gamma; \delta)^\top$.

La condition suffisante pour la stationnarité au second ordre de $\{x_t\}$ est :

$$\alpha + \beta + \gamma/2 < 1,$$

qui fut obtenue par Cipollini, Engle et Gallo (2006).

Dans la formulation alternative, il peut être intéressant d'obtenir des conditions sur les coefficients $\omega, \alpha, \gamma, \delta$ et β . On remarque que la positivité est obtenue si $\beta \geq 0$ et si on impose que :

$$\omega + \alpha x_t + \gamma x_t I(r_t < 0) + \delta x_t^{\frac{1}{2}} \text{signe}(r_t) \geq 0, \quad (1.6.12)$$

et ce quelque soit x_t et r_t . Afin d'étudier cette question, le lemme suivant est utile.

Lemme 1.6.1. *La relation*

$$\sum_{i=1}^T (a_i x_i^2 + b_i x_i) + c \geq 0$$

est satisfaite pour tous les $x_i \in [0, +\infty)$, $i = 1, \dots, T$, si et seulement si les coefficients a_i , b_i et c satisfont les conditions suivantes :

C1. Les a_i sont positifs, c'est-à-dire $a_i \geq 0$, $i = 1, \dots, T$;

C2. Les b_i sont positifs, $b_i \geq 0$, $i = 1, \dots, T$, pour tous les i tels que $a_i = 0$;

$$C3. \quad c - \frac{1}{4} \sum_{i=1}^T \frac{b_i^2}{a_i} I(b_i < 0) I(a_i > 0) \geq 0.$$

La preuve de ce résultat se trouve dans Cipollini, Engle et Gallo (2006).

En utilisant le lemme 1.6.1, on trouve les conditions suffisantes pour la non-négativité de μ_t :

$$(1) \quad \beta \geq 0, \alpha \geq 0, \alpha + \gamma \geq 0;$$

$$(2) \quad \text{Si } \alpha = 0 \text{ alors } \delta \geq 0. \text{ Si } \alpha + \gamma = 0 \text{ alors } \delta \leq 0;$$

$$(3) \quad \omega - \frac{\delta^2}{4} \left\{ \frac{I(\delta < 0) I(\alpha > 0)}{\alpha} + \frac{I(\delta > 0) I(\alpha + \gamma > 0)}{\alpha + \gamma} \right\} \geq 0.$$

En effet, distinguons dans (1.6.12) les cas $r_t \geq 0$ et $r_t < 0$. Si $r_t \geq 0$, alors nous devons étudier quand

$$\omega + \delta x_t^{\frac{1}{2}} + \alpha x_t \geq 0.$$

Une utilisation du lemme 1.6.1 avec $T = 1$ donne :

- (1) $\alpha \geq 0$;
- (2) Si $\alpha = 0$ alors $\delta \geq 0$;
- (3) $\omega - \frac{1}{4} \frac{\delta^2}{\alpha} I(\delta < 0) I(\alpha > 0) \geq 0$.

Si maintenant $r_t < 0$, alors nous devons étudier quand :

$$\omega - \delta x_t^{\frac{1}{2}} + (\alpha + \gamma)x_t \geq 0.$$

Une autre utilisation du lemme 1.6.1 donne les conditions :

- (1) $\alpha + \gamma \geq 0$;
- (2) Si $\alpha + \gamma = 0$ alors $\delta \leq 0$;
- (3) $\omega - \frac{1}{4} \frac{\delta^2}{(\alpha + \gamma)} I(\delta > 0) I(\alpha + \gamma > 0) \geq 0$.

Quand on combine les deux ensembles des conditions, on trouve que les conditions suffisantes pour la non-négativité de μ_t sont celles avancées précédemment.

Passons maintenant à la méthode d'estimation des paramètres.

1.6.3. Estimation et inférence du modèle MEM

Considérons une spécification $\text{gamma}(\phi, \phi)$ pour la loi de l'innovation ϵ_t . Ainsi, on déduit que :

$$x_t | \mathcal{F}_{t-1} \sim \text{gamma} \left(\phi, \frac{\phi}{\mu_t} \right).$$

La fonction de vraisemblance, notée L_t , de x_t est donnée par l'expression suivante :

$$L_t = \frac{\left(\frac{\phi}{\mu_t} \right)^\phi}{\Gamma(\phi)} x_t^{\phi-1} \exp \left(-\phi \frac{x_t}{\mu_t} \right),$$

et son logarithme, noté l_t , est égal à l'expression :

$$l_t = \ln L_t = \phi \ln \phi - \ln \Gamma(\phi) + (\phi - 1) \ln x_t - \phi \left(\ln \mu_t + \frac{x_t}{\mu_t} \right).$$

Considérons $\lambda = (\theta^\top; \phi^\top)^\top$, le vecteur λ contient deux sous-vecteurs θ et ϕ , ce qui revient à résoudre les deux sous-systèmes :

$$\begin{aligned} \nabla_\lambda l_t = 0 &\iff \nabla_\theta l_t = 0, \\ &\nabla_\phi l_t = 0. \end{aligned}$$

L'opérateur ∇_λ représente les dérivées par rapport aux composantes de λ . Donc,

$$\begin{aligned} \nabla_\theta l_t &= \phi \nabla_\theta \mu_t \left(\frac{x_t - \mu_t}{\mu_t^2} \right), \\ \nabla_\phi l_t &= \ln \phi + 1 - \Psi(\phi) + \ln \left(\frac{x_t}{\mu_t} \right) - \frac{x_t}{\mu_t}. \end{aligned}$$

où $\Psi(\phi) = \frac{\Gamma'(\phi)}{\Gamma(\phi)}$ est la fonction parfois nommée digamma. Voir Cipollini, Engle et Gallo (2006).

Ainsi, le système du premier ordre revient à résoudre les équations suivantes :

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \nabla_\theta \mu_t \frac{x_t - \mu_t}{\mu_t^2} &= 0, \\ \ln \phi + 1 - \Psi(\phi) + \frac{1}{T} \sum_{t=1}^T \left[\ln \left(\frac{x_t}{\mu_t} \right) - \frac{x_t}{\mu_t} \right] &= 0. \end{aligned}$$

Un examen des conditions du premier ordre mène à la remarque suivante : le système est indépendant du paramètre ϕ dans la distribution gamma. Ceci implique que quelque soit la valeur de ϕ , toute distribution gamma dans la famille $\{\text{gamma}(\phi, \phi), \phi > 0\}$ pour le terme d'erreur d'un modèle MEM mènera aux mêmes estimations ponctuelles.

La contribution de x_t à la matrice Hessienne est donnée par l'expression suivante :

$$\mathbf{H}_t = \begin{pmatrix} \mathbf{H}_{t,\theta\theta'} & \mathbf{H}_{t,\theta\phi} \\ \mathbf{H}_{t,\theta\phi}^\top & \mathbf{H}_{t,\phi\phi} \end{pmatrix},$$

dont les différentes matrices blocs sont précisément données par les expressions :

$$\begin{aligned} \mathbf{H}_{t,\theta\theta'} = \nabla_{\theta\theta'} l_t &= \phi \left(\frac{-2x_t + \mu_t}{\mu_t^3} \nabla_{\theta} \mu_t \nabla_{\theta'} \mu_t + \frac{x_t - \mu_t}{\mu_t^2} \nabla_{\theta\theta'} \mu_t \right), \\ \mathbf{H}_{t,\theta\phi} = \nabla_{\theta\phi} l_t &= \frac{x_t - \mu_t}{\mu_t^2} \nabla_{\theta} \mu_t, \\ \mathbf{H}_{t,\phi\phi} = \nabla_{\phi\phi} l_t &= \frac{1}{\phi} - \Psi'(\phi), \end{aligned}$$

où la fonction $\Psi'(\phi)$ est la fonction trigamma. Voir Cipollini, Engle et Gallo (2006). La matrice des variances et covariances asymptotique de l'estimation par maximum de vraisemblance est obtenue selon certaines conditions de régularité par l'expression suivante :

$$V_{\infty} = \begin{pmatrix} \phi \lim_{t \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \frac{1}{\mu_t^2} \nabla_{\theta} \mu_t \nabla_{\theta'} \mu_t & \mathbf{0} \\ \mathbf{0} & \Psi'(\phi) - \frac{1}{\phi} \end{pmatrix}^{-1}.$$

Notons que même si ϕ n'est pas impliqué dans l'estimation de θ , la variance de $\hat{\theta}$ est proportionnelle à $\frac{1}{\phi}$. Notons aussi que les estimateurs $\hat{\theta}$ et $\hat{\phi}$ obtenus par maximum de vraisemblance sont asymptotiquement non-corrélés.

Dans le chapitre suivant, nous présentons les modèles vectoriels et multiplicatifs avec erreurs non-négatives vMEM qui sont une simple généralisation des modèles MEM. Dans un cadre multivarié, nous introduisons la définition du processus et la notion de stationnarité de $\{\mathbf{x}_t\}$, la notion de la non-négativité de μ_t , ainsi que les méthodes d'estimation existantes pour ces modèles.

Chapitre 2

MODÈLES vMEM

Dans ce chapitre, nous introduisons la classe de modèles vMEM. Comme mentionné précédemment, un tel modèle nécessite la spécification d'une loi multivariée dont toutes les composantes sont non-négatives. Nous présentons dans un premier lieu la distribution gamma multivariée proposée et étudiée dans Cheriyan (1941) et Ramabhadran (1951). Cependant, compte tenu des limites de cette loi, nous considérons l'utilisation des copules. Nous rappelons quelques définitions de base qui nous permettront de décrire le concept de copule. Puis, nous résumons les concepts et les propriétés des copules les plus importants dans l'étude qui seront nécessaires pour la spécification de la distribution des innovations dans le cas des modèles d'erreurs multivariés. Ensuite, nous présentons en détail les modèles vMEM introduits par Engle (2002) et développés par Engle et Gallo (2006). Nous présentons ensuite trois méthodes d'estimation. Pour l'approche paramétrique, nous considérons la méthode du maximum de vraisemblance (plus précisément la méthode du maximum de vraisemblance concentrée), tandis que pour l'approche semi-paramétrique nous présentons la méthode des moments généralisés (GMM) et la méthode des erreurs standard associées aux paramètres (WNLS).

2.1. DISTRIBUTION GAMMA MULTIVARIÉE

Une approche naturelle consiste à spécifier directement la distribution du vecteur aléatoire ϵ_t . Cependant, comme les composantes doivent être non-négatives, on ne peut pas utiliser une distribution comme la normale multivariée. En fait, compte tenu des propriétés de la loi gamma, il peut être utile de considérer la loi gamma multivariée. La littérature suggère que plusieurs généralisations sont possibles. On trouve par exemple la loi gamma bivariée développée par Cheriyan (1941). Elle a été généralisée au cas complètement multivarié par Ramabhadran (1951). Cette distribution, notée GammaCR, est

détaillée dans l'ouvrage de Johnson, Kotz et Balakrishnan (2000), qui porte sur les propriétés de plusieurs lois multivariées continues. Cette distribution a aussi été considérée par Cipollini, Engle et Gallo (2006, Annexe C).

Afin de construire un vecteur aléatoire $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{Kt})^\top$ selon la loi GammaCR, on considère les variables aléatoires $Y_{0t}, Y_{1t}, \dots, Y_{Kt}$ indépendantes et de distribution gamma. On a alors :

$$Y_{0t} \sim \text{gamma}(\phi_0, 1), \quad (2.1.1)$$

$$Y_{it} \sim \text{gamma}(\phi_i - \phi_0, 1), \quad (2.1.2)$$

et ce pour $i = 1, \dots, K$. Les paramètres ϕ_0 et ϕ_i doivent être spécifiés et satisfont les contraintes $0 < \phi_0 < \phi_i, i = 1, \dots, K$. Posons

$$\epsilon_{it} = \frac{Y_{it} + Y_{0t}}{\phi_i}, i = 1, \dots, K.$$

On dira que ϵ_t admet une distribution GammaCR avec paramètres ϕ_0 et $\boldsymbol{\phi} = (\phi_1, \dots, \phi_K)^\top$.

On note cette distribution :

$$\epsilon_t \sim \text{GammaCR}(\phi_0, \boldsymbol{\phi}).$$

Cipollini, Engle et Gallo (2006) présentent une version plus générale de la distribution de Cheriyan et Ramabhadran, qui repose sur un paramètre additionnel de dimension K . Cependant, ils considèrent la distribution $\text{GammaCR}(\boldsymbol{\phi}, \boldsymbol{\phi})$ dans leurs développements et motivations.

On note que la dépendance est introduite compte tenu de la variable aléatoire Y_{0t} .

La fonction de densité est donnée par :

$$f(\boldsymbol{\epsilon}_t) = \frac{1}{\Gamma(\phi_0)} \prod_{i=1}^K \frac{\phi_i}{\Gamma(\phi_i - \phi_0)} e^{-\sum_{i=1}^K \phi_i \epsilon_{i,t}} \int_0^v y_0^{\phi_0-1} e^{(K-1)y_0} \prod_{i=1}^K (\phi_i \epsilon_{i,t} - y_0)^{\phi_i - \phi_0 - 1} dy_0,$$

où $v = \min(\phi_1 \epsilon_{1,t}, \dots, \phi_K \epsilon_{K,t})$. Remarquons que la fonction de densité de

cette distribution est formulée en fonction de produits et d'une intégrale. Cette fonction de densité devient rapidement compliquée à calculer à mesure que K devient grand.

Comme mentionné par Cipollini, Engle et Gallo (2006), toutes les lois marginales de la distribution GammaCR sont gamma, c'est-à-dire que :

$$\epsilon_{it} \sim \text{gamma}(\phi_i, \phi_i).$$

On peut facilement déterminer la structure de dépendance entre les composantes de ϵ_t . En effet

$$\text{Cov}(\epsilon_{i,t}, \epsilon_{j,t}) = \text{Cov}\left(\frac{Y_{it} + Y_{0t}}{\phi_i}, \frac{Y_{jt} + Y_{0t}}{\phi_j}\right), \quad (2.1.3)$$

$$= \frac{1}{\phi_i \phi_j} \text{Cov}(Y_{0t}, Y_{0t}), \quad (2.1.4)$$

$$= \frac{\text{Var}(Y_{0t})}{\phi_i \phi_j}, \quad (2.1.5)$$

$$= \frac{\phi_0}{\phi_i \phi_j}. \quad (2.1.6)$$

On note que ces covariances sont toutes strictement positives, ce qui peut être considéré comme un désavantage.

La corrélation entre $\epsilon_{i,t}$ et $\epsilon_{j,t}$, $i \neq j$, est donnée par :

$$\rho(\epsilon_{i,t}, \epsilon_{j,t}) = \frac{\phi_0}{\sqrt{\phi_i \phi_j}}.$$

Puisque $\epsilon_{it} \sim \text{gamma}(\phi_i, \phi_i)$, on trouve que $\text{Var}(\epsilon_{it}) = 1/\phi_i$, $i = 1, \dots, K$. On remarque ainsi une dépendance de la mesure de corrélation avec les variances des marges. Ainsi, si les marges sont toutes $\text{gamma}(1,1)$, avec $\phi_0 = 0,5$, alors la corrélation entre deux composantes est toujours de 50%. Ces restrictions sur la structure de corrélation sont suffisamment importantes pour considérer une solution alternative reposant sur la théorie des copules.

2.2. PROPRIÉTÉS FONDAMENTALES DES COPULES

Dans cette partie, après un bref historique du concept de copule, nous définissons ce qu'est une copule, et nous fournissons les propriétés les plus importantes pour notre étude.

Le concept de copule est lié à un théorème qui a été exposé dans Sklar (1959), dans un article rédigé en français. Les copules sont des fonctions qui relient ensemble les distributions univariées afin de créer des distributions multivariées. Avec une fonction de copule, il est possible de distinguer les rôles respectifs des fonctions de distributions multivariées avec les distributions marginales univariées (les marges). Un ouvrage qui fait un survol des débuts de cette théorie à 1999 est Nelsen (1999). On note que l'étude des copules et leur applications en statistique a vraiment connu des développements ces quinze dernières années. En plus des applications en statistique multivariée, la finance, l'actuariat et l'économétrie font maintenant appel à ces méthodes. Une simple application de Google avec les mots-clés "copula Sklar" donne plus de 25000 résultats.

Le terme copule en soi (ou *copula* en anglais) vient du mot latin "copŭlæ", qui signifie liaison, lien, alliance ou union. On peut consulter le dictionnaire Cassell's Latin Dictionary.

La définition suivante stipule qu'une copule n'est rien d'autre qu'une fonction de répartition avec marges uniformes.

Définition 2.2.1. *Une copule est définie comme étant une fonction de répartition multivariée ayant des lois marginales uniformes sur $[0, 1]$.*

La théorie des copules est un ensemble de techniques très puissantes afin de comprendre la dépendance de distributions à plusieurs dimensions. Il est possible de construire des distributions, et de les simuler, et de considérer ainsi des familles de distributions bidimensionnelles et multidimensionnelles dont les marges sont choisies arbitrairement par l'analyste.

La théorie de la statistique a longtemps considéré la loi normale comme étant centrale. De plus, l'extension de la loi normale univariée au cas multivarié donne une distribution multivariée aux propriétés tout à fait remarquables. Si un vecteur est de distribution multinormale, alors les marges sont normales et même les lois conditionnelles sont également de lois normales. Voir par exemple Johnson et Wichern (1992).

Avant de présenter les propriétés des copules à n dimensions, nous rappelons la définition de la fonction de répartition conjointe et des lois marginales qui

nous permettront de décrire le concept de copule.

2.2.1. Définition de la fonction de répartition conjointe et des lois marginales

Soit \mathbf{X} un vecteur aléatoire en K -dimensions $\mathbf{X} = (X_1, X_2, \dots, X_K)^\top$. On appelle alors fonction de répartition conjointe de \mathbf{X} , la fonction de K variables réelles définie par :

$$F(x_1, x_2, \dots, x_K) = P(X_1 \leq x_1, \dots, X_K \leq x_K).$$

Cette fonction de répartition conjointe nous permet d'établir les fonctions de répartition F_1, \dots, F_K des variables aléatoires marginales X_1, \dots, X_K (donc les fonctions de répartition marginales) comme suit :

$$F_i(x_i) = P(X_i \leq x_i) = F(\infty, \dots, \infty, x_i, \dots, \infty, \dots, \infty).$$

Dans la suite, nous allons donner la définition d'une copule de dimension K .

Définition 2.2.2. Une copule de dimension K (ou K -copule) est une fonction C définie de I^K dans I avec $I = [0, 1]$ telle que :

- (1) Pour tout vecteur aléatoire $\mathbf{u} = (u_1, \dots, u_K) \in I^K$, $C(\mathbf{u}) = 0$ si au moins une des coordonnées est nulle.
- (2) Si toutes les coordonnées de \mathbf{u} sont un à l'exception de u_k alors :
 $C(\mathbf{u}) = C(1, \dots, 1, u_k, 1, \dots, 1) = u_k$ pour tout $k \in 1, \dots, K$.

Nous présentons maintenant le théorème de Sklar, qui comme mentionné est central dans l'étude des distributions multivariées .

Théorème 2.2.1 (Le théorème de Sklar). Soit H une fonction de répartition en K dimensions ayant des distributions marginales F_1, F_2, \dots, F_K . Alors il existe une K -copule C telle que pour tout $\mathbf{x} \in [-\infty, \infty]^K$:

$$H(x_1, x_2, \dots, x_K) = C(F_1(x_1), \dots, F_K(x_K)).$$

Si les marginales sont continues alors C est unique ; sinon C est uniquement déterminée en fonction de $\text{Ran}F_1 * \text{Ran}F_2 * \dots * \text{Ran}F_K$, où l'expression Ran est l'abréviation du support (ou range) de F qui est un sous-ensemble de $[-\infty, +\infty]$.

Inversement, si C est une copule à K dimensions et que F_1, \dots, F_K sont des fonctions de distribution, alors la fonction H définie comme ci-dessus est une fonction de répartition à K -dimensions ayant pour marginales F_1, F_2, \dots, F_K .

Avant d'aborder le corollaire du théorème de Sklar qui utilise des fonctions inverses généralisées, il convient tout d'abord de donner une définition de celle-ci.

Définition 2.2.3. Si F est une fonction de répartition, alors une fonction inverse généralisée de F pourra être toute fonction $F^{(-1)}$ définie sur $I = [0, 1]$ telle que :

- (1) Si $t \in \text{Ran}F$, alors $F^{(-1)}(t)$ est un nombre $x \in [-\infty, \infty]$ tel que $F(x) = t$.
Donc, pour tout $t \in \text{Ran}F$, $F(F^{(-1)}(t)) = t$.
- (2) Si $t \notin \text{Ran}F$ alors $F^{(-1)}(t) = \inf \{x \mid F(x) \geq t\} = \sup \{x \mid F(x) \leq t\}$.

Corollaire 2.2.1 (Inversion du théorème de Sklar). Définissons H, C, F_1, \dots, F_K comme dans le théorème de Sklar précédent et notons $F_1^{(-1)}, \dots, F_K^{(-1)}$ les fonctions inverses généralisées de F_1, \dots, F_K . Alors, pour tout vecteur $\mathbf{u} \in [0, 1]^K$:

$$C(\mathbf{u}_1, \dots, \mathbf{u}_K) = H(F_1^{(-1)}(\mathbf{u}_1), \dots, F_K^{(-1)}(\mathbf{u}_K)).$$

Le théorème de Sklar est utilisé pour construire des copules à partir des distributions multidimensionnelles.

Les propriétés suivantes d'une copule multivariée sont particulièrement importantes.

Théorème 2.2.2 (Les bornes de Fréchet-Hoeffding). Les bornes d'une copule de dimension K s'énoncent comme suit :

$$\max \left(\sum_{i=1}^K u_i + 1 - K, 0 \right) \leq C(\mathbf{u}) \leq \min(u_1, \dots, u_K).$$

Les bornes de Fréchet-Hoeffding sont utiles car elles stipulent les cas extrêmes pour la fonction de copule C . Une question naturelle consiste à savoir si ces bornes sont atteintes et s'il y a des interprétations intéressantes à ces cas limites. L'ouvrage de Joe (1997) est utile à cette fin. Il ressort que la borne supérieure est toujours une fonction de répartition, et donc peut servir afin de construire des lois multivariées. Voir le théorème 3.2 dans Joe (1997, p. 58). Le cas de la borne inférieure est beaucoup plus compliqué. Pour les distributions bivariées, la borne inférieure de Fréchet-Hoeffding est une fonction de répartition. Cependant, ce résultat n'est plus vrai dans le cas $K > 3$. Cependant des conditions peuvent être formulées pour que la borne inférieure soit une fonction de répartition. Voir Joe (1997, pp. 60-65).

Dans le théorème qui suit, nous abordons la copule associée à l'indépendance entre les composantes du vecteur aléatoire.

Théorème 2.2.3 (Notion d'indépendance). *Pour $K \geq 2$, soient X_1, X_2, \dots, X_K des variables aléatoires continues. Alors :*

- (1) X_1, X_2, \dots, X_K sont indépendantes si et seulement si la K -copule de X_1, X_2, \dots, X_K est $\prod^K(\mathbf{u}) = u_1 u_2 \cdots u_K$.
- (2) Chacune des variables aléatoires X_1, X_2, \dots, X_K est une fonction absolument croissante de n'importe quelle des autres si et seulement si la K -copule de X_1, X_2, \dots, X_K a K -copule égale à $\min(u_1, u_2, \dots, u_K)$.

Ce résultat est énoncé dans Nelsen (1999, p.43). Dans le cas de distributions continues, il est utile de présenter la densité d'une copule multivariée. Ceci fait l'objet de la prochaine définition.

Définition 2.2.4 (Densité de la copule multivariée). *La densité f d'une fonction de répartition F est définie comme suit :*

$$f(x_1, \dots, x_K) = \frac{\partial F(x_1, \dots, x_K)}{\partial x_1 \dots \partial x_K}.$$

La densité de la copule, fonction que l'on note $c(u_1, \dots, u_K)$, s'exprime donc comme suit :

$$c(u_1, \dots, u_K) = \frac{\partial C(u_1, \dots, u_K)}{\partial u_1 \dots \partial u_K}.$$

Si nous appelons f_i la densité de la i -ème marginale, alors la densité f de F s'exprime alors comme suit :

$$f(x_1, \dots, x_K) = c(F_1(x_1), \dots, F_K(x_K)) \times \prod_{i=1}^K f_i(x_i).$$

Cette formule est très importante lorsque la méthode du maximum de vraisemblance est adoptée dans le contexte pleinement paramétrique. Ce résultat important est très utile dans Cipollini, Engle et Gallo (2006).

Il existe plusieurs types de copules. On peut énoncer les copules archimédiennes, les copules elliptiques et les copules à valeurs extrêmes. Nous nous intéressons aux copules elliptiques qui sont associées aux distributions elliptiques. Cette famille de copules contient la copule gaussienne ou normale ainsi que la copule de Student. Les copules elliptiques sont attrayantes parce qu'elles peuvent être utilisées également lorsque le nombre d'observations est modérément grand. De même, elles peuvent décrire la dépendance dans les queues des distributions. Voir McNeil, Frey et Embrechts (2005).

Nous nous intéressons à présenter la densité de la copule gaussienne ainsi que la méthode pour la simuler.

Théorème 2.2.4 (Densité d'une copule normale multivariée). *La densité d'une copule normale multivariée s'écrit comme suit :*

$$c(\mathbf{u}) = |\mathbf{R}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \mathbf{q}^\top (\mathbf{R}^{-1} - \mathbf{I}) \mathbf{q}\right],$$

où le terme \mathbf{R} est la matrice de corrélation, la valeur de q_i est $q_i = \Phi^{-1}(u_i)$ tel que $\Phi(\cdot)$ représente la fonction de répartition de la loi normale $\mathcal{N}(0, 1)$.

Dans notre étude de simulations, il sera particulièrement important de simuler une copule gaussienne dont les marges sont gamma. Nous considérons dans la prochaine section la simulation d'une copule normale multivariée dont les marges sont premièrement de distributions arbitraires, deuxièmement normales et troisièmement lorsque les marges sont de distributions gamma.

2.2.2. Simuler une copule normale

L'algorithme que nous avons mis en œuvre afin de simuler une copule gaussienne multivariée repose sur les étapes suivantes.

- (1) Choisir une matrice de corrélation arbitraire Σ telle que Σ est une matrice carrée symétrique de diagonale 1 et qui a comme élément $\rho_{ij} = \rho_{ji}$. Le coefficient ρ_{ij} représente la corrélation entre la variable i et la variable j .
- (2) Calculer la décomposition de Cholesky qui consiste à trouver une matrice A telle que $\Sigma = A^T A$.
- (3) Générer K variables aléatoires centrées réduites iid $\tilde{X}_1, \dots, \tilde{X}_K$.
- (4) Calculer $(X_1, \dots, X_K)^T = \mathbf{X} = A\tilde{\mathbf{X}}$ où $\tilde{\mathbf{X}} = (\tilde{X}_1, \dots, \tilde{X}_K)^T$.
- (5) Poser $U_i = \Phi(X_i)$, $i = 1, \dots, K$, où $\Phi(\cdot)$ est la fonction de répartition de loi normale $\mathcal{N}(0, 1)$ et $\mathbf{U} = (\Phi(X_1), \dots, \Phi(X_K))^T$.
- (6) Soit $X_i = \Phi^{-1}(U_i)$, où $U_i \sim \mathcal{U}(0, 1)$. Alors la fonction de répartition de X_i est une loi normale $\mathcal{N}(0, 1)$. Ce résultat classique permet donc de trouver les marges.

Dans l'algorithme précédent, s'il est souhaité que les distributions marginales soient normales, il suffit de faire les quatre premières étapes et de calculer l'inverse de la fonction de répartition de la loi normale. Ceci correspond à utiliser la fonction `qnorm(·)` en R. Cependant, si une copule gaussienne avec des marges de lois gamma est ce qui doit être simulée, alors c'est l'inverse de la fonction de répartition d'une loi gamma qui doit être calculé, en utilisant la fonction `qgamma(·)` en R. Voir aussi McNeil, Frey et Embrechts (2005) pour plus de détails.

Dans la prochaine section, nous présenterons la définition des modèles vMEM et nous exposerons les principales caractéristiques de ces modèles.

2.3. PRÉSENTATION DU MODÈLE vMEM

Dans le chapitre précédent, nous avons présenté les modèles MEM qui permettent de modéliser une série chronologique univariée. Donc les modèles

MEM sont utiles lorsque l'on dispose d'une réalisation finie d'une variable dans le temps. Cependant, dans les applications, plusieurs variables sont observées dans le temps et il peut être pertinent de procéder à une modélisation simultanée. Parmi les applications évidentes, l'étude des relations de causalité entre les variables peut s'avérer d'intérêt. De plus, si certaines variables sont reliées, elles peuvent être utilisées afin de proposer des prévisions plus précises. De manière générale, la modélisation de plusieurs séries chronologiques univariées amène l'étude des séries chronologiques multivariées.

Les modèles vectoriels et multiplicatifs avec erreurs non-négatives (vMEM) permettent de modéliser des vecteurs aléatoires avec composantes qui sont toutes non-négatives. Ils ont été introduits par Engle (2002) et développés par Engle et Gallo (2006). Ces modèles représentent une généralisation au cas multivarié des modèles MEM introduits par Engle (2002). Il s'agit d'une représentation appropriée des processus non-négatifs qui ont des interactions dynamiques entre eux. La prochaine section décrit des exemples d'utilisation de la classe des modèles vMEM.

2.3.1. Exemples d'utilisation de l'approche vMEM

Il est souvent d'intérêt de modéliser conjointement la dynamique de deux ou plusieurs séries temporelles non-négatives. Nous présentons dans cette partie quatre exemples où une approche multivariée MEM semble plus naturelle.

(1) Étude de la volatilité

Une vaste littérature porte sur les données à haute fréquence où les séries financières sont des rendements dont on veut décrire la volatilité. Un objectif pourrait être de mesurer les interactions dynamiques entre différentes mesures financières telles la volatilité réalisée, la différence entre le prix le plus haut et le prix le plus bas (donc étudier l'étendue ou le *range* en anglais), les rendements absolus, etc. Ces mesures expliquent l'information contenue dans l'activité intra-journalière de la volatilité. Ainsi, avec les modèles vMEM, on peut considérer un modèle global où ces mesures peuvent interagir. Plus précisément, nous pouvons étudier si ces mesures sont indépendantes des autres, si ces mesures s'expliquent en fonction de leur propre passé, où encore si ces

mesures sont expliquées par des variables retardées d'autres mesures.

(2) Transfert de volatilité

Une vaste littérature tente de décrire des mécanismes de transmission de la volatilité (en anglais, *volatility spillover*). Par transfert de volatilité, il est question de transfert ou de contagion à travers les marchés. Avec les modèles vMEM, nous pouvons construire un modèle d'interdépendance où nous utilisons un proxy de la volatilité (par exemple la variation journalière) pour les différents marchés et où l'on tente d'analyser les interactions selon le modèle choisi.

(3) Dynamique de l'exécution des commandes

Dans les marchés boursiers, il y a un compromis entre la possibilité de soumettre une commande au meilleur prix contre le risque que ces commandes ne soient jamais exécutées. On pourrait s'intéresser alors à la distribution de la quantité de stocks qui vont s'exécuter dans une période de temps donnée à une distance donnée du prix courant. On parle alors de profondeur d'exécution. La présence de zéros est possible ici car il y a des situations où la quantité de stocks pouvant être exécutés à une certaine distance donnée pour le prix actuel est tout simplement nulle. Il peut être pertinent d'étudier la dynamique entre les quantités, les prix, etc. Il pourrait être pertinent de procéder à une modélisation vMEM dans un contexte de prévision. Voir Noss (2007).

(4) Dynamiques entre les transactions, les durées, les volumes et la volatilité

Dans un contexte où les données sont observées à ultra haute fréquence, les transactions surviennent à des instants irréguliers. On a déjà introduit les modèles ACD afin de décrire les durées. De façon réaliste, les analystes ne disposeraient pas seulement des durées, mais également des mesures de volume et les mesures de rendement. Cette situation suggère un modèle multivarié pour les quantités en cause. Une modélisation possible a été proposée par Manganelli (2002). Une question naturelle porte sur la dynamique existante entre les différentes mesures.

Plus spécifiquement, il peut être pertinent de vérifier si l'interaction entre les variables en cause peut finalement révéler la vitesse de transmission de l'information privée dans les prix. En utilisant une modélisation vMEM, le vecteur aléatoire pourrait avoir comme composantes les durées, les volumes et les volatilités. Dans le modèle vMEM, on peut intégrer les retards des variables en cause dans le système, et la matrice de variance du terme d'erreur permet de tenir compte de corrélations contemporaines entre les composantes du terme d'erreur.

Dans la prochaine section, nous présentons la définition du modèle vMEM. Nous étudions les conditions de stationnarité ainsi que les conditions de non-négativité. Des méthodes d'estimation sont aussi discutées. Les résultats exposés dans ce qui suit proviennent de Cipollini, Engle et Gallo (2006).

2.3.2. Définition et formulation

Soit $\{\mathbf{x}_t\}$ un processus dont les composantes sont non-négatives. Le vecteur aléatoire \mathbf{x}_t est de dimension $K \times 1$, tel que $\mathbf{x}_t = (x_{1,t}, \dots, x_{K,t})^\top$, et donc $x_{i,t} \geq 0$ avec probabilité un. On observe la série chronologique \mathbf{x}_t , $t = 1, \dots, T$. On dira que $\{\mathbf{x}_t\}$ suit un modèle MEM vectoriel, ou encore vMEM, s'il s'exprime comme :

$$\mathbf{x}_t = \boldsymbol{\mu}_t \odot \boldsymbol{\epsilon}_t, \quad (2.3.1)$$

$$= \text{diag}(\boldsymbol{\mu}_t) \boldsymbol{\epsilon}_t, \quad (2.3.2)$$

où le symbole ' \odot ' désigne le produit d'Hadamard. On rappelle que pour deux matrices A et B , chacune de dimension $n \times m$, alors

$$A \odot B = C,$$

où C est de dimension $n \times m$ et telle que $C = (c_{ij})$, $c_{ij} = a_{ij}b_{ij}$, avec $A = (a_{ij})$ et $B = (b_{ij})$. C'est donc simplement le produit élément par élément de deux matrices de même dimension. Pour un vecteur $K \times 1$ \mathbf{a} , $\text{diag}(\mathbf{a})$ représente la matrice diagonale $K \times K$ avec les composantes de \mathbf{a} sur la diagonale principale.

Conditionnellement à l'information disponible jusqu'à l'instant $t - 1$, notée encore \mathcal{F}_{t-1} , la moyenne conditionnelle est $\boldsymbol{\mu}_t$, que l'on suppose satisfaire la structure :

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}(\mathcal{F}_{t-1}, \boldsymbol{\theta}),$$

où $\boldsymbol{\mu}_t$ est un vecteur aléatoire de dimension $K \times 1$. Le vecteur de paramètres est contenu dans $\boldsymbol{\theta}$.

Le terme d'erreur $\{\boldsymbol{\epsilon}_t\}$ est un processus stochastique que l'on peut supposer iid, tel que $\boldsymbol{\epsilon}_t$ est de dimension $K \times 1$. Le support de $\boldsymbol{\epsilon}_t$ est $[0, +\infty)^K$. On suppose que $E(\boldsymbol{\epsilon}_t) = \mathbf{1}$ où le vecteur $\mathbf{1}$ est tel que $\mathbf{1} = (1, \dots, 1)^\top$ et $\text{Var}(\boldsymbol{\epsilon}_t) = \Sigma$.

On note

$$\boldsymbol{\epsilon}_t \sim D(\mathbf{1}, \Sigma). \quad (2.3.3)$$

Ainsi, on obtient que

$$E(\mathbf{x}_t | \mathcal{F}_{t-1}) = E(\text{diag}(\boldsymbol{\mu}_t)\boldsymbol{\epsilon}_t | \mathcal{F}_{t-1}), \quad (2.3.4)$$

$$= \text{diag}(\boldsymbol{\mu}_t)E(\boldsymbol{\epsilon}_t | \mathcal{F}_{t-1}), \quad (2.3.5)$$

$$= \text{diag}(\boldsymbol{\mu}_t)\mathbf{1}, \quad (2.3.6)$$

$$= \boldsymbol{\mu}_t. \quad (2.3.7)$$

De même

$$\text{Var}(\mathbf{x}_t | \mathcal{F}_{t-1}) = \text{Var}(\text{diag}(\boldsymbol{\mu}_t)\boldsymbol{\epsilon}_t | \mathcal{F}_{t-1}), \quad (2.3.8)$$

$$= \text{diag}(\boldsymbol{\mu}_t)\text{Var}(\boldsymbol{\epsilon}_t | \mathcal{F}_{t-1})\text{diag}(\boldsymbol{\mu}_t), \quad (2.3.9)$$

$$= \text{diag}(\boldsymbol{\mu}_t)\Sigma\text{diag}(\boldsymbol{\mu}_t), \quad (2.3.10)$$

$$= \boldsymbol{\mu}_t\boldsymbol{\mu}_t^\top \odot \Sigma, \quad (2.3.11)$$

où l'on utilise la relation bien connue $\text{Var}(A\mathbf{x}) = A\text{Var}(\mathbf{x})A^\top$ et les propriétés du produit d'Hadamard.

2.3.3. Spécification de $\boldsymbol{\mu}_t$

Comme dans le cas univarié des modèles vMEM, la spécification la plus simple d'ordre (1,1) de la moyenne conditionnelle $\boldsymbol{\mu}_t$ est la suivante :

$$\boldsymbol{\mu}_t = \boldsymbol{\omega} + \boldsymbol{\alpha}\mathbf{x}_{t-1} + \boldsymbol{\beta}\boldsymbol{\mu}_{t-1},$$

où le vecteur $\boldsymbol{\omega}$ a comme dimension $K \times 1$ et les matrices $\boldsymbol{\alpha}$ et $\boldsymbol{\beta}$ ont pour dimension $K \times K$.

Si nous sommes intéressé à étudier les rendements, il est pertinent de tenir compte du signe de rendements financiers en fonction du temps. Considérons la série chronologique des rendements associés, que l'on note $\{r_t\}$. La spécification générale de la moyenne conditionnelle est la suivante :

$$\boldsymbol{\mu}_t = \boldsymbol{\omega} + \boldsymbol{\alpha}\mathbf{x}_{t-1} + \boldsymbol{\gamma}\mathbf{x}_{t-1}^{(-)} + \boldsymbol{\delta}\mathbf{x}_{t-1}^{(s)} + \boldsymbol{\beta}\boldsymbol{\mu}_{t-1}. \quad (2.3.12)$$

Dans la spécification (2.3.12), les fonctions $\mathbf{x}_t^{(-)}$ et $\mathbf{x}_t^{(s)}$ sont égales à $\mathbf{x}_t^{(-)} = \mathbf{x}_t I(r_t < 0)$ et $\mathbf{x}_t^{(s)} = \mathbf{x}_t^{1/2} \text{signe}(r_t)$, avec $I(B)$ représentant la variable indicatrice de l'ensemble B . Les matrices $\boldsymbol{\delta}$ et $\boldsymbol{\gamma}$ ont des interprétations dans la littérature financière liées à des considérations d'asymétrie. Les matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ et $\boldsymbol{\delta}$ ont pour dimension $K \times K$. On note que le vecteur $\boldsymbol{\omega}$ a pour dimension $K \times 1$.

La spécification générale de $\boldsymbol{\mu}_t$ peut être écrite de la manière suivante :

$$\boldsymbol{\mu}_t = \boldsymbol{\omega} + \boldsymbol{\alpha}^* \mathbf{x}_t^* + \boldsymbol{\beta}\boldsymbol{\mu}_{t-1},$$

tel que le vecteur $\mathbf{x}_t^* = (\mathbf{x}_t; \mathbf{x}_t^{(-)}; \mathbf{x}_t^{(s)})$ a pour dimension $3K \times 1$ et la matrice $\boldsymbol{\alpha}^* = (\boldsymbol{\alpha}; \boldsymbol{\gamma}; \boldsymbol{\beta})$ a pour dimension $K \times (3K)$. Soit le vecteur $\boldsymbol{\theta} = (\boldsymbol{\omega}; \boldsymbol{\alpha}^*; \boldsymbol{\beta})$. Comme dans le cas univarié des modèles MEM, l'espace des paramètres pour les vMEM doit être restreint afin de s'assurer que $\boldsymbol{\mu}_t \geq \mathbf{0}$. Pour un vecteur \mathbf{x} , la notation $\mathbf{x} \geq \mathbf{0}$ veut dire que chaque composante du vecteur est non-négative.

Pour un processus stochastique admettant une structure vMEM, la condition suffisante de stationnarité est que toutes les valeurs propres de

$$\mathbf{A} = \boldsymbol{\alpha} + \boldsymbol{\beta} + \boldsymbol{\gamma}/2$$

soient plus petites que un en module.

Les conditions suffisantes de non-négativité des composantes de $\boldsymbol{\mu}_t$ reposent sur les résultats déjà obtenus dans le cadre des modèles MEM. Ainsi, $\boldsymbol{\mu}_t \geq \mathbf{0}$, $\forall t$, si les conditions suivantes sont satisfaites pour tout $i, j = 1, \dots, K$.

(1) Pour tout j , $\beta_{ij} \geq 0$, $\alpha_{ij} \geq 0$, $\alpha_{ij} + \gamma_{ij} \geq 0$;

(2) Si $\alpha_{ij} = 0$ alors $\delta_{ij} \geq 0$; Si $\alpha_{ij} + \gamma_{ij} = 0$ alors $\delta_{ij} \leq 0$;

$$(3) \quad \omega_i - \frac{1}{4} \sum_{j=1}^k \delta_{ij}^2 \left[\frac{I(\delta_{ij} < 0)I(\alpha_{ij} > 0)}{\alpha_{ij}} + \frac{I(\delta_{ij} > 0)I(\alpha_{ij} + \gamma_{ij} > 0)}{\alpha_{ij} + \gamma_{ij}} \right] \geq 0.$$

Pour plus de détails, voir Cipollini, Engle et Gallo (2006). Il semble assez évident que ces conditions étaient déjà difficiles à utiliser pour les MEM. Les choses se compliquent substantiellement dans le cas multivarié. À mesure que K devient grand, les conditions précédentes sont très compliquées à vérifier en pratique.

2.3.4. Spécification de ϵ_t

Dans cette section, nous considérons la spécification de la distribution du terme d'erreur ϵ_t pour le modèle vMEM. L'extension naturelle est de considérer :

$$\epsilon_{i,t} | \mathcal{F}_{t-1} \sim \text{gamma}(\phi_i, \phi_i), i = 1, \dots, K.$$

Comme mentionné dans la section 2.1, une possibilité est d'utiliser une distribution gamma multivariée, qui est telle que les distributions marginales sont gamma. Cependant, compte tenu des restrictions sur la structure de dépendance, une autre possibilité est d'utiliser les copules. Dans ce qui suit, on considère l'utilisation de la copule normale avec marges gamma.

2.3.4.1. Copule normale de distributions marginales gamma

Une autre manière valide pour définir la distribution de ϵ_t est de supposer que tous les densités marginales univariées ont comme distribution une $\text{gamma}(\phi_i, \phi_i)$, $i = 1, \dots, K$ et d'utiliser les copules. La copule est l'outil permettant d'extraire la structure de dépendance et le comportement marginal. Plusieurs copules ont été proposées dans des travaux théoriques et appliqués (Embrechts, Lindskog et McNeil (2001) et Bouyé, Durrleman, Nikeghbali, Riboulet et Roncalli (2000)). Les copules, normale ou student, provenant de la famille elliptique, sont flexibles et possèdent de nombreuses propriétés intéressantes comme la résolubilité analytique et la facilité de la simulation.

Dans notre étude, nous considérons la copule normale. Nous supposons que :

$$\epsilon_t \sim \mathcal{N}(\mathbf{R}) - \prod_i^K \text{gamma}(\phi_i, \phi_i).$$

Cette notation veut simplement dire que la structure de dépendance est celle

d'une copule normale, mais que les distributions marginales sont :

$$\text{gamma}(\phi_i, \phi_i), i = 1, \dots, K.$$

La matrice R permet de spécifier la matrice de corrélation.

La variance-covariance conditionnelle de ϵ_t a un élément global approximativement égal à :

$$C(\epsilon_{i,t}, \epsilon_{j,t}) \simeq \frac{R_{ij}}{\sqrt{\Phi_i \Phi_j}}.$$

De plus, les corrélations sont approximativement :

$$\rho(\epsilon_{i,t}, \epsilon_{j,t}) \simeq R_{ij}.$$

Contrairement à la distribution gamma multivariée, la copule normale avec marges gamma permet une structure de covariance et corrélation plus souple. En particulier, des corrélations négatives sont possibles. De plus, les corrélations ne dépendent plus des variances des distributions marginales, et il y a moins de contraintes complexes sur les paramètres de la distribution. De façon intéressante, l'écriture de la fonction de densité est relativement aisée, ce qui rend possible l'écriture de fonctions de vraisemblance. L'estimation dans les modèles vMEM est abordée à la prochaine section.

2.4. ESTIMATION DU MODÈLE VMEM

Plusieurs méthodes sont possibles pour estimer les modèles MEM multivariés. Si nous spécifions la distribution de ϵ_t , nous pouvons utiliser la méthode du maximum de vraisemblance. Si nous ne spécifions pas la distribution de ϵ_t , alors il est possible de considérer des méthodes semi-paramétriques telles que la méthode des moments généralisés (GMM).

Nous commençons par présenter la méthode du maximum de vraisemblance.

2.4.1. Maximum de vraisemblance

Soit ϵ_t tel que leur fonction de densité est donnée par :

$$f(\epsilon_t) = c(F_1(\epsilon_{1,t}), \dots, F_K(\epsilon_{K,t})) \prod_{i=1}^K f_i(\epsilon_{i,t}),$$

avec $f_i(\cdot)$ et $F_i(\cdot)$ qui indiquent respectivement la fonction de densité et la fonction de répartition de la i ème composante de ϵ_t , et $c(\cdot)$ est la fonction de densité de la copule choisie.

Dans notre étude, nous utilisons une copule normale et des marginales dont leurs distributions univariées sont $\text{gamma}(\phi_i, \phi_i)$, $i = 1, \dots, K$.

Pour chaque i , les fonctions de densité et de répartition sont respectivement :

$$f_i(\epsilon_{i,t}) = \frac{\phi_i^{\phi_i}}{\Gamma(\phi_i)} \epsilon_{i,t}^{\phi_i-1} \exp(-\phi_i \epsilon_{i,t}),$$

et

$$F_i(\epsilon_{i,t}) = \Gamma(\phi_i; \phi_i \epsilon_{i,t}),$$

où la fonction $\Gamma(\zeta; x)$ indique la fonction de répartition d'une distribution $\text{gamma}(\zeta, 1)$ calculée à x .

La densité de la copule normale est :

$$c(\mathbf{u}_t) = |\mathbf{R}|^{-1/2} \exp \left[-\frac{1}{2} \mathbf{q}_t^\top (\mathbf{R}^{-1} - \mathbf{I}) \mathbf{q}_t \right],$$

où

$$\mathbf{q}_t = (\Phi^{-1}(F_1(\epsilon_{1,t})); \dots; \Phi^{-1}(F_K(\epsilon_{K,t}))).$$

Dans l'expression précédente, $\Phi^{-1}(x)$ indique la fonction quantile de la distribution normale de moyenne 0 et de variance 1 calculée à x . Le vecteur $\mathbf{u}_t = (F_1(\epsilon_{1,t}), \dots, F_K(\epsilon_{K,t}))^\top$ représente un vecteur de dimension K des fonctions de répartition univariées de $\epsilon_{i,t}$ avec $i = 1, \dots, K$.

Ainsi, la fonction de densité de $\mathbf{x}_t \mid \mathcal{F}_{t-1}$ est la suivante :

$$f(\mathbf{x}_t \mid \mathcal{F}_{t-1}) = c(F_1(x_{1,t}/\mu_{1,t}), \dots, F_K(x_{K,t}/\mu_{K,t})) \prod_{i=1}^K \frac{f_i(x_{i,t}/\mu_{i,t})}{\mu_{i,t}},$$

où la fonction $\frac{f_i(x_{i,t}/\mu_{i,t})}{\mu_{i,t}}$ indique la fonction de densité d'une loi gamma(ϕ_i , $\phi_i/\mu_{i,t}$).

La log-vraisemblance pour un échantillon de taille T se détaille comme suit :

$$\begin{aligned} \mathfrak{l} &= \sum_{t=1}^T \mathfrak{l}_t, \\ &= \sum_{t=1}^T \ln f(\mathbf{x}_t \mid \mathcal{F}_{t-1}), \\ &= \sum_{t=1}^T \ln c(F_1(x_{1,t}/\mu_{1,t}), \dots, F_K(x_{K,t}/\mu_{K,t})), \\ &+ \sum_{t=1}^T \sum_{i=1}^K [\ln f_i(x_{i,t}/\mu_{i,t}) - \ln \mu_{i,t}]. \end{aligned}$$

Puisque la copule est normale, nous pouvons écrire :

$$\begin{aligned} \mathfrak{l}_t &= \frac{1}{2} \ln |\mathbf{R}^{-1}| - \frac{1}{2} \mathbf{q}_t^\top \mathbf{R}^{-1} \mathbf{q}_t + \frac{1}{2} \mathbf{q}_t^\top \mathbf{q}_t \\ &+ \sum_{i=1}^K \left[\phi_i \ln \phi_i - \ln \Gamma(\phi_i) - \ln x_{i,t} + \phi_i \left(\ln x_{i,t} + \phi_i \left(\ln x_{i,t} - \ln \mu_{i,t} - \frac{x_{i,t}}{\mu_{i,t}} \right) \right) \right], \\ &= (\text{Contribution de la copule})_t + (\text{Contribution des marginales})_t. \end{aligned}$$

La contribution des marginales dépend de θ et $\boldsymbol{\phi}$, tandis que la contribution de la copule normale dépend de \mathbf{R} , θ et $\boldsymbol{\phi}$. Ça implique

$$\frac{\partial \mathfrak{l}}{\partial \mathbf{R}^{-1}} = \frac{1}{2} (\mathbf{T}\mathbf{R} - \mathbf{q}^\top \mathbf{q}) = 0 \Rightarrow \hat{\mathbf{R}} = \frac{\mathbf{q}^\top \mathbf{q}}{\mathbf{T}},$$

où la matrice $\mathbf{T} \times \mathbf{K}$ notée \mathbf{q} est telle que $\mathbf{q}^\top = (\mathbf{q}_1, \dots, \mathbf{q}_T)$. Cependant, l'estimateur de \mathbf{R} ci-dessus est obtenu sans imposer aucune contrainte par rapport à sa nature comme une matrice de corrélation (par exemple que les éléments

diagonaux sont un et que R est une matrice définie positive). D'un autre côté, une maximisation directe de la contribution de la copule normale en tenant compte des contraintes que les éléments diagonaux sont un et que R est une matrice définie positive ne donne pas une solution explicite et complique la mise en œuvre empirique.

Un compromis acceptable qui devrait accroître l'efficacité, bien que formellement il ne peut pas être interprété comme un estimateur du maximum de vraisemblance, est de normaliser l'estimateur obtenu ci-dessus, \hat{R} , afin de le transformer en une matrice de corrélation :

$$\tilde{R} = D_Q^{-\frac{1}{2}} Q D_Q^{-\frac{1}{2}},$$

avec

$$Q = \frac{\mathbf{q}^\top \mathbf{q}}{T}, \text{ où } D_Q = \text{diag}(Q_{11}, \dots, Q_{KK}).$$

En utilisant le nouvel estimateur de R , on obtient la log-vraisemblance concentrée suivante :

$$\begin{aligned} \mathfrak{l}_c &= \frac{T}{2} \left[-\ln |\tilde{R}| - \text{trace}(\tilde{R}Q) + \text{trace}(Q) \right] + (\text{contribution des marginales}), \\ &= -\frac{T}{2} \left[\ln |\mathbf{q}^\top \mathbf{q}| - \sum_{i=1}^K \ln(\mathbf{q}_{\cdot i}^\top \mathbf{q}_{\cdot i}) \right] + (\text{contribution des marginales}). \end{aligned}$$

Plus de détails sur cette construction se trouve dans Cipollini, Engle et Gallo (2006). Un aspect important de cette méthode porte sur le calcul des dérivées de la fonction de vraisemblance. Nous exposons les quantités utiles dans la prochaine section.

2.4.1.1. Les dérivées de la log vraisemblance concentrée

Les paramètres à estimer sont $\boldsymbol{\phi}$ et $\boldsymbol{\theta}$, où le vecteur $\boldsymbol{\phi}$ est un vecteur qui représente les paramètres de la distribution gamma des ϵ_t . Le vecteur $\boldsymbol{\theta}$ se retrouve dans l'équation de μ_t . La dérivée de la log vraisemblance concentrée sera donc par rapport à $\boldsymbol{\phi}$ et $\boldsymbol{\theta}$. On trouve :

$$\nabla_{\theta} \mathcal{l}_c = \sum_{t=1}^T \nabla_{\theta} \boldsymbol{\mu}_t^{\top} \mathbf{a}_t, \quad (2.4.1)$$

où

$$\mathbf{a}_t = (\mathbf{a}_{1,t}, \dots, \mathbf{a}_{K,t})^{\top}, \quad (2.4.2)$$

et le vecteur \mathbf{a}_t a pour composantes :

$$\mathbf{a}_{i,t} = \frac{f_i^{(1)}\left(\frac{x_{i,t}}{\mu_{i,t}}; \phi_i\right) \mathbf{b}_{i,t} + f_i^{(2)}\left(\frac{x_{i,t}}{\mu_{i,t}}; \phi_i\right)}{\mu_{i,t}}, \quad (2.4.3)$$

$i = 1, \dots, K$. Les fonctions $f_i^{(1)}$ et $f_i^{(2)}$ satisfont :

$$f_i^{(1)}(x_{i,t}/\mu_{i,t}; \phi_i) = \text{dgamma}\left(\frac{x_{i,t}}{\mu_{i,t}}; \phi_i + 1; \phi_i\right), \quad (2.4.4)$$

$$f_i^{(2)}(x_{i,t}/\mu_{i,t}; \phi_i) = \phi_i \left(\frac{x_{i,t}}{\mu_{i,t}} - 1\right), \quad (2.4.5)$$

$$\mathbf{b}_{i,t} = -\frac{(\mathbf{C}\mathbf{q}_t)_i}{\phi(q_{i,t})}. \quad (2.4.6)$$

Dans la formule (2.4.6), la matrice \mathbf{C} est :

$$\mathbf{C} = \mathbf{Q}^{-1} \mathbf{D}_Q^{1/2} \mathbf{Q} \mathbf{D}_Q^{1/2} \mathbf{Q}^{-1} - \mathbf{Q}^{-1} + \mathbf{I}_K - \tilde{\mathbf{R}}^{-1} + \mathbf{D}_Q^{-1} - \mathbf{D}_Q^{-1/2} \text{diag}(\mathbf{Q}^{-1} \mathbf{D}_Q^{1/2} \mathbf{Q}).$$

Ainsi, on trouve l'expression suivante :

$$\nabla_{\phi_i} \mathcal{l}_c = \sum_{t=1}^T \left[-\nabla_{\phi_i} F_i\left(\frac{x_{i,t}}{\mu_{i,t}}; \phi_i\right) \mathbf{b}_{i,t} + \ln(\phi_i) - \psi(\phi_i) + \ln\left(\frac{x_{i,t}}{\mu_{i,t}}\right) - \frac{x_{i,t}}{\mu_{i,t}} + 1 \right]. \quad (2.4.7)$$

Les dérivées de $\nabla_{\phi_i} F_i(x_{i,t}/\mu_{i,t}; \phi_i)$ doivent être calculées numériquement. Cependant, estimer les paramètres à l'aide du maximum de vraisemblance concentrée n'est pas facile, particulièrement si le nombre de paramètres K est grand.

Parfois nous nous intéressons à analyser la dynamique de la moyenne conditionnelle μ_t sans spécifier la distribution multivariée pour le terme d'erreur.

Pour cela, une approche alternative semi-paramétrique est possible. Cette approche est semi-paramétrique dans le sens que contrairement à l'approche imposant une loi multivariée, il n'est pas nécessaire de spécifier une distribution multivariée pour le terme d'erreur. Voir Cipollini, Engle et Gallo (2012). Les paramètres d'intérêt sont donc les paramètres du vecteur θ de dimension p qui mesurent la dynamique de μ_t . Elle est basée sur les moments d'ordre un et deux de la variable $\mathbf{x}_t | \mathcal{F}_{t-1}$ qui sont tels que :

$$E(\mathbf{x}_t | \mathcal{F}_{t-1}) = \boldsymbol{\mu}_t, \quad (2.4.8)$$

$$\text{Var}(\mathbf{x}_t | \mathcal{F}_{t-1}) = \boldsymbol{\mu}_t \boldsymbol{\mu}_t^\top \odot \boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\mu}_t) \boldsymbol{\Sigma} \text{diag}(\boldsymbol{\mu}_t). \quad (2.4.9)$$

Nous allons présenter dans la prochaine section deux méthodes d'estimation qui sont dans cette catégorie : la méthode des moments généralisés (GMM) et aussi la méthode des erreurs standard associées aux paramètres (WNLS).

2.4.2. Méthode des moments généralisés (GMM)

La méthode des moments généralisés (GMM) a été introduite pour estimer des paramètres définis par des conditions de moments. Le principe de la méthode des moments généralisés consiste à choisir une valeur de θ permettant de rendre la moyenne empirique la plus proche possible de zéro. Nous commençons par décrire cette méthode.

Soit la variable \mathbf{u}_t telle que :

$$\mathbf{u}_t = \mathbf{x}_t \oslash \boldsymbol{\mu}_t - \mathbf{1}, \quad (2.4.10)$$

où le signe ' \oslash ' représente la division élément par élément. Plus particulièrement, pour deux matrices $A = (a_{ij})$ et $B = (b_{ij})$ de même dimensions telles que $b_{ij} > 0$, on définit $A \oslash B = C = (c_{ij})$ où $c_{ij} = a_{ij}/b_{ij}$. D'après les équations (2.4.8) et (2.4.9), la variable \mathbf{u}_t est une différence de martingales conditionnelle homoscédastique telle que :

$$E(\mathbf{u}_t | \mathcal{F}_{t-1}) = \mathbf{0}, \quad (2.4.11)$$

$$\text{Var}(\mathbf{u}_t | \mathcal{F}_{t-1}) = \boldsymbol{\Sigma}. \quad (2.4.12)$$

Le moment conditionnel d'ordre un peut être une clé pour l'estimation (voir Wooldridge (1994)), mais il ne suffit pas pour estimer k paramètres du vecteur θ . Cependant, une matrice \mathbf{G}_t qui dépend de \mathcal{F}_{t-1} et non-corrélée avec \mathbf{u}_t , donne le résultat suivant :

$$E(\mathbf{G}_t \mathbf{u}_t | \mathcal{F}_{t-1}) = \mathbf{0}. \quad (2.4.13)$$

Dans la littérature sur les méthodes de type GMM la matrice \mathbf{G}_t est appelée instrument.

La structure des modèles vMEM permet des simplifications considérables, liée en particulier à \mathbf{u}_t étant une différence de martingales. Cette caractéristique implique que $\mathbf{G}_t \mathbf{u}_t$ est une différence de martingales (voir Wooldridge (1994)), et le théorème de la double espérance donne donc $E(\mathbf{G}_t \mathbf{u}_t) = \mathbf{0}$. Selon la méthode des moments généralisés, on doit résoudre en utilisant les versions empiriques le système :

$$\frac{1}{T} \sum_{t=1}^T \mathbf{G}_t \mathbf{u}_t = \mathbf{0}. \quad (2.4.14)$$

L'équation (2.4.14) nous donne L moments conditionnels à résoudre. Si $L = k$, nous avons autant d'équations que la dimension k où le nombre k représente la dimension du vecteur θ .

Selon Wooldridge (1994), la matrice des variances et covariances est la suivante :

$$\text{Avar}(\hat{\theta}_T) = \frac{1}{T} (\mathbf{S}^\top \mathbf{V}^{-1} \mathbf{S})^{-1}, \quad (2.4.15)$$

où les matrices \mathbf{S} et \mathbf{V} sont :

$$S = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E(\nabla_{\theta}^{\top} \mathbf{G}_t \mathbf{u}_t), \quad (2.4.16)$$

$$V = \lim_{T \rightarrow \infty} \frac{1}{T} \text{Var} \left(\sum_{t=1}^T \mathbf{G}_t \mathbf{u}_t \right), \quad (2.4.17)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E \{ (\mathbf{G}_t \mathbf{u}_t) (\mathbf{G}_t \mathbf{u}_t)^{\top} \}. \quad (2.4.18)$$

Puisque la fonction $\mathbf{G}_t \mathbf{u}_t$ est une différence de martingales, alors on peut passer de (2.4.17) à (2.4.18).

La structure de différence de martingales de \mathbf{u}_t donne aussi une simple formulation pour un choix efficace de l'instrument \mathbf{G}_t , où le terme *efficace* signifie produire la plus petite variance asymptotique entre les estimateurs de GMM posée par les fonctions $\mathbf{G}_t \mathbf{u}_t$ comme en (2.4.14). Alors, le choix efficace de l'instrument \mathbf{G}_t est le suivant :

$$\mathbf{G}_t^* = -E(\nabla_{\theta} \mathbf{u}_t^{\top} | \mathcal{F}_{t-1}) \text{Var}(\mathbf{u}_t | \mathcal{F}_{t-1})^{-1}. \quad (2.4.19)$$

En utilisant l'équation (2.4.10), on trouve que

$$\nabla_{\theta} \mathbf{u}_t^{\top} = -\nabla_{\theta} \boldsymbol{\mu}_t^{\top} \text{diag}(\boldsymbol{\mu}_t)^{-1} \text{diag}(\mathbf{u}_t + \mathbf{1}). \quad (2.4.20)$$

Et donc l'instrument \mathbf{G}_t^* sera égal à :

$$\mathbf{G}_t^* = -\nabla_{\theta} \boldsymbol{\mu}_t^{\top} \text{diag}(\boldsymbol{\mu}_t)^{-1} \Sigma^{-1}. \quad (2.4.21)$$

Retournons maintenant à l'équation (2.4.14) et remplaçant l'instrument \mathbf{G}_t par l'instrument \mathbf{G}_t^* . Donc on peut estimer le vecteur θ en résolvant le système non-linéaire suivant :

$$\frac{1}{T} \sum_{t=1}^T \nabla_{\theta} \boldsymbol{\mu}_t^{\top} \text{diag}(\boldsymbol{\mu}_t)^{-1} \Sigma^{-1} \mathbf{u}_t = \mathbf{0}. \quad (2.4.22)$$

La matrice de variance asymptotique est :

$$\text{Avar}(\hat{\theta}_T) = \frac{1}{T} \left[\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E(\nabla_{\theta} \mu_t^{\top} \text{diag}(\mu_t) \Sigma \text{diag}(\mu_t))^{-1} \nabla_{\theta^{\top}} \mu_t \right]^{-1}. \quad (2.4.23)$$

Remarquons que l'équation (2.4.22) dépend de Σ dont les éléments sont inconnus. Pour estimer les paramètres de la moyenne conditionnelle, nous devons considérer les deux étapes suivantes :

Dans la première étape, on considère $\Sigma = I_K$ telle que la matrice I_K est la matrice identité d'ordre K , où le nombre K représente le nombre de variables étudiées. Nous résolvons ensuite le système non-linéaire suivant :

$$\frac{1}{T} \sum_{t=1}^T \nabla_{\theta} \mu_t^{\top} \text{diag}(\mu_t)^{-1} \mathbf{u}_t = \mathbf{0}. \quad (2.4.24)$$

Nous obtenons alors un estimateur $\hat{\theta}$ pour le vecteur θ .

Dans la deuxième étape, nous utilisons le vecteur $\hat{\theta}$ estimé afin de calculer tout d'abord \mathbf{u}_t (2.4.10) et ensuite pour déterminer $\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{u}}_t \hat{\mathbf{u}}_t^{\top}$. Enfin, nous remplaçons dans l'équation 2.4.22 la matrice $\hat{\Sigma}$ telle que le système non-linéaire à résoudre est le suivant :

$$\frac{1}{T} \sum_{t=1}^T \nabla_{\theta} \mu_t^{\top} \text{diag}(\mu_t)^{-1} \hat{\Sigma}^{-1} \mathbf{u}_t = \mathbf{0}. \quad (2.4.25)$$

Les deux étapes sont répétées de manière itérative jusqu'à convergence. Donc la méthode semi-paramétrique GMM consiste à estimer le vecteur θ en résolvant l'équation (2.4.22) par les deux étapes décrites ci-dessus. Toutefois, la même condition peut également être obtenue au moyen d'autres méthodes d'inférence. Une de ces méthodes est la méthode des erreurs standard associées aux paramètres (WNLS).

2.4.3. Méthode des erreurs standard associées aux paramètres (WNLS)

Comme la méthode des moments généralisés, la méthode des erreurs standard associées aux paramètres peut être utilisée pour les modèles vectoriels et multiplicatifs avec erreurs non-négatives afin d'estimer les paramètres de la moyenne conditionnelle μ_t sans spécifier la distribution du terme d'erreur. Cette méthode, notée WNLS, est détaillée dans l'ouvrage de Wooldridge (1994). Elle prend en considération les moments conditionnels d'espérance (2.4.8) ainsi

que la variance (2.4.9) de \mathbf{x}_t .

Selon Wooldridge (1994), cette méthode consiste à minimiser le système non-linéaire suivant :

$$\min_{\theta} \frac{1}{2T} \sum_{t=1}^T (\mathbf{x}_t - \boldsymbol{\mu}_t)^\top \widehat{\mathbf{W}}_t^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_t), \quad (2.4.26)$$

où la matrice \mathbf{W}_t est définie positive, symétrique et dépend de \mathcal{F}_{t-1} , le vecteur θ est un vecteur qui contient tous les paramètres de la moyenne conditionnelle à estimer.

Pour avoir un estimateur optimal du vecteur θ , ceci nécessite de trouver une valeur convenable et efficace pour la matrice \mathbf{W}_t . Selon Wooldridge (1994), la matrice \mathbf{W}_t est égale à la variance conditionnelle (2.3.8) de \mathbf{x}_t .

En remplaçant \mathbf{W}_t par $\text{diag}(\boldsymbol{\mu}_t)\Sigma\text{diag}(\boldsymbol{\mu}_t)$, la nouvelle équation à minimiser est la suivante :

$$\min_{\theta} \frac{1}{2T} \sum_{t=1}^T (\mathbf{x}_t - \boldsymbol{\mu}_t)^\top (\text{diag}(\boldsymbol{\mu}_t)\Sigma\text{diag}(\boldsymbol{\mu}_t))^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_t), \quad (2.4.27)$$

Or l'équation (2.4.27) dépend de la matrice inconnue Σ . Alors, il est nécessaire tout d'abord d'estimer la matrice Σ et ensuite l'utiliser pour estimer l'équation (2.4.27). Afin de faire cette estimation, nous devons considérer deux étapes.

Dans la première étape, on pose $\mathbf{W}_t = \text{diag}(\boldsymbol{\mu}_t)\mathbf{I}_K\text{diag}(\boldsymbol{\mu}_t)$ tel que $\Sigma = \mathbf{I}_K$ où \mathbf{I}_K est la matrice identité d'ordre K , et le nombre K représente le nombre de variables étudiées. Puis nous minimisons l'équation suivante :

$$\min_{\theta} \frac{1}{2T} \sum_{t=1}^T (\mathbf{x}_t - \boldsymbol{\mu}_t)^\top (\text{diag}(\boldsymbol{\mu}_t)\mathbf{I}_K\text{diag}(\boldsymbol{\mu}_t))^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_t). \quad (2.4.28)$$

Dans la deuxième étape, nous utilisons le vecteur $\hat{\theta}$ estimé dans la première étape pour calculer la matrice inconnue Σ . Une fois que la matrice Σ est calculée, nous minimisons le système non-linéaire suivant :

$$\min_{\theta} \frac{1}{2T} \sum_{t=1}^T (\mathbf{x}_t - \boldsymbol{\mu}_t)^\top (\text{diag}(\boldsymbol{\mu}_t) \hat{\Sigma} \text{diag}(\boldsymbol{\mu}_t))^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_t). \quad (2.4.29)$$

Dans cette section, nous avons exposé la méthode du maximum de vraisemblance concentrée, la méthode des moments généralisés (GMM) et la méthode des erreurs standard associées aux paramètres (WNLS). Toutes ces méthodes d'estimation ont été programmées dans le langage R, et le code principal est fourni en Annexe. Dans le prochain chapitre, nous allons étudier empiriquement les trois méthodes d'estimation à l'aide d'une brève étude de simulations Monte Carlo.

Chapitre 3

SIMULATIONS

Dans les chapitres précédents, nous avons introduit les modèles MEM et plus particulièrement les modèles multivariés vMEM. Nous avons discuté en détails d'un aspect particulièrement important, à savoir comment spécifier la loi du terme d'erreur, qui est une distribution multivariée dont toutes les composantes sont non-négatives. Nous avons présenté plusieurs méthodes d'estimation dans ce contexte. Une contribution importante du mémoire est la comparaison empirique de ces méthodes d'estimation. Ainsi, les biais des estimateurs, en plus de l'efficacité des estimateurs, sont étudiés dans ce chapitre. Une discussion importante porte sur la convergence des méthodes d'estimation. En fait, c'est vraiment à ce niveau que nous avons rencontré le plus de difficultés, car dans plusieurs situations les algorithmes éprouvaient des difficultés de convergence. Ceci semble en accord avec les articles discutant des propriétés empiriques de ces méthodes.

3.1. DESCRIPTION DE LA SIMULATION

Plusieurs processus ont été considérés, s'exprimant tous comme des modèles vMEM(1,1) de la forme suivante :

$$\mu_t = \omega + \alpha x_{t-1} + \beta \mu_{t-1},$$

où ω a pour dimension 3×1 , et α et β sont des matrices de paramètres 3×3 . Les matrices de paramètres ont été choisies en s'inspirant d'autres études empiriques. Voir Cipollini et Gallo (2010) et Cipollini, Engle et Gallo (2012).

Nous devons insister sur le fait que nous avons considéré des modèles trivariés, donc avec $K = 3$, ce qui a entraîné des complications numériques.

Les tailles échantillonales nécessaires étaient relativement grandes, à savoir $T = 1000$ et $T = 5000$. Compte tenu que les modèles trivariés reposent sur un maximum possible de 22 paramètres indépendants, les tailles considérées sont raisonnables. Rappelons qu'une modélisation GARCH univariée repose souvent sur un grand nombre d'observations. Une modélisation vMEM à trois dimensions nécessite également des tailles échantillonales assez grandes.

Concernant la loi du terme d'erreur, nous avons considéré dans tous les cas une copule gaussienne avec des marges diverses. Pour les distributions marginales, nous avons considéré des lois gamma avec plusieurs choix de paramètres et des lois exponentielles. La copule gaussienne repose sur la matrice R telle que :

$$R = \begin{pmatrix} 1 & 0.37 & 0.21 \\ 0.37 & 1 & 0.18 \\ 0.21 & 0.18 & 1 \end{pmatrix}.$$

Dans cette situation, les composantes du vecteur ϵ_t sont dépendantes sans être fortement dépendantes. Nous avons essayé d'autres choix de matrice R sans qu'il y ait des changements majeurs.

Les processus vMEM(1,1) et les lois marginales sont décrits comme suit :

$$\begin{pmatrix} av_t \\ bv_t \\ cv_t \end{pmatrix} = \left[\begin{pmatrix} 0.0206 \\ 0.0772 \\ 0.1172 \end{pmatrix} + \begin{pmatrix} 0 & 0.1564 & 0 \\ 0 & 0.1283 & 0 \\ 0 & 0 & 0.2418 \end{pmatrix} \begin{pmatrix} av_{t-1} \\ bv_{t-1} \\ cv_{t-1} \end{pmatrix} + \begin{pmatrix} 0.7900 & & \\ & 0.7795 & \\ & & 0.6450 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{av_t} \\ \mu_{t-1}^{bv_t} \\ \mu_{t-1}^{cv_t} \end{pmatrix} \right] \odot \begin{pmatrix} \epsilon_{av,t} \\ \epsilon_{bv,t} \\ \epsilon_{cv,t} \end{pmatrix}; \quad (3.1.1)$$

$$\begin{pmatrix} \epsilon_{av,t} \\ \epsilon_{bv,t} \\ \epsilon_{cv,t} \end{pmatrix} \sim \begin{pmatrix} \text{gamma}(4,0,4,0) \\ \text{gamma}(5,5,5,5) \\ \text{gamma}(3,1,3,1) \end{pmatrix};$$

$$\begin{pmatrix} dv_t \\ ev_t \\ fv_t \end{pmatrix} = \left[\begin{pmatrix} 0.0206 \\ 0.0772 \\ 0.1172 \end{pmatrix} + \begin{pmatrix} 0 & 0.1564 & 0 \\ 0 & 0.1283 & 0 \\ 0 & 0 & 0.2418 \end{pmatrix} \begin{pmatrix} dv_{t-1} \\ ev_{t-1} \\ fv_{t-1} \end{pmatrix} + \begin{pmatrix} 0.7900 & & \\ & 0.7795 & \\ & & 0.6450 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{dv_t} \\ \mu_{t-1}^{ev_t} \\ \mu_{t-1}^{fv_t} \end{pmatrix} \right] \odot \begin{pmatrix} \epsilon_{dv,t} \\ \epsilon_{ev,t} \\ \epsilon_{fv,t} \end{pmatrix}; \quad (3.1.2)$$

$$\begin{pmatrix} \epsilon_{dv,t} \\ \epsilon_{ev,t} \\ \epsilon_{fv,t} \end{pmatrix} \sim \begin{pmatrix} \text{gamma}(27.9853,27.9853) \\ \text{gamma}(28.6715,28.6715) \\ \text{gamma}(36.3381,36.3381) \end{pmatrix};$$

$$\begin{pmatrix} gv_t \\ hv_t \\ iv_t \end{pmatrix} = \left[\begin{pmatrix} 0.0206 \\ 0.0772 \\ 0.1172 \end{pmatrix} + \begin{pmatrix} 0 & 0.1564 & 0 \\ 0 & 0.1283 & 0 \\ 0 & 0 & 0.2418 \end{pmatrix} \begin{pmatrix} gv_{t-1} \\ hv_{t-1} \\ iv_{t-1} \end{pmatrix} + \begin{pmatrix} 0.7900 & & \\ & 0.7795 & \\ & & 0.6450 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{gv_t} \\ \mu_{t-1}^{hv_t} \\ \mu_{t-1}^{iv_t} \end{pmatrix} \right] \odot \begin{pmatrix} \epsilon_{gv,t} \\ \epsilon_{hv,t} \\ \epsilon_{iv,t} \end{pmatrix}; \quad (3.1.3)$$

$$\begin{pmatrix} \epsilon_{gv,t} \\ \epsilon_{hv,t} \\ \epsilon_{iv,t} \end{pmatrix} \sim \begin{pmatrix} \text{exponentielle}(1) \\ \text{exponentielle}(1) \\ \text{exponentielle}(1) \end{pmatrix};$$

$$\begin{pmatrix} jv_t \\ kv_t \\ lv_t \end{pmatrix} = \left[\begin{pmatrix} 0.0206 \\ 0.0406 \\ 0.0982 \end{pmatrix} + \begin{pmatrix} 0.1961 & 0.0529 & 0 \\ 0.1864 & 0.0694 & 0 \\ 0.0775 & 0 & 0.1770 \end{pmatrix} \begin{pmatrix} jv_{t-1} \\ kv_{t-1} \\ lv_{t-1} \end{pmatrix} + \begin{pmatrix} 0.6675 & & \\ & 0.6715 & \\ & & 0.6618 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{jv_t} \\ \mu_{t-1}^{kv_t} \\ \mu_{t-1}^{lv_t} \end{pmatrix} \right] \odot \begin{pmatrix} \epsilon_{jv,t} \\ \epsilon_{kv,t} \\ \epsilon_{lv,t} \end{pmatrix}; \quad (3.1.4)$$

$$\begin{pmatrix} \epsilon_{jv,t} \\ \epsilon_{kv,t} \\ \epsilon_{lv,t} \end{pmatrix} \sim \begin{pmatrix} \text{gamma}(27.0821, 27.0821) \\ \text{gamma}(28.1324, 28.1324) \\ \text{gamma}(35.8882, 35.8882) \end{pmatrix};$$

$$\begin{pmatrix} mv_t \\ nv_t \\ ov_t \end{pmatrix} = \left[\begin{pmatrix} 0.0206 \\ 0.0406 \\ 0.0982 \end{pmatrix} + \begin{pmatrix} 0.1961 & 0.0529 & 0 \\ 0.1864 & 0.0694 & 0 \\ 0.0775 & 0 & 0.1770 \end{pmatrix} \begin{pmatrix} mv_{t-1} \\ nv_{t-1} \\ ov_{t-1} \end{pmatrix} + \begin{pmatrix} 0.6675 & & \\ & 0.6715 & \\ & & 0.6618 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{mv_t} \\ \mu_{t-1}^{nv_t} \\ \mu_{t-1}^{ov_t} \end{pmatrix} \right] \odot \begin{pmatrix} \epsilon_{mv,t} \\ \epsilon_{nv,t} \\ \epsilon_{ov,t} \end{pmatrix}; \quad (3.1.5)$$

$$\begin{pmatrix} \epsilon_{mv,t} \\ \epsilon_{nv,t} \\ \epsilon_{ov,t} \end{pmatrix} \sim \begin{pmatrix} \text{gamma}(24.50, 24.50) \\ \text{gamma}(35.63, 35.63) \\ \text{gamma}(12.98, 12.98) \end{pmatrix};$$

$$\begin{pmatrix} pv_t \\ qv_t \\ rv_t \end{pmatrix} = \left[\begin{pmatrix} 0.0206 \\ 0.0406 \\ 0.0982 \end{pmatrix} + \begin{pmatrix} 0.1961 & 0.0529 & 0 \\ 0.1864 & 0.0694 & 0 \\ 0.0775 & 0 & 0.1770 \end{pmatrix} \begin{pmatrix} pv_{t-1} \\ qv_{t-1} \\ rv_{t-1} \end{pmatrix} + \begin{pmatrix} 0.6675 & & \\ & 0.6715 & \\ & & 0.6618 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{pv_t} \\ \mu_{t-1}^{qv_t} \\ \mu_{t-1}^{rv_t} \end{pmatrix} \right] \odot \begin{pmatrix} \epsilon_{pv,t} \\ \epsilon_{qv,t} \\ \epsilon_{rv,t} \end{pmatrix}; \quad (3.1.6)$$

$$\begin{pmatrix} \epsilon_{pv,t} \\ \epsilon_{qv,t} \\ \epsilon_{rv,t} \end{pmatrix} \sim \begin{pmatrix} \text{exponentielle}(1) \\ \text{exponentielle}(1) \\ \text{exponentielle}(1) \end{pmatrix};$$

$$\begin{pmatrix} sv_t \\ tv_t \\ uv_t \end{pmatrix} = \left[\begin{pmatrix} 0.020 \\ 0.077 \\ 0.120 \end{pmatrix} + \begin{pmatrix} 0.61 & & \\ & 0.54 & \\ & & 0.4 \end{pmatrix} \begin{pmatrix} sv_{t-1} \\ tv_{t-1} \\ uv_{t-1} \end{pmatrix} + \begin{pmatrix} 0.20 & & \\ & 0.10 & \\ & & 0.25 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{sv_t} \\ \mu_{t-1}^{tv_t} \\ \mu_{t-1}^{uv_t} \end{pmatrix} \right] \odot \begin{pmatrix} \epsilon_{sv,t} \\ \epsilon_{tv,t} \\ \epsilon_{uv,t} \end{pmatrix}; \quad (3.1.7)$$

$$\begin{pmatrix} \epsilon_{sv,t} \\ \epsilon_{tv,t} \\ \epsilon_{uv,t} \end{pmatrix} \sim \begin{pmatrix} \text{gamma}(27.9853, 27.9853) \\ \text{gamma}(28.6715, 28.6715) \\ \text{gamma}(36.3381, 36.3381) \end{pmatrix}.$$

Les processus (3.1.1), (3.1.2) et (3.1.3) sont des modèles vMEM(1,1) qui ont la même structure au niveau de la moyenne conditionnelle μ_t . Ces modèles se distinguent par le terme d'erreur. Ainsi les lois marginales sont différentes.

Les variances diffèrent ainsi que les corrélations entre les composantes. Voir la sous-section 2.3.4.1.

Les processus (3.1.4), (3.1.5) et (3.1.6) sont des modèles vMEM(1,1) assez similaires aux modèles (3.1.1), (3.1.2) et (3.1.3), sauf que le nombre de paramètres non-nuls à estimer est plus important.

Finalement, le processus (3.1.7) est un modèle vMEM(1,1) tel que la diagonale de la matrice β a des coefficients plus petits que 0,5. Comparativement aux autres modèles, moins de persistance est donc considérée.

Afin de simuler les modèles, nous avons dû déterminer les moyennes initiales μ_0 et μ_1 . Elles satisfont :

$$\mu_0 = (I - \alpha - \beta)^{-1} \omega, \quad (3.1.8)$$

$$\mu_1 = \omega + (\alpha + \beta)\mu_0. \quad (3.1.9)$$

Par la suite, nous avons calculé

$$x_1 = \text{diag}(\mu_1)\epsilon_1,$$

pour ϵ_1 un vecteur aléatoire selon la loi considérée. Par la suite, nous avons calculé récursivement :

$$\mu_t = \omega + \alpha x_{t-1} + \beta \mu_{t-1},$$

et

$$x_t = \mu_t \epsilon_t.$$

Pour chaque réalisation, nous avons simulé $T_1 = 2T + 1$ observations, et nous avons conservé les T dernières observations, pour $T \in \{1000, 5000\}$.

Les trois méthodes d'estimation exposées dans le chapitre 2 ont été programmées : la méthode pleinement paramétrique du maximum de vraisemblance, où tous les paramètres sont estimés, incluant les paramètres dans la loi du terme d'erreur ; les méthodes semi-paramétriques GMM ainsi que WNLS ont aussi été incluses dans notre étude.

Toute la programmation a été réalisée avec le logiciel R. Tous les codes sont en Annexe. La prochaine section fait état de quelques constatations liées à des problèmes de convergence.

3.1.1. Problèmes numériques

De manière générale, l'estimation des modèles vMEM était particulièrement compliquée. La méthode du maximum de vraisemblance reposait sur le calcul des dérivées premières, et la méthode n'était pas très stable. Cipollini et Gallo (2010, p. 2485) abordent ces questions, et eux-mêmes ont dû combiner des approches numériques afin d'obtenir la convergence. Cette constatation n'était pas très pratique dans un contexte de simulation. Basée sur une seule réalisation, il était possible d'ajuster de diverses façons des estimateurs afin de trouver l'optimum. Cependant, automatiser cela pour un grand nombre de simulations était particulièrement difficile. Pour le maximum de vraisemblance, nous avons utilisé l'algorithme BHHH (voir Berndt, Hall, Hall et Hausman (1974)). Nous avons rencontré des messages d'erreur suggérant que l'algorithme errait à trouver un optimum (local). La forme de la vraisemblance semblait particulièrement plate à certains endroits.

Pour la méthode GMM, un système non-linéaire devait être résolu. Cependant, compte tenu des restrictions sur l'espace paramétrique, les calculs des matrices inverses $\text{diag}(\boldsymbol{\mu}_t)$ et $\text{diag}(\boldsymbol{\mu}_t)\boldsymbol{\Sigma}^{-1}$ n'étaient pas toujours possible. Il semblait que la matrice $\text{diag}(\boldsymbol{\mu}_t)$ ou $\boldsymbol{\Sigma}$ convergeait vers des matrices singulières. Ce problème était plus prononcé pour $T = 1000$ que $T = 5000$, ce qui était rassurant (mais ne résolvait pas toujours la convergence pour $T = 1000$). Concernant la méthode WNLS, il n'y avait pas beaucoup de problème. Cependant, pour $T = 1000$, il était parfois difficile de calculer la matrice inverse $(\text{diag}(\boldsymbol{\mu}_t)\boldsymbol{\Sigma}\text{diag}(\boldsymbol{\mu}_t))$.

Le tableau 3.1 ci-dessous présente le nombre de cas où la convergence a été obtenue, basé sur 1000 simulations.

Puisque nous réalisons des simulations, nous connaissons les vraies valeurs des modèles que l'on a pu utiliser comme valeurs initiales. Bien entendu, en pratique, le choix des valeurs initiales peut être problématique.

D'après le tableau 3.1, nous remarquons l'importance d'augmenter le nombre

TABLE 3.1. Nombre de cas sur 1000 où il y a eu convergence pour les sept modèles de 1000 réalisations par rapport au nombre d'observations .

Nombre effectif de Réalisations pour les trois méthodes						
Modèles	T = 1000			T = 5000		
	BHHH	GMM	WNLS	BHHH	GMM	WNLS
Modèle 1	198	613	1000	598	987	1000
Modèle 2	289	288	1000	706	981	1000
Modèle 3	226	433	992	632	661	991
Modèle 4	171	106	1000	559	961	1000
Modèle 5	160	173	1000	412	996	1000
Modèle 6	99	524	997	478	966	1000
Modèle 7	494	327	1000	905	883	1000

d'observations. Par exemple, pour le modèle 4, le nombre de réalisations sans erreur sur 1000 est 106 pour $T = 1000$ et 961 pour $T = 5000$. Plus T augmente, plus le nombre de messages d'erreurs a diminué dans nos simulations.

3.2. PRÉSENTATIONS DES RÉSULTATS

Nous commençons par analyser les résultats des estimations pour les processus simulés. Les résultats des trois premiers modèles (3.1.1), (3.1.2) et (3.1.3) se trouvent dans les tableaux 3.2, 3.3 et 3.4. Pour tous les processus, nous avons calculé le biais, l'erreur quadratique moyenne (EQM), la variance échantillonnale (VÉ) et la moyenne des estimateurs de variance (MVE).

D'après les résultats des simulations, nous trouvons que si T augmente, le biais des paramètres de la moyenne conditionnelle et la variance échantillonnale diminuent pour les trois méthodes d'estimation et ainsi automatiquement l'erreur quadratique moyenne MSE diminue, car elle s'exprime en fonction du biais de l'estimateur et de sa variance. Aussi les résultats nous montrent que plus les ϕ_i sont petits, plus le biais et la variance échantillonnale et la moyenne des estimateurs de variance des ϕ_i sont réduits. En comparant les méthodes, au niveau des paramètres de la moyenne conditionnelle et pour chaque T , nous trouvons que la méthode paramétrique a le plus petit biais parmi les trois méthodes et que la méthode GMM a la variance échantillonnale la plus petite parmi les trois méthodes. De plus, plus les ϕ_i sont élevés pour le modèle (3.1.2), plus les méthodes GMM et WNLS ont un biais et une variance échantillonnale plus petites pour chaque T . Tandis que pour la méthode BHHH, cela est valide justement pour la variance échantillonnale. Passons maintenant au processus vMEM (3.1.4), (3.1.5) et (3.1.6).

Les résultats des processus vMEM (3.1.4), (3.1.5) et (3.1.6) se trouvent dans les tableaux 3.5, 3.6 et 3.7. D'après les résultats des simulations pour les modèles (3.1.4), (3.1.5) et (3.1.6), nous trouvons aussi que si T augmente, le biais des paramètres de la moyenne conditionnelle, la variance échantillonnale et l'erreur quadratique moyenne diminuent pour les trois méthodes d'estimation. Parmi les trois méthodes, pour chaque T , nous trouvons que la méthode GMM est la méthode qui donne la variance échantillonnale et la moyenne des estimateurs de variance les plus petites. Pour les processus vMEM (3.1.4), (3.1.5) et (3.1.6), il n'existe pas une méthode précise qui a le plus petit biais pour tous les paramètres de la moyenne conditionnelle. Ils sont divisés entre la méthode BHHH et GMM. Donc la méthode paramétrique n'est pas toujours la meilleure méthode à utiliser pour estimer un processus vMEM. Concernant les valeurs de ϕ_i , nous trouvons que plus les ϕ_i sont petits, plus le biais, la variance échantillonnale et la moyenne des estimateurs de variance des paramètres ϕ_i sont réduits.

Pour le dernier modèle simulé (3.1.7), les paramètres β sont très petits (inférieur à 0,5). Nous remarquons que si les valeurs du paramètre β sont petites, le nombre de cas où il y a convergence de la méthode paramétrique est proche de 50% pour $T = 1000$, et est égale à 90,5% pour $T = 5000$ (voir tableau 3.1).

À partir des résultats, nous pouvons dire que le plus remarquable est que si T augmente le nombre de situations où il y a convergence augmente, tandis que le biais, l'erreur quadratique moyenne, la variance échantillonnale et la moyenne des estimateurs de variance diminuent. De plus, pour chaque T , au niveau des paramètres de la moyenne conditionnelle, la méthode GMM donne les plus petites variances échantillonnales et moyennes des estimateurs de variances parmi les trois méthodes, au moins dans nos études empiriques.

TABLE 3.2. Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 1 défini par (3.1.1) pour $T = 1000$ et 5000.

Val.ini	Modèle 1; T = 1000														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 4$	3,4806	51,94	2897,70	20,02	42,43	0,0275	-0,69	3,68	0,32	0,37	0,0435	-2,29	19,84	1,46	1,70
$\phi_2 = 5, 5$	6,2920	-79,20	7172,92	90,09	85,98	0,0894	-1,22	8,09	0,66	0,90	0,1212	-4,40	42,27	2,29	3,75
$\phi_3 = 3, 1$	3,3143	-21,43	652,59	19,32	24,16	0,1244	-0,72	8,87	0,83	0,85	0,1892	-7,20	81,24	2,94	3,18
$\omega_1 = 0, 0206$	0,0243	-0,37	3,62	0,35	1,85	0,1647	-0,83	8,86	0,82	0,72	0,2574	-10,10	128,42	2,63	3,05
$\omega_2 = 0, 0772$	0,0856	-0,84	11,27	1,06	1,86	0,1338	-0,55	5,08	0,48	0,56	0,1746	-4,63	34,50	1,31	2,13
$\omega_3 = 0, 1172$	0,1215	-0,43	9,73	0,95	1,01	0,2403	0,15	8,29	0,83	0,80	0,3691	-12,73	188,31	2,63	3,01
$\alpha_{12} = 0, 1564$	0,1616	-0,52	8,45	0,82	3,66	0,7738	1,62	22,26	1,96	1,89	0,7742	1,58	31,57	2,91	3,34
$\alpha_{22} = 0, 1283$	0,1300	-0,17	7,25	0,72	1,13	0,7579	2,16	25,09	2,04	2,66	0,7588	2,07	41,07	3,68	5,98
$\alpha_{33} = 0, 2418$	0,2421	-0,03	9,70	0,97	1,02	0,6373	0,77	21,54	2,10	2,17	0,6348	1,02	32,20	3,12	3,48
$\beta_{11} = 0, 7900$	0,7821	0,79	19,28	1,87	9,19										
$\beta_{22} = 0, 7795$	0,7658	1,37	33,31	3,14	5,40										
$\beta_{33} = 0, 6450$	0,6379	0,71	24,95	2,44	2,52										
Val.ini	Modèle 1; T = 5000														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 4$	3,4356	56,44	3233,68	4,87	9,14	0,0219	-0,13	0,61	0,06	0,06	0,0347	-1,41	0,41	0,21	0,26
$\phi_2 = 5, 5$	6,2630	-76,30	5978,16	15,71	17,16	0,0784	-0,12	1,28	0,13	0,14	0,1051	-2,79	1,08	0,31	0,53
$\phi_3 = 3, 1$	3,2729	-17,29	339,20	4,03	4,79	0,1182	-0,10	1,54	0,15	0,15	0,1800	-6,28	4,45	0,51	0,55
$\omega_1 = 0, 0206$	0,0216	-0,10	0,91	0,09	0,30	0,1595	-0,31	1,47	0,14	0,14	0,2514	-9,50	9,54	0,51	0,59
$\omega_2 = 0, 0772$	0,0783	-0,11	1,55	0,15	0,28	0,1278	0,05	1,06	0,11	0,11	0,1712	-4,29	2,09	0,25	0,41
$\omega_3 = 0, 1172$	0,1173	-0,01	2,14	0,21	0,17	0,2403	0,15	1,64	0,16	0,16	0,3668	-12,50	16,22	0,59	0,60
$\alpha_{12} = 0, 1564$	0,1593	-0,29	2,01	0,19	0,72	0,7876	0,24	3,42	0,34	0,33	0,7866	0,34	0,50	0,49	0,57
$\alpha_{22} = 0, 1283$	0,1275	0,08	1,35	0,13	0,22	0,7769	0,26	4,22	0,42	0,44	0,7763	0,32	0,56	0,55	0,95
$\alpha_{33} = 0, 2418$	0,2400	0,18	2,23	0,22	0,20	0,6431	0,19	4,26	0,42	0,40	0,6428	0,22	0,65	0,64	0,64
$\beta_{11} = 0, 7900$	0,7883	0,17	4,99	0,50	1,65										
$\beta_{22} = 0, 7795$	0,7773	0,22	5,28	0,52	0,89										
$\beta_{33} = 0, 6450$	0,6444	0,06	5,88	0,59	0,47										

Val.ini= Valeurs initiales ; Est. = Estimateurs ; EQM= Erreur quadratique moyenne ; VÉ = Variance échantillonnale ; MVE= Moyenne des estimateurs de variance.

TABLE 3.3. Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 2 défini par (3.1.2) pour $T = 1000$ et 5000.

Val.ini	Modèle 2; $T = 1000$														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 27,9853$	24,2254	375,99	150803,40	943,17	1584,60	0,0222	-0,16	1,49	0,15	0,23	0,0252	-0,46	3,20	0,30	0,84
$\phi_2 = 28,6715$	32,8293	-415,78	195559,09	2268,70	2320,28	0,0798	-0,26	4,13	0,41	0,70	0,0912	-1,40	12,69	1,07	2,86
$\phi_3 = 36,3381$	38,8290	-249,09	91729,00	2968,22	3125,66	0,1199	-0,27	6,06	0,60	0,75	0,1297	-1,25	11,15	0,96	2,51
$\omega_1 = 0,0206$	0,0233	-0,27	2,78	0,27	1,91	0,0222	-0,16	1,49	0,15	0,23	0,0252	-0,46	3,20	0,30	0,84
$\omega_2 = 0,0772$	0,0872	-1,00	12,20	1,12	2,73	0,0798	-0,26	4,13	0,41	0,70	0,0912	-1,40	12,69	1,07	2,86
$\omega_3 = 0,1172$	0,1204	-0,32	9,78	0,97	1,01	0,1199	-0,27	6,06	0,60	0,75	0,1297	-1,25	11,15	0,96	2,51
$\alpha_{12} = 0,1564$	0,1592	-0,28	5,40	0,53	3,60	0,1590	-0,26	4,05	0,40	0,48	0,1681	-1,17	7,06	0,57	1,54
$\alpha_{22} = 0,1283$	0,1318	-0,35	6,75	0,66	1,65	0,1325	-0,42	4,55	0,44	0,51	0,1351	-0,68	6,59	0,61	1,64
$\alpha_{33} = 0,2418$	0,2394	0,24	8,41	0,84	0,85	0,2430	-0,12	6,27	0,63	0,66	0,2506	-0,88	8,40	0,76	2,05
$\beta_{11} = 0,7900$	0,7834	0,66	14,63	1,42	9,41	0,7851	0,49	8,66	0,84	1,23	0,7843	0,57	13,58	1,32	3,69
$\beta_{22} = 0,7795$	0,7637	1,58	35,09	3,26	7,70	0,7719	0,76	15,32	1,47	2,15	0,7669	1,26	29,04	2,74	7,34
$\beta_{33} = 0,6450$	0,6442	0,08	23,48	2,35	2,39	0,6411	0,39	15,39	1,52	1,85	0,6374	0,76	21,06	2,05	5,55

Val.ini	Modèle 2; $T = 5000$														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 27,9853$	23,8568	412,85	172758	231,76	328,70	0,0215	-0,09	0,46	0,05	0,04	0,0228	-0,22	0,60	0,06	0,14
$\phi_2 = 28,6715$	32,8328	-416,13	177683,80	451,81	469,17	0,0787	-0,15	1,30	0,13	0,13	0,0824	-0,52	1,77	0,15	0,43
$\phi_3 = 36,3381$	38,5070	-216,89	53081,23	604,08	613,41	0,1188	-0,16	1,48	0,15	0,14	0,1233	-0,61	2,01	0,16	0,44
$\omega_1 = 0,0206$	0,0216	-0,10	0,65	0,06	0,34	0,0215	-0,09	0,46	0,05	0,04	0,0228	-0,22	0,60	0,06	0,14
$\omega_2 = 0,0772$	0,0786	-0,14	1,81	0,18	0,43	0,0787	-0,15	1,30	0,13	0,13	0,0824	-0,52	1,77	0,15	0,43
$\omega_3 = 0,1172$	0,1186	-0,14	1,72	0,17	0,18	0,1188	-0,16	1,48	0,15	0,14	0,1233	-0,61	2,01	0,16	0,44
$\alpha_{12} = 0,1564$	0,1573	-0,09	1,38	0,14	0,74	0,1571	-0,07	0,96	0,10	0,10	0,1674	-1,10	2,41	0,12	0,31
$\alpha_{22} = 0,1283$	0,1285	-0,02	1,46	0,15	0,32	0,1285	-0,02	1,01	0,10	0,10	0,1347	-0,64	1,60	0,12	0,32
$\alpha_{33} = 0,2418$	0,2412	0,06	1,60	0,16	0,17	0,2415	0,03	1,31	0,13	0,13	0,2506	-0,88	2,25	0,15	0,41
$\beta_{11} = 0,7900$	0,7880	0,20	3,55	0,35	1,83	0,7883	0,17	2,51	0,25	0,24	0,7881	0,19	2,70	0,27	0,68
$\beta_{22} = 0,7795$	0,7773	0,22	5,92	1,34	1,34	0,7772	0,23	4,14	0,41	0,42	0,7774	0,21	4,41	0,44	1,21
$\beta_{33} = 0,6450$	0,6440	0,10	4,25	0,42	0,46	0,6436	0,14	3,68	0,37	0,36	0,6434	0,16	3,94	0,39	1,03

Val.ini= Valeurs initiales ; Est. = Estimateurs ; EQM= Erreur quadratique moyenne ; VÉ = Variance échantillonnale ; MVE= Moyenne des estimateurs de variance.

TABLE 3.4. Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 3 défini par (3.1.3) pour $T = 1000$ et 5000.

Val.ini	Modèle 3; T = 1000														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 1, 0000$	0,8835	11,65	146,37	1,07	1,83	0,0262	-0,56	2,29	0,20	0,23	0,1669	-14,63	446,88	23,27	2,14
$\phi_2 = 1, 0000$	1,1167	-11,67	159,14	2,29	2,39	0,0923	-1,51	11,09	0,88	1,05	0,4079	-33,07	1530,21	43,68	8,17
$\phi_3 = 1, 0000$	1,0540	-5,40	46,59	1,74	1,95	0,1232	-0,60	8,91	0,86	0,91	0,4753	-35,81	1563,15	28,06	4,18
$\omega_1 = 0, 0206$	0,0233	-0,27	2,62	0,25	0,65	0,1701	-1,37	8,19	0,63	0,68	0,8545	-69,81	5074,87	20,14	5,92
$\omega_2 = 0, 0772$	0,0843	-0,71	13,12	1,26	1,83	0,1367	-0,84	7,65	0,69	0,76	0,5356	-40,73	1908,72	25,00	4,65
$\omega_3 = 0, 1172$	0,1166	0,06	10,02	1,00	1,02	0,2412	0,06	11,95	1,19	1,14	0,8962	-65,44	4430,42	14,78	5,51
$\alpha_{12} = 0, 1564$	0,1660	-0,96	8,97	0,81	2,07	0,7774	1,26	12,82	1,12	1,39	0,7695	2,05	36,56	3,24	0,43
$\alpha_{22} = 0, 1283$	0,1261	0,22	10,48	1,04	1,32	0,7467	3,28	41,84	3,11	3,40	0,7328	4,67	106,74	8,50	1,50
$\alpha_{33} = 0, 2418$	0,2369	0,49	13,62	1,34	1,35	0,6336	1,14	27,80	2,65	2,68	0,6423	0,27	38,86	3,88	0,71
$\beta_{11} = 0, 7900$	0,7860	0,40	13,78	1,36	4,11										
$\beta_{22} = 0, 7795$	0,7630	1,65	49,61	4,69	6,20										
$\beta_{33} = 0, 6450$	0,6424	0,26	33,26	3,32	3,07										
Val.ini	Modèle 3; T = 5000														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 1, 0000$	0,8784	12,16	150,44	0,25	0,37	0,0227	-0,21	0,41	0,04	0,04	0,1222	-10,16	128,02	2,47	0,27
$\phi_2 = 1, 0000$	1,1154	-11,54	137,48	0,44	0,48	0,0794	-0,22	1,28	0,12	0,15	0,3200	-24,28	647,58	5,80	0,89
$\phi_3 = 1, 0000$	1,0480	-4,80	26,47	0,34	0,38	0,1168	0,04	1,72	0,17	0,16	0,4572	-34,00	1203,22	4,69	0,68
$\omega_1 = 0, 0206$	0,0219	-0,13	0,53	0,05	0,11	0,1666	-1,02	2,36	0,13	0,13	0,8599	-70,35	5023,23	7,44	1,16
$\omega_2 = 0, 0772$	0,0753	0,19	1,77	0,17	0,26	0,2384	0,05	1,22	0,12	0,14	0,5095	-38,12	1503,51	5,04	0,85
$\omega_3 = 0, 1172$	0,1166	0,06	2,12	0,21	0,18	0,7865	0,35	2,77	0,26	0,22	0,9228	-68,10	4685,52	4,76	1,12
$\alpha_{12} = 0, 1564$	0,1649	-0,85	2,49	0,18	0,39	0,2384	0,34	2,65	0,25	0,26	0,7843	0,57	6,05	0,57	0,07
$\alpha_{22} = 0, 1283$	0,1234	0,49	2,01	0,18	0,24	0,7698	0,97	5,40	0,45	0,56	0,7701	0,94	13,17	1,23	0,19
$\alpha_{33} = 0, 2418$	0,2382	0,36	3,09	0,30	0,26	0,6424	0,26	5,76	0,57	0,49	0,6447	0,03	8,19	0,82	0,12
$\beta_{11} = 0, 7900$	0,7892	0,08	3,57	0,36	0,72										
$\beta_{22} = 0, 7795$	0,7782	0,13	6,67	0,67	0,96										
$\beta_{33} = 0, 6450$	0,6424	0,26	7,10	0,70	0,57										

Val.ini= Valeurs initiales ; Est. = Estimateurs ; EQM= Erreur quadratique moyenne ; VÉ = Variance échantillonnale ; MVE= Moyenne des estimateurs de variance.

TABLE 3.5. Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 4 défini par (3.1.4) pour $T = 1000$ et 5000.

Val.ini	Modèle 4; T = 1000														
	Méthode BHHH					Méthode GMM					Méthode WNLS				
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 27, 0821$	23,5182	356,39	136468,27	945,64	1425,96	0,0203	0,03	0,31	0,03	0,04	0,0242	-0,36	0,74	0,06	0,16
$\phi_2 = 28, 1324$	32,5289	-439,65	213943,25	2064,98	2223,07	0,0398	0,08	0,63	0,06	0,07	0,0446	-0,40	1,23	0,11	0,28
$\phi_3 = 35, 8882$	38,3465	-245,83	95782,82	3535,11	3068,91	0,0883	0,99	4,35	0,34	0,58	0,1102	-1,20	11,13	0,97	2,58
$\omega_1 = 0, 0206$	0,0229	-0,23	0,71	0,07	0,79	0,2011	-0,50	5,68	0,54	0,69	0,2060	-0,99	9,82	0,88	2,26
$\omega_2 = 0, 0406$	0,0436	-0,30	1,46	0,14	0,47	0,1896	-0,32	5,85	0,57	0,73	0,1994	-1,30	10,68	0,90	2,42
$\omega_3 = 0, 0982$	0,1013	-0,31	8,52	0,84	1,42	0,0806	-0,31	8,69	0,86	0,89	0,0841	-0,66	12,91	1,25	3,33
$\alpha_{11} = 0, 1961$	0,1953	0,08	8,18	0,82	9,94	0,0548	-0,19	5,13	0,51	0,54	0,0587	-0,58	6,78	0,64	1,76
$\alpha_{21} = 0, 1864$	0,1899	-0,35	11,39	1,13	4,18	0,0669	0,25	6,04	0,60	0,62	0,0688	0,06	7,74	0,77	2,02
$\alpha_{31} = 0, 0775$	0,0796	-0,21	12,54	1,25	4,55	0,1769	0,01	7,13	0,71	0,62	0,1822	-0,52	7,81	0,75	2,00
$\alpha_{12} = 0, 0529$	0,0552	-0,23	6,47	0,64	4,06	0,6626	0,49	11,88	1,16	1,72	0,6612	0,63	20,06	1,97	5,36
$\alpha_{22} = 0, 0694$	0,0688	0,06	7,64	0,76	2,00	0,6734	-0,19	15,76	1,57	1,78	0,6667	0,48	22,85	2,26	5,63
$\alpha_{33} = 0, 177$	0,1752	0,18	7,29	0,73	0,86	0,6737	-1,19	21,93	2,05	2,48	0,6510	1,08	35,42	3,43	9,09
$\beta_{11} = 0, 6675$	0,6580	0,95	23,95	2,30	16,27										
$\beta_{22} = 0, 6715$	0,6603	1,12	31,22	3,00	7,03										
$\beta_{33} = 0, 6618$	0,6583	0,35	34,77	3,46	4,53										

Val.ini	Modèle 4; T = 5000														
	Méthode BHHH					Méthode GMM					Méthode WNLS				
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 27, 0821$	23,1546	392,75	156433,13	217,98	302,50	0,0212	-0,06	0,09	0,01	0,01	0,0226	-0,20	0,14	0,01	0,03
$\phi_2 = 28, 1324$	32,2594	-412,70	174244,87	392,11	446,63	0,0408	-0,02	0,15	0,01	0,02	0,0429	-0,23	0,22	0,02	0,05
$\phi_3 = 35, 8882$	38,0517	-216,35	52413,19	560,44	595,93	0,0999	-0,17	1,49	0,15	0,14	0,1038	-0,56	1,90	0,16	0,44
$\omega_1 = 0, 0206$	0,0214	-0,08	0,14	0,01	0,16	0,1960	0,01	1,34	0,13	0,14	0,2093	-1,32	3,38	0,17	0,45
$\omega_2 = 0, 0406$	0,0408	-0,02	0,20	0,02	0,09	0,1866	-0,02	1,43	0,14	0,15	0,1960	-0,96	2,56	0,16	0,48
$\omega_3 = 0, 0982$	0,1001	-0,19	1,81	0,18	0,28	0,0779	-0,04	2,06	0,21	0,19	0,0810	-0,35	2,43	0,23	0,61
$\alpha_{11} = 0, 1961$	0,1968	-0,07	1,98	0,20	0,91	0,0537	-0,08	1,13	0,11	0,11	0,0572	-0,43	1,57	0,14	0,35
$\alpha_{21} = 0, 1864$	0,1872	-0,08	1,99	0,20	0,91	0,0689	0,05	1,28	0,13	0,13	0,0723	-0,29	1,59	0,15	0,40
$\alpha_{31} = 0, 0775$	0,0789	-0,14	2,51	0,25	1,01	0,1767	0,03	1,40	0,14	0,13	0,1832	-0,62	1,84	0,15	0,40
$\alpha_{12} = 0, 0529$	0,0535	-0,06	1,50	0,15	0,83	0,6654	0,21	3,74	0,37	0,35	0,6651	0,24	4,05	0,40	1,01
$\alpha_{22} = 0, 0694$	0,0686	0,08	1,76	0,17	0,40	0,6707	0,08	3,82	0,38	0,37	0,6707	0,08	3,92	0,39	1,07
$\alpha_{33} = 0, 177$	0,1771	-0,01	1,58	0,16	0,17	0,6595	0,23	6,12	0,61	0,57	0,6594	0,24	6,21	0,62	1,65
$\beta_{11} = 0, 6675$	0,6642	0,33	5,66	0,56	3,19										
$\beta_{22} = 0, 6715$	0,6705	0,10	5,07	0,51	1,34										
$\beta_{33} = 0, 6618$	0,6584	0,34	7,37	0,73	0,87										

Val.ini= Valeurs initiales ; Est. = Estimateurs ; EQM= Erreur quadratique moyenne ; VÉ = Variance échantillonnale ; MVE= Moyenne des estimateurs de variance.

TABLE 3.6. Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 5 défini par (3.1.5) pour $T = 1000$ et 5000.

Val.ini	Modèle 5; $T = 1000$						Modèle 5; $T = 5000$								
	Méthode BHHH			Méthode GMM			Méthode BHHH			Méthode GMM					
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 24, 5$	21,3488	315,12	108606,11	930,31	1416,59	0,0208	-0,02	0,39	0,04	0,05	0,0244	-0,38	0,82	0,07	0,17
$\phi_2 = 35, 63$	40,8609	-523,09	292748,56	1912,37	3531,61	0,0409	-0,03	0,44	0,04	0,06	0,0439	-0,33	0,88	0,08	0,20
$\phi_3 = 12, 98$	14,0100	-103,00	17002,92	639,48	461,51	0,0885	0,97	4,85	0,39	0,73	0,1166	-1,84	17,50	1,41	3,34
$\omega_1 = 0, 0206$	0,0230	-0,24	85,88	8,58	0,85	0,2032	-0,71	6,14	0,56	0,70	0,2081	-1,20	10,51	0,91	2,28
$\omega_2 = 0, 0406$	0,0424	-0,18	89,43	8,94	0,23	0,1933	-0,69	5,53	0,50	0,53	0,1976	-1,12	7,54	0,63	1,73
$\omega_3 = 0, 0982$	0,1007	-0,25	383,59	38,35	2,28	0,0842	-0,67	16,66	1,62	2,02	0,0904	-1,29	33,98	3,23	8,20
$\alpha_{11} = 0, 1864$	0,1917	0,44	311,26	31,11	10,18	0,0538	-0,09	5,61	0,56	0,77	0,0600	-0,71	9,62	0,91	2,46
$\alpha_{31} = 0, 0775$	0,0808	-0,33	587,45	58,73	14,90	0,0655	0,39	4,86	0,47	0,63	0,0686	0,08	7,47	0,75	1,98
$\alpha_{12} = 0, 0529$	0,0564	-0,35	306,78	30,67	4,40	0,1753	0,17	5,44	0,54	0,65	0,1927	-1,57	11,65	0,92	2,19
$\alpha_{22} = 0, 0694$	0,0685	0,09	282,49	28,25	1,30	0,6601	0,74	15,52	1,50	1,93	0,6604	0,71	22,56	2,20	5,86
$\alpha_{33} = 0, 177$	0,1711	0,59	276,85	27,65	0,80	0,6687	0,28	11,80	1,17	1,45	0,6684	0,31	17,72	1,76	4,45
$\beta_{11} = 0, 6675$	0,6597	0,78	520,20	51,96	14,04	0,6738	-1,20	19,45	1,80	2,84	0,6490	1,28	42,89	4,12	9,67
$\beta_{22} = 0, 6715$	0,6674	0,41	445,94	44,58	3,56										
$\beta_{33} = 0, 6618$	0,6625	-0,07	665,05	66,50	4,36										

Val.ini	Méthode BHHH			Méthode GMM			Méthode BHHH			Méthode GMM					
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 24, 5$	21,0345	346,55	121940,64	184,50	333,68	0,0212	-0,06	0,10	0,01	0,01	0,0228	-0,22	0,16	0,01	0,03
$\phi_2 = 35, 63$	40,9364	-530,64	289307,09	772,91	721,15	0,0409	-0,03	0,11	0,01	0,01	0,0426	-0,20	0,16	0,01	0,04
$\phi_3 = 12, 98$	13,7541	-77,41	6758,67	76,68	95,27	0,1002	-0,20	1,84	0,18	0,17	0,1095	-1,13	3,62	0,23	0,56
$\omega_1 = 0, 0206$	0,0211	-0,05	0,16	0,02	0,17	0,1961	0,00	1,40	0,14	0,14	0,2113	-1,52	4,02	0,17	0,45
$\omega_2 = 0, 0406$	0,0408	-0,02	0,15	0,02	0,04	0,1868	-0,04	1,07	0,11	0,11	0,1949	-0,85	1,88	0,12	0,34
$\omega_3 = 0, 0982$	0,0983	-0,01	2,77	0,28	0,43	0,0793	-0,18	4,95	0,49	0,45	0,0853	-0,78	6,95	0,63	1,49
$\alpha_{11} = 0, 1864$	0,1955	0,06	1,82	0,18	2,22	0,0540	-0,11	1,60	0,16	0,15	0,0579	-0,50	2,24	0,20	0,48
$\alpha_{31} = 0, 0775$	0,0756	0,19	6,90	0,69	3,10	0,0693	0,01	1,22	0,12	0,13	0,0720	-0,26	1,53	0,15	0,39
$\alpha_{12} = 0, 0529$	0,0531	-0,02	1,94	0,19	0,86	0,1772	-0,02	1,44	0,14	0,14	0,1933	-1,63	4,52	0,19	0,43
$\alpha_{22} = 0, 0694$	0,0697	-0,03	1,71	0,17	0,26	0,6647	0,28	4,30	0,42	0,39	0,6649	0,26	4,55	0,45	1,10
$\alpha_{33} = 0, 177$	0,1765	0,05	1,94	0,19	0,17	0,6701	0,14	3,15	0,31	0,30	0,6710	0,05	3,14	0,31	0,86
$\beta_{11} = 0, 6675$	0,6668	0,07	4,96	0,50	2,53	0,6577	0,41	7,04	0,69	0,65	0,6570	0,28	7,62	0,75	1,74
$\beta_{22} = 0, 6715$	0,6700	0,15	3,96	0,39	0,65										
$\beta_{33} = 0, 6618$	0,6625	-0,07	9,60	0,96	0,78										

Val.ini= Valeurs initiales ; Est. = Estimateurs ; EQM= Erreur quadratique moyenne ; VÉ = Variance échantillonnale ; MVE= Moyenne des estimateurs de variance.

TABLE 3.7. Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 6 défini par (3.1.6) pour $T = 1000$ et 5000.

Val.ini	Modèle 6; $T = 1000$																	
	Méthode BHHH						Méthode GMM						Méthode WNLS					
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$			
$\phi_1 = 1$	0,8870	11,30	139,69	1,21	1,84	0,0237	-0,31	0,65	0,06	0,05	0,1218	-10,12	142,38	4,01	0,67			
$\phi_2 = 1$	1,1121	-11,21	149,36	2,36	2,39	0,0424	-0,18	0,87	0,08	0,11	0,1730	-13,24	237,42	6,21	0,97			
$\phi_3 = 1$	1,0559	-5,59	49,92	1,86	1,97	0,1094	-1,12	11,72	1,05	0,99	0,4231	-32,49	1399,41	34,37	5,96			
$\omega_1 = 0, 0, 0206$	0,0228	-0,22	0,68	0,06	0,16	0,0237	-0,31	0,65	0,06	0,05	0,1218	-10,12	142,38	4,01	0,67			
$\omega_2 = 0, 0, 0406$	0,0432	-0,26	1,86	0,18	0,22	0,0424	-0,18	0,87	0,08	0,11	0,1730	-13,24	237,42	6,21	0,97			
$\omega_3 = 0, 0, 0982$	0,1022	-0,40	11,13	1,10	1,16	0,1094	-1,12	11,72	1,05	0,99	0,4231	-32,49	1399,41	34,37	5,96			
$\alpha_{11} = 0, 1, 961$	0,2036	-0,75	12,85	1,23	3,55	0,2080	-1,19	12,85	1,14	1,14	0,9208	-72,47	5415,66	16,44	14,20			
$\alpha_{11} = 0, 1, 864$	0,1919	-0,55	14,60	1,43	2,38	0,1916	-0,52	9,53	0,93	1,17	0,7168	-53,04	3109,56	29,63	11,24			
$\alpha_{31} = 0, 0, 775$	0,0856	-0,81	20,54	1,99	1,96	0,0862	-0,87	13,94	1,32	1,43	0,3072	-22,97	935,50	40,79	15,73			
$\alpha_{12} = 0, 0, 0529$	0,0575	-0,46	8,04	0,78	1,90	0,0591	-0,62	6,34	0,60	0,65	0,3411	-28,82	1205,99	37,53	4,58			
$\alpha_{22} = 0, 0, 0694$	0,0613	0,81	12,12	1,15	1,45	0,0670	0,24	7,88	0,78	0,76	0,2794	-21,00	721,97	28,09	4,35			
$\alpha_{33} = 0, 1, 177$	0,1698	0,72	13,26	1,28	1,18	0,1831	-0,61	11,29	1,09	1,03	0,6863	-50,93	2854,80	26,11	4,78			
$\beta_{11} = 0, 0, 6675$	0,6611	0,64	23,63	2,32	6,61	0,6564	1,11	21,08	1,99	2,16	0,6662	0,13	42,99	4,30	0,67			
$\beta_{22} = 0, 0, 6715$	0,6512	2,03	42,69	3,86	5,49	0,6554	1,61	25,65	2,31	2,55	0,6564	1,51	75,60	7,33	1,03			
$\beta_{33} = 0, 0, 6618$	0,6516	1,02	55,11	5,41	4,95	0,6341	2,77	50,61	4,29	4,06	0,6399	2,19	86,08	8,13	1,31			

Val.ini	Modèle 6; $T = 5000$																	
	Méthode BHHH						Méthode GMM						Méthode WNLS					
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$			
$\phi_1 = 1$	0,8795	12,05	147,73	0,25	0,37	0,0222	-0,16	0,13	0,01	0,01	0,1104	-8,98	86,49	0,59	0,11			
$\phi_2 = 1$	1,1153	-11,53	137,33	0,44	0,48	0,0394	0,12	0,18	0,02	0,02	0,1588	-11,82	148,33	0,87	0,15			
$\phi_3 = 1$	1,0469	-4,69	25,51	0,35	0,38	0,0986	-0,04	1,82	0,18	0,16	0,3876	-28,94	891,34	5,38	0,90			
$\omega_1 = 0, 0, 0206$	0,0220	-0,14	0,15	0,01	0,03	0,0207	-1,06	3,36	0,22	0,22	0,9820	-78,59	6193,21	1,67	2,87			
$\omega_2 = 0, 0, 0406$	0,0389	0,17	0,26	0,02	0,03	0,0394	0,12	0,18	0,02	0,02	0,1588	-11,82	148,33	0,87	0,15			
$\omega_3 = 0, 0, 0982$	0,0985	-0,03	2,40	0,24	0,19	0,0986	-0,04	1,82	0,18	0,16	0,3876	-28,94	891,34	5,38	0,90			
$\alpha_{11} = 0, 1, 961$	0,2059	-0,98	4,29	0,33	0,67	0,2067	-1,06	3,36	0,22	0,22	0,9820	-78,59	6193,21	1,67	2,87			
$\alpha_{21} = 0, 1, 864$	0,1793	0,71	3,26	0,28	0,40	0,1803	0,61	2,27	0,19	0,22	0,7087	-52,23	2793,81	6,61	2,13			
$\alpha_{31} = 0, 0, 775$	0,0778	-0,03	3,60	0,36	0,31	0,0770	0,05	2,61	0,26	0,24	0,2951	-21,76	547,68	7,43	2,50			
$\alpha_{12} = 0, 0, 0529$	0,0555	-0,26	1,79	0,17	0,36	0,0564	-0,35	1,43	0,13	0,13	0,3089	-25,60	747,00	9,14	0,87			
$\alpha_{22} = 0, 0, 0694$	0,0658	0,36	2,04	0,19	0,26	0,0666	0,28	1,59	0,15	0,15	0,2744	-20,50	479,10	5,87	0,84			
$\alpha_{33} = 0, 1, 177$	0,1741	0,29	2,86	0,28	0,23	0,1748	0,22	2,51	0,25	0,20	0,6838	-50,68	2636,96	6,87	0,94			
$\beta_{11} = 0, 0, 6675$	0,6664	0,11	5,91	0,59	1,17	0,6645	0,30	4,46	0,44	0,41	0,6757	-0,82	8,24	0,76	0,12			
$\beta_{22} = 0, 0, 6715$	0,6724	-0,09	6,09	0,61	0,87	0,6702	0,13	4,57	0,46	0,48	0,6720	-0,05	10,70	1,07	0,19			
$\beta_{33} = 0, 0, 6618$	0,6576	0,42	11,36	1,12	0,88	0,6579	0,39	8,85	0,87	0,74	0,6579	0,39	16,00	1,58	0,22			

Val.ini= Valeurs initiales ; Est. = Estimateurs ; EQM= Erreur quadratique moyenne ; VÉ = Variance échantillonnale ; MVE= Moyenne des estimateurs de variance.

TABLE 3.8. Propriétés empiriques des estimations ponctuelles et estimations de variances lors d'une simulation de taille 1000 pour le modèle 7 défini par (3.1.7) pour $T = 1000$ et 5000.

Val.ini	Modèle 7; $T = 1000$														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 27, 9853$	24,1070	387,83	159885,92	947,09	1656,30	0,0201	-0,01	0,07	0,01	0,01	0,0219	-0,19	0,12	0,01	0,02
$\phi_2 = 28, 6715$	32,8818	-421,03	200461,26	2319,84	2397,85	0,0769	0,01	0,67	0,07	0,07	0,0815	-0,45	0,98	0,08	0,22
$\phi_3 = 36, 3381$	38,8987	-256,06	97207,28	3164,13	3184,99	0,1173	0,27	2,65	0,26	0,27	0,1219	-0,19	3,38	0,33	0,85
$\omega_1 = 0, 02$	0,0206	-0,06	0,08	0,01	0,03	0,6107	-0,07	8,22	0,82	0,87	0,6470	-3,70	24,17	1,05	2,72
$\omega_2 = 0, 077$	0,0780	-0,10	0,81	0,08	0,17	0,5377	0,23	9,54	0,95	0,90	0,5636	-2,36	16,33	1,08	2,78
$\omega_3 = 0, 12$	0,1176	0,24	3,13	1,05	5,59	0,4084	-0,84	13,31	1,26	0,95	0,4140	-1,40	13,37	1,14	2,93
$\alpha_{11} = 0, 61$	0,6078	0,22	10,53	1,02	2,24	0,1999	0,01	14,49	1,45	1,54	0,1976	0,24	16,60	1,65	4,31
$\alpha_{22} = 0, 54$	0,5374	0,26	10,26	1,02	2,24	0,1018	-0,18	22,61	2,26	2,61	0,0977	0,23	25,78	2,57	7,37
$\alpha_{33} = 0, 4$	0,3998	0,02	11,11	1,11	1,27	0,2355	1,45	45,56	4,35	3,93	0,2427	0,73	46,96	4,64	11,58
$\beta_{11} = 0, 2$	0,1982	0,18	17,74	1,77	8,30										
$\beta_{22} = 0, 1$	0,0963	0,37	28,89	2,87	5,98										
$\beta_{33} = 0, 25$	0,2418	0,82	47,76	4,71	5,21										
Val.ini	Modèle 7; $T = 5000$														
	Méthode BHHH				Méthode GMM				Méthode WNLS						
	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$	Est.	Biais $\times 10^{-2}$	EQM $\times 10^{-4}$	VÉ $\times 10^{-3}$	MVE $\times 10^{-3}$
$\phi_1 = 27, 9853$	23,8230	416,23	178087,75	483,65	334,55	0,0202	-0,02	0,01	0,00	0,00	0,0215	-0,15	0,04	0,00	0,00
$\phi_2 = 28, 6715$	32,7904	-411,89	176325,29	667,17	472,86	0,0770	0,00	0,16	0,02	0,01	0,0808	-0,38	0,31	0,02	0,04
$\phi_3 = 36, 3381$	38,5253	-218,72	55461,37	762,11	616,43	0,1155	0,45	0,83	0,06	0,05	0,1196	0,04	0,57	0,06	0,16
$\omega_1 = 0, 02$	0,0202	-0,02	12,22	1,22	0,01	0,6111	-0,11	1,79	0,18	0,17	0,6503	-4,03	18,38	0,21	0,55
$\omega_2 = 0, 077$	0,0771	-0,01	40,15	4,02	0,03	0,5387	0,13	2,17	0,21	0,18	0,5653	-2,53	8,60	0,22	0,56
$\omega_3 = 0, 12$	0,1153	0,47	79,08	7,89	0,07	0,3996	0,04	2,44	0,24	0,19	0,4140	-1,40	4,21	0,22	0,59
$\alpha_{11} = 0, 61$	0,6110	-0,10	143,44	14,34	1,14	0,1987	0,13	3,18	0,32	0,31	0,1987	0,13	3,35	0,33	0,85
$\alpha_{22} = 0, 54$	0,5389	0,11	146,40	14,64	0,44	0,1000	0,00	6,14	0,61	0,52	0,0994	0,06	5,69	0,57	1,46
$\alpha_{33} = 0, 4$	0,3996	0,04	153,31	15,33	0,25	0,2483	0,17	9,73	0,97	0,80	0,2496	0,04	8,43	0,84	2,28
$\beta_{11} = 0, 2$	0,1988	0,12	194,70	19,47	1,63										
$\beta_{22} = 0, 1$	0,0992	0,08	248,59	24,86	1,15										
$\beta_{33} = 0, 25$	0,2492	0,08	309,02	30,90	1,01										

Val.ini= Valeurs initiales ; Est. = Estimateurs ; EQM= Erreur quadratique moyenne ; VÉ = Variance échantillonnale ; MVE= Moyenne des estimateurs de variance.

Chapitre 4

ANALYSES DE DONNÉES RÉELLES

Dans ce chapitre, nous illustrons la théorie vue au deuxième chapitre sur des données financières. Les séries sont présentées dans la section 4.1. Les données choisies sont tirées de l'article de Cipollini, Engle et Gallo (2012) et sont disponibles sur le site de *Journal of Applied Econometrics*¹. Trois indices boursiers sont considérés, à savoir le S&P500, le S&P400 et le FTSE100. Chaque série financière est composée de trois indices, impliquant des séries trivariées. La période d'observation s'étend de janvier 1996 à février 2009.

4.1. ANALYSE ET MODÉLISATION DE CHAQUE SÉRIE

Les indices boursiers sont calculés comme dans Cipollini, Engle et Gallo (2012). Plus précisément, nous définissons trois variables annualisées définies comme suit :

$$|r_t| = |\text{rendement}_t| \times \sqrt{252} \times 100, \quad (4.1.1)$$

$$rv_t = \sqrt{\text{variance réalisée}_t} \times \sqrt{252} \times 100, \quad (4.1.2)$$

$$hl_t = \ln \left(\frac{\text{prix le plus élevé}_t}{\text{prix le plus bas}_t} \right) \times \sqrt{252} \times 100. \quad (4.1.3)$$

Ainsi, la variable $|r_t|$ désigne le rendement absolu (*Absolute return* en anglais) au temps t . Il est bien connu qu'une étude de la volatilité repose sur le comportement des rendements carrés. Une mesure alternative peut être les rendements absolus. En effet, dans le cas où $X \sim N(0, \sigma^2)$, on peut montrer que $E|X| = \sigma\sqrt{2/\pi}$, qui est une fonction de σ^2 . Voir Hogg et Craig (1978, p. 115). C'est le genre de mesure qui a été étudiée dans Cipollini, Engle et Gallo (2012). Les rendements ont été obtenus directement du Oxford-Man Institute (OMI).

1. <http://qed.econ.queensu.ca/jae/datasets/cipollini001/>.

La variable rv_t désigne la mesure de volatilité de type volatilité réalisée selon la méthode du noyau développée par Barndorff-Nielsen, Hansen, Lunde et Shephard (2008). Dans sa version la plus simple et naturelle, la volatilité réalisée est calculée en utilisant la somme des carrés des rendements intrajournaliers pour un jour donné. Voir Andersen, Bollerslev, Christoffersen et Diebold (2006). La variable hl_t désigne quant à elle la différence entre le prix le plus élevé et le prix le plus bas (*range* en anglais) au temps t . Notons que pour une année donnée, il y a 252 jours ouvrables en bourse.

Après exclusion des données manquantes dues pour la plupart aux jours fériés, nous avons retenu 3261 observations pour l'indice boursier S&P500, 2840 observations pour l'indice boursier S&P400 et 3251 observations pour l'indice boursier FTSE100. Il est à noter que ces variables sont déjà traitées dans la base de données sur le site du *Journal of Applied Econometrics*.

Dans l'annexe B, les différentes séries financières sont présentées. Les trois indices boursiers S&P500, S&P400 et FTSE100 sont représentés visuellement dans les graphiques B.1, B.2 et B.3, respectivement. Pour chaque graphique, on trouve la série de rendement absolu, la variance réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas. On remarque que ces séries affichent des caractéristiques courantes des séries financières, comme les regroupements en paquets (ou en grappes) de fortes variations, ainsi que de faibles variations. Ces phénomènes sont souvent appelés des effets ARCH (ou des effets de grappes ARCH ou GARCH). Ces constatations suggèrent donc qu'une spécification du modèle comme un vMEM avec des dynamiques GARCH serait potentiellement utile.

L'analyse qui a été faite en premier lieu consiste à considérer le modèle vectoriel et multiplicatif avec erreurs non-négatives le plus simple vMEM(1,1). Nous modélisons conjointement pour chaque indice boursier les trois séries formées par le rendement absolu, la volatilité réalisée selon la méthode du noyau et l'étendue. Les paramètres à estimer α , γ et β de la moyenne conditionnelle sont des matrices carrées diagonales d'ordre trois. Ainsi, la moyenne

conditionnelle a la forme suivante :

$$\begin{pmatrix} \mu_t^{r|} \\ \mu_t^{rv} \\ \mu_t^{hl} \end{pmatrix} = \boldsymbol{\omega} + \boldsymbol{\alpha} \begin{pmatrix} |r|_{t-1} \\ rv_{t-1} \\ hl_{t-1} \end{pmatrix} + \boldsymbol{\gamma} \mathbf{r}_{t-1}^{(-)} + \boldsymbol{\beta} \begin{pmatrix} \mu_{t-1}^{r|} \\ \mu_{t-1}^{rv} \\ \mu_{t-1}^{hl} \end{pmatrix}, \quad (4.1.4)$$

$$= \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & & \\ & \alpha_{22} & \\ & & \alpha_{33} \end{pmatrix} \begin{pmatrix} |r|_{t-1} \\ rv_{t-1} \\ hl_{t-1} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & & \\ & \gamma_{22} & \\ & & \gamma_{33} \end{pmatrix} \mathbf{r}_{t-1}^{(-)} + \begin{pmatrix} \beta_{11} & & \\ & \beta_{22} & \\ & & \beta_{33} \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{r|} \\ \mu_{t-1}^{rv} \\ \mu_{t-1}^{hl} \end{pmatrix}, \quad (4.1.5)$$

$$= \boldsymbol{\mu}_0 + (\boldsymbol{\alpha} + \boldsymbol{\gamma}/2 + \boldsymbol{\beta}) \begin{pmatrix} \mu_{t-1}^{r|} - \mu_0^{r|} \\ \mu_{t-1}^{rv} - \mu_0^{rv} \\ \mu_{t-1}^{hl} - \mu_0^{hl} \end{pmatrix} + \boldsymbol{\alpha} \begin{pmatrix} |r|_{t-1} - \mu_{t-1}^{r|} \\ rv_{t-1} - \mu_{t-1}^{rv} \\ hl_{t-1} - \mu_{t-1}^{hl} \end{pmatrix} + \boldsymbol{\gamma} \begin{pmatrix} |r|_{t-1}^{(-)} - \mu_{t-1}^{r|}/2 \\ rv_{t-1}^{(-)} - \mu_{t-1}^{rv}/2 \\ hl_{t-1}^{(-)} - \mu_{t-1}^{hl}/2 \end{pmatrix}; \quad (4.1.6)$$

où le vecteur $\boldsymbol{\omega}$ est égal à $\boldsymbol{\omega} = [\mathbf{I} - (\boldsymbol{\alpha} + \boldsymbol{\gamma} + \boldsymbol{\beta})] \boldsymbol{\mu}_0$ et le terme $\boldsymbol{\gamma} \mathbf{r}_{t-1}^{(-)}$ vise à saisir les effets asymétriques associés avec le signe de la variable rendement. Le vecteur $\mathbf{r}_{t-1}^{(-)}$ contient le vecteur $\begin{pmatrix} |r|_{t-1} \\ rv_{t-1} \\ hl_{t-1} \end{pmatrix}$ multiplié par une variable indicatrice. Si le rendement journalier est négatif alors cette variable vaut un sinon elle vaut zéro.

Dans la première analyse, la forme de la moyenne conditionnelle retenue correspond à la formule (4.1.6). L'estimation des paramètres a été effectuée en utilisant deux méthodes. Nous avons retenu la méthode semi-paramétrique GMM, ainsi que la méthode paramétrique reposant sur la copule normale. Les paramètres estimés du modèle vMEM(1,1) pour les trois indices boursiers sont présentés dans les tables 4.1 et 4.2, respectivement.

TABLE 4.1. Estimation des paramètres du modèle vMEM(1,1) en utilisant la méthode semi-paramétrique GMM.

Param.	S&P500			S&P400			FTSE100		
	Val. init.	Val. est.	$g(\hat{\theta}) \times 10^{-8}$	Val. init.	Val. est.	$g(\hat{\theta}) \times 10^{-8}$	Val. init.	Val. est.	$g(\hat{\theta}) \times 10^{-8}$
α_{11}	0,0037	0,0103	-0,59	-0,0033	0,0107	0,22	-0,0075	0,0023	0,45
α_{22}	0,2110	0,2218	-0,79	0,1960	0,2267	0,87	0,1434	0,1714	-8,24
α_{33}	0,0538	0,0640	0,06	0,0508	0,0623	2,71	0,0457	0,0590	4,79
γ_{11}	0,0815	0,0783	-0,12	0,0859	0,0874	0,96	0,1001	0,0851	0,56
γ_{22}	0,0798	0,0756	-1,17	0,0760	0,0626	2,57	0,0745	0,0602	-5,69
γ_{33}	0,0890	0,0841	-0,08	0,0909	0,0887	0,15	0,0904	0,0746	2,02
β_{11}	0,9392	0,9319	-0,13	0,9395	0,9221	0,49	0,9408	0,9434	0,50
β_{22}	0,7221	0,7100	-0,71	0,7321	0,7228	0,71	0,8003	0,7835	-8,10
β_{33}	0,8799	0,8691	-0,24	0,8762	0,8697	4,67	0,8909	0,8899	0,55

Param.=Paramètres ; Val. init.= Valeurs initiales ; Val. est.=Valeurs estimées.

TABLE 4.2. Estimation des paramètres du modèle vMEM(1,1) en utilisant la méthode paramétrique maximum de vraisemblance.

Param.	S&P500		S&P400		FTSE100	
	Val. init.	Val. est.	Val. init.	Val. est.	Val. init.	Val. est.
ϕ_1	1,2600	1,2599	1,3100	1,3338	1,3700	1,3200
ϕ_2	13,9000	13,9024	15,0800	15,4617	15,6400	18,5945
ϕ_3	6,1000	6,1006	5,8800	6,0342	7,2100	7,1089
α_{11}	0,0037	0,0038	-0,0033	-0,0036	-0,0075	0,0464
α_{22}	0,2110	0,2110	0,1960	0,1735	0,1434	0,0915
α_{33}	0,0538	0,0540	0,0508	0,0140	0,0457	0,0206
γ_{11}	0,0815	0,0812	0,0859	0,0600	0,1001	-0,0717
γ_{22}	0,0798	0,0795	0,0760	0,0445	0,0745	0,0497
γ_{33}	0,0890	0,0894	0,0909	0,0812	0,0904	0,0175
β_{11}	0,9392	0,9391	0,9395	0,9532	0,9408	0,9032
β_{22}	0,7221	0,7221	0,7321	0,7840	0,8003	0,8656
β_{33}	0,8799	0,8804	0,8762	0,9245	0,8909	0,9504

Param.= Paramètres ; Val. init.= Valeurs initiales ; Val. est.= Valeurs estimées.

Les paramètres estimés sont très proches des valeurs initiales. De plus, si on remplace le vecteur des valeurs initiales θ par le vecteur des paramètres estimés $\hat{\theta}$ dans le système non-linéaire suivant :

$$g(\theta) = \frac{1}{T} \sum_{t=1}^T \nabla_{\theta} \mu_t^{\top} \text{diag}(\mu_t)^{-1} \Sigma^{-1} \mathbf{u}_t, \quad (4.1.7)$$

nous trouvons les résultats de $g(\hat{\theta})$ dans le tableau 4.1. D'après les résultats, nous remarquons que les paramètres estimés annulent bien le système non-linéaire, suggérant que l'estimation obtenue est bien solution au système à résoudre selon la méthode des moments généralisés (GMM).

Une fois les modèles estimés, nous avons pu déterminer les résidus :

$$\hat{\epsilon} = \mathbf{X} / \hat{\mu}.$$

Sous la spécification correcte du modèle, la série multivariée $\hat{\epsilon}$ doit être iid avec la distribution $D(\cdot)$. Les statistiques de portemanteau comme la statistique de Ljung-Box multivariée fondée sur les résidus du modèle vMEM peuvent être utilisés pour analyser si la spécification est capable de capturer les propriétés dynamiques du processus. En fonction de s délais, la statistique de Ljung-Box multivariée est définie par :

$$LB(s) = T(T + 2) \sum_{j=1}^s \left(\frac{1}{T-j} \right) \text{trace} [\hat{C}_{0j} \hat{C}_{00}^{-1} \hat{C}_{0j}^{\top} \hat{C}_{00}^{-1}],$$

où le nombre d'observation est T et le terme \hat{C}_{0j} est comme suit :

$$\hat{C}_{0j} = T^{-1} \sum_{t=j+1}^T \hat{\epsilon}_t \hat{\epsilon}_{t-j}^{\top}.$$

Sous l'hypothèse nulle H_0 , cette statistique suit la loi $\chi_{1-\alpha, h}^2$ où le nombre de degrés de liberté h est égal à (nombre de séries)² \times nombre de délais. Notons que la littérature sur les vMEM s'est contentée jusqu'à maintenant d'utiliser des statistiques de diagnostic pour séries univariées, appliquées par exemple à chaque série. Il nous semble préférable d'utiliser des statistiques permettant de valider des modèles multivariés.

Dans notre analyse, les statistiques de Ljung-Box multivariées ont été calculées. D'après les tableaux 4.3 et 4.4, nous pouvons constater qu'on rejette pour tous les délais l'hypothèse nulle, c'est-à-dire les résidus ne sont pas indépendants. Et donc, une amélioration du modèle sera nécessaire pour chacun des indices boursiers. Bref, une nouvelle analyse a été faite en second lieu mais avec des ordres de modèles différents de la première analyse.

TABLE 4.3. Test de Ljung-Box multivarié des indices boursiers étudiés par la méthode semi-paramétrique GMM.

m	S&P500			S&P400			FTSE100		
	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP
1	240,5612	9	0	168,2936	9	0	158,1225	9	0
2	303,8516	18	0	206,204	18	0	186,0608	18	0
6	381,3356	54	0	286,0369	54	0	257,9064	54	0
12	473,1189	108	0	369,6323	108	0	357,6720	108	0
22	582,8857	198	0	433,7651	198	0	433,0275	198	0
32	672,0003	288	0	516,9439	288	0	540,7688	288	0
46	824,5094	414	0	650,2106	414	0	647,9470	414	0

m = délais; Q(m)= statistique de Ljung-Box multivariée; D.d.l = Degré de liberté; VP=Valeur-P.

TABLE 4.4. Test de Ljung-Box multivarié des indices boursiers étudiés par la méthode paramétrique maximum de vraisemblance.

m	S&P500			S&P400			FTSE100		
	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP
1	231,9886	9	0	223,406	9	0	495,0769	9	0
2	292,4416	18	0	293,8604	18	0	923,1369	18	0
6	375,6917	54	0	484,5746	54	0	2287,5697	54	0
12	469,0426	108	0	643,2882	108	0	3613,5179	108	0
22	580,7734	198	0	739,3924	198	0	4971,0142	198	0
32	671,6154	288	0	847,4767	288	0	5958,6990	288	0
46	816,7298	414	0	997,7667	414	0	6680,7772	414	0

m = délais ; Q(m)= statistique de Ljung-Box multivariée ; D.d.l = Degré de liberté ; VP=Valeur-P.

En nous inspirant des essais faits par Cipollini, Engle et Gallo (2012) sur le choix convenable des modèles vMEMs des trois indices boursiers, nous avons appliqué un modèle vMEM(2,1) pour chaque indice boursier étudié où les matrices carrées α_1 , α_2 , γ et β sont des matrices non diagonales dont les coefficients en dehors de la diagonale principale ne sont pas nuls. La moyenne conditionnelle μ_t a la forme suivante :

$$\begin{pmatrix} \mu_t^{r|} \\ \mu_t^{v|} \\ \mu_t^{h|} \end{pmatrix} = \omega + \alpha_1 \begin{pmatrix} r|_{t-1} \\ rv|_{t-1} \\ hl|_{t-1} \end{pmatrix} + \alpha_2 \begin{pmatrix} r|_{t-2} \\ rv|_{t-2} \\ hl|_{t-2} \end{pmatrix} + \gamma r_{t-1}^{(-)} + \beta \begin{pmatrix} \mu_{t-1}^{r|} \\ \mu_{t-1}^{v|} \\ \mu_{t-1}^{h|} \end{pmatrix}. \quad (4.1.8)$$

L'estimation des paramètres de la deuxième analyse a été effectuée en utilisant la méthode semi-paramétrique GMM et la méthode paramétrique maximum de vraisemblance pour les trois indices boursiers (voir tableaux 4.5, 4.6 et 4.7). Pour la méthode GMM, nous avons rencontré le même message d'erreur déjà vu dans la simulation concernant les calculs des matrices inverses $\text{diag}(\mu_t)$ et $\text{diag}(\mu_t)\Sigma^{-1}$ qui n'étaient pas possible. Dans ces cas, nous avons utilisé la matrice inverse généralisée au lieu de la matrice inverse. D'après les tableaux 4.5, 4.6 et 4.7, nous remarquons que les valeurs estimés sont très proches des valeurs initiales et aussi les valeurs de $g(\hat{\theta})$ sont très proches de zéros. Donc, les méthodes paramétrique et semi-paramétrique estiment bien les modèles vMEM(2,1) malgré l'existence de plusieurs paramètres.

Ensuite, à partir des résidus estimés pour les différents indices boursiers, les statistiques de test de Ljung-Box multivariées ont été calculées et présentées

dans les tableaux 4.7 et 4.9. Cette fois-ci, nous trouvons, à l'aide des deux méthodes paramétrique et non-paramétrique, que les résidus semblent compatibles avec l'indépendance pour la majorité des délais m au niveau de 5%. Notons que les valeurs-p sont les probabilités correspondantes des valeurs des statistiques. Donc l'amélioration des valeurs-p montre que les modèles vMEM(2,1) décrivent mieux les séries trivariées utilisées.

TABLE 4.5. Estimation des paramètres du modèle vMEM(2,1) en utilisant la méthode semi-paramétrique GMM et la méthode paramétrique maximum de vraisemblance de l'indice boursier S&P500.

S&P500				
Param.	Val. init.	Val. est. I	$g(\hat{\theta}) \times 10^{-8}$	Val. est. II
α_{11}^1	-0,1045	-0,1080	-1,70	-0,1033
α_{21}^1	0,0166	0,0179	-3,40	0,0165
α_{12}^1	0,3956	0,3852	-6,94	0,4029
α_{22}^1	0,2849	0,2873	-6,10	0,2870
α_{32}^1	0,3548	0,3561	0,37	0,3348
α_{13}^1	-0,0559	-0,0414	-4,76	-0,0581
α_{23}^1	0,0506	0,0445	-5,28	0,0494
α_{33}^1	-0,0931	-0,0939	-0,17	-0,0801
α_{11}^2	0,0771	0,0786	-2,95	0,0771
α_{21}^2	0,0226	0,0311	-2,36	0,0236
α_{12}^2	-0,1395	-0,1434	-6,50	-0,1427
α_{22}^2	-0,1587	-0,1433	-5,46	-0,1589
α_{32}^2	-0,1934	-0,1906	0,59	-0,1828
α_{23}^2	0,0618	0,0593	-4,63	0,0597
α_{33}^2	0,0987	0,1001	1,19	0,0940
γ_{11}	0,0464	0,0501	-3,06	0,0437
γ_{22}	0,0751	0,0784	-11,85	0,0741
γ_{33}	0,0719	0,0723	0,23	0,0756
β_{11}	0,8132	0,8123	-3,96	0,8117
β_{21}	0,3274	0,3977	-3,09	0,3240
β_{22}	0,2795	0,1779	-3,20	0,2847
β_{33}	0,7898	0,7862	0,80	0,7894
ϕ_1	1,3000			1,3002
ϕ_2	14,6600			14,6571
ϕ_3	6,4300			6,4339

Param. = Paramètres ; Val. init.= Valeurs initiales ;

Val. est. = Valeurs estimées ; I = méthode semi-paramétrique ;

II = méthode paramétrique.

TABLE 4.6. Estimation des paramètres du modèle vMEM(2,1) en utilisant la méthode semi-paramétrique GMM et la méthode paramétrique maximum de vraisemblance de l'indice boursier S&P400.

S&P400				
Param.	Val. init.	Val. est. I	$g(\hat{\theta}) \times 10^{-5}$	Val. est. II
α_{11}^1	0,0012	0,0035	5,50	0,0015
α_{12}^1	0,3422	0,3458	4,72	0,3435
α_{22}^1	0,3151	0,3244	2,62	0,3151
α_{32}^1	0,3517	0,3514	5,40	0,3390
α_{23}^1	0,0652	0,0663	6,78	0,0654
α_{33}^1	0,0692	0,0782	1,74	0,0704
α_{22}^2	-0,132	-0,1369	7,11	-0,1323
α_{32}^2	-0,1105	-0,1023	3,09	-0,1000
α_{13}^2	0,047	0,0523	0,98	0,0463
α_{23}^2	-0,0552	-0,0571	2,88	-0,0558
β_{11}	0,9245	0,9150	3,31	0,9247
β_{21}	-0,0224	-0,0212	4,00	-0,0219
β_{12}	-0,3336	-0,3348	0,62	-0,3343
β_{22}	0,8209	0,8167	6,13	0,8214
β_{32}	-0,2112	-0,2146	1,19	-0,2106
β_{33}	0,8901	0,8771	3,97	0,8896
ϕ_1	1,3100			1,3107
ϕ_2	15,5600			15,5605
ϕ_3	5,8500			5,8550

Param. = Paramètres ; Val. init.= Valeurs initiales ;

Val. est. = Valeurs estimées ; I = méthode semi-paramétrique ;

II = méthode paramétrique.

TABLE 4.7. Estimation des paramètres du modèle vMEM(2,1) en utilisant la méthode semi-paramétrique GMM et la méthode paramétrique maximum de vraisemblance de l'indice boursier FTSE100.

FTSE100				
Param.	Val. init.	Val. est. I	$g(\hat{\theta}) \times 10^{-7}$	Val. est. II
α_{11}^1	-0,0470	-0,0470	6,35	-0,0383
α_{21}^1	0,0163	0,0154	-35,30	0,0338
α_{12}^1	0,3594	0,3672	8,60	0,3188
α_{22}^1	0,2984	0,2954	-113,61	0,3230
α_{32}^1	0,3333	0,3314	31,28	0,3292
α_{23}^1	0,0838	0,0875	-52,56	0,0591
α_{33}^1	0,0487	0,0506	2,59	0,0516
α_{11}^2	0,0741	0,0714	6,68	0,1768
α_{31}^2	0,0277	0,0264	-3,74	0,0671
α_{22}^2	-0,1757	-0,1714	-119,90	-0,2303
α_{32}^2	-0,1504	-0,1431	27,88	-0,1849
α_{23}^2	-0,0592	-0,0625	-72,79	-0,0334
β_{11}	0,8258	0,8179	6,66	0,7761
β_{12}	-0,2154	-0,2113	3,98	-0,2740
β_{22}	0,9772	0,9770	-104,52	0,9756
β_{23}	-0,1322	-0,1323	-62,20	-0,1261
β_{33}	0,7573	0,7523	30,34	0,7423
ϕ_1	1,3500			1,3654
ϕ_2	16,5000			17,1221
ϕ_3	7,1300			7,3078

Param. = Paramètres ; Val. init.= Valeurs initiales ;

Val. est. = Valeurs estimées ; I = méthode semi-paramétrique ;

II = méthode paramétrique.

TABLE 4.8. Test de Ljung-Box multivarié du modèle vMEM(2,1) des indices boursiers étudiés par la méthode semi-paramétrique.

m	S&P500			S&P400			FTSE100		
	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP
1	6,41	9	0,70	21,32	9	0,01	6,00	9	0,74
2	17,49	18	0,49	33,60	18	0,01	16,73	18	0,54
6	51,06	54	0,59	56,98	54	0,36	40,07	54	0,92
12	103,39	108	0,61	119,74	108	0,21	110,95	108	0,40
22	211,32	198	0,25	188,18	198	0,68	191,07	198	0,63
32	301,28	288	0,28	261,72	288	0,86	292,61	288	0,41
46	444,99	414	0,14	379,25	414	0,89	406,03	414	0,60

m = délais ; Q(m)= statistique de Ljung-Box multivariée ; D.d.l=Degré de liberté ; VP=Valeur-P.

TABLE 4.9. Test de Ljung-Box multivarié du modèle vMEM(2,1) des indices boursiers étudiés par la méthode paramétrique.

m	S&P500			S&P400			FTSE100		
	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP	Q(m)	D.d.l	VP
1	7,34	9	0,60	21,44	9	0,01	6,54	9	0,69
2	18,69	18	0,41	34,47	18	0,01	43,03	18	0,00
6	51,48	54	0,57	60,64	54	0,25	74,14	54	0,04
12	102,15	108	0,64	122,09	108	0,17	147,69	108	0,01
22	207,19	198	0,31	190,94	198	0,63	229,77	198	0,06
32	293,95	288	0,39	263,47	288	0,85	339,32	288	0,02
46	433,89	414	0,24	381,71	414	0,87	455,16	414	0,08

m = délais ; Q(m)= statistique de Ljung-Box multivariée ; D.d.l=Degré de liberté ; VP=Valeur-P.

CONCLUSION

Dans ce mémoire, nous nous sommes intéressés aux modèles vectoriels et multiplicatifs avec erreurs non-négatives pour des séries chronologiques multivariées à composantes non-négatives. En premier lieu, nous avons exposé les modèles GARCH et ACD qui constituent des cas particuliers des modèles MEM. Ainsi, nous avons présenté les modèles MEM et leurs généralisations vMEM introduits par Engle (2002) et développés par Engle et Gallo (2006). Pour estimer les modèles vMEM, nous avons exposé deux approches d'estimation présentées par Cipollini, Engle et Gallo (2006). La première approche proposée était l'approche paramétrique qui consistait à compléter le modèle avec la spécification de la fonction de densité conjointe pour le vecteur des termes d'erreurs. La méthode paramétrique utilisée était la méthode du maximum de vraisemblance. La deuxième approche était l'approche semi-paramétrique qui consistait à estimer les paramètres de la moyenne conditionnelle sans spécifier une distribution multivariée pour le terme d'erreur. Deux méthodes semi-paramétriques ont été présentées à savoir la méthode des moments généralisés (GMM) et la méthode des erreurs standard associées aux paramètres (WNLS).

En second lieu, des simulations Monte Carlo ont été réalisées pour analyser les résultats théoriques. Nous avons considéré sept modèles vMEM(1,1) pour deux tailles échantillonnelles différentes $T = 1000$ et $T = 5000$. Le but était de comparer les deux approches, paramétrique et semi-paramétrique, et de vérifier si la taille de l'échantillon a une influence sur les paramètres estimés. Les simulations ont montré que si le nombre d'observations T augmente le biais, l'erreur quadratique moyenne, la variance échantillonnale et la moyenne des estimateurs de variance diminuent. De plus, pour chaque T , au niveau des paramètres de la moyenne conditionnelle, la méthode GMM donne les plus petites variances échantillonnelles et moyennes des estimateurs de variances parmi les trois méthodes, au moins dans nos études empiriques.

En troisième lieu, nous avons appliqué les méthodes d'estimation à des données financières de trois indices boursiers S&P500, S&P400 et FTSE100. Pour chaque indice boursier, nous avons considéré une série multivariée formée par trois séries chronologiques qui sont le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas. Tout d'abord, nous avons estimé pour chaque indice boursier un modèle $vMEM(1,1)$ à l'aide des méthodes paramétrique et semi-paramétrique. Après avoir calculé les séries résiduelles estimées, nous avons constaté qu'elles n'étaient pas compatibles avec l'hypothèse d'indépendance. Ensuite, une deuxième analyse a été appliquée mais cette fois-ci avec un modèle $vMEM(2,1)$. Les résultats trouvés montrent que les paramètres estimés sont très proches des valeurs initiales et que les erreurs estimées du modèle $vMEM(2,1)$ sont maintenant compatibles avec l'hypothèse d'indépendance pour les trois indices étudiés.

BIBLIOGRAPHIE

- [1] Andersen, T. G., Bollerslev, T., Christoffersen, P. F. et Diebold, F. X. (2005), "Volatility forecasting". *NBER working paper* No. 11188.
- [2] Andersen, T. G., Bollerslev, T., Christoffersen, P. F. et Diebold, F. X. (2006), "Volatility and correlation forecasting". Dans : Elliott G., Granger C. W. J. et Timmermann A., Éditeurs, *Handbook of Economic Forecasting*, North-Holland : Amsterdam, pp. 777- 878.
- [3] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. et Shephard, N. (2008), "Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise". *Econometrica* **76**, 1481-1536.
- [4] Bauwens, L. et Giot, P. (2000), "The logarithmic ACD model : an application to the bid-ask quote process of three NYSE stocks". *Annales d'Économie et de Statistique* **60**, 117-149.
- [5] Bauwens, L. et Giot, P. (2001), *Econometric Modelling of Stock Market Intraday Activity*, Advanced Studies in Theoretical and Applied Econometrics, Kluwer : Dordrecht.
- [6] Berndt, E., Hall, B., Hall, R. et Hausman, J. (1974), "Estimation and Inference in Nonlinear Structural Models". *Annals of Economic and Social Measurement* **3**, 653-665.
- [7] Bollerslev, T. (1986), "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics* **31**, 307-327.
- [8] Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. et Roncalli, T. (2000), "Copulas for finance : A reading guide and some applications". Rapport technique, groupe de recherche opérationnelle, Crédit Lyonnais, Paris.
- [9] Brockwell, P. J. et Davis, R. A. (1991), *Time Series : Theory and Methods*, Springer-Verlag : New York.
- [10] Brownlees, C. T., Cipollini, F. et Gallo, G. M. (2011), "Multiplicative error models". *Econometrics Working Papers Archive* 2011_03.
- [11] Cheriyan, K. (1941), "A bivariate correlated gamma-type distribution function". *Journal of the Indian Mathematical Society* **5**, 133-144.

- [12] Chou, R. Y. (2005), "Forecasting financial volatilities with extreme values : The conditional autoregressive range (carr) model". *Journal of Money, Credit and Banking* **37**, 561-582.
- [13] Cipollini, F., Engle, R. F. et Gallo, G. M. (2006), "Vector multiplicative error models : representation and inference". *NBER Working Paper* No. 12690.
- [14] Cipollini, F., Engle, R. F. et Gallo, G. M. (2007), "A model for multivariate non-negative valued processes in financial econometrics". *Econometrics Working Papers Archive* wp200716.
- [15] Cipollini, F., Engle, R. F. et Gallo, G. M. (2012), "Semi parametric vector mem". *Journal of Applied Econometrics*, doi :10.1002/jae.2292.
- [16] Cipollini, F. et Gallo, G. M. (2010), "Automated variable selection in vector multiplicative error models". *Computational Statistics & Data Analysis* **54**, 2470-2486.
- [17] Diamond, D. W. et Verrecchia, R. E. (1987), "Constraints on short-selling and asset price adjustments to private information". *Journal of Financial Economics* **18**, 277-311.
- [18] Ding, Z., Granger, C. W. J. et Engle, R. F. (1993), "A long memory property of stock market returns and a new model". *Journal of Empirical Finance* **1**, 83-106.
- [19] Easley, D. et O'Hara, M. (1992), "Time and the process of security price adjustment". *The Journal of Finance* **19**, 69-90.
- [20] Embrechts, P., Lindskog, F. et McNeil, A. J. (2001). "Modelling dependence with copulas". Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich.
- [21] Engle, R. F. (2002), "New frontiers for ARCH models". *Journal of Applied Econometrics* **17**, 425-446.
- [22] Engle, R. F. et Gallo, G. M. (2006), "A multiple indicators model for volatility using intra-daily data". *Journal of Econometrics* **131**, 3-27.
- [23] Engle, R. F., Gallo, G. M. et Velucchi, M. (2012), "Volatility spillovers in East Asian financial markets : a MEM based approach". *The Review of Economics and Statistics* **94**, 222-233.
- [24] Engle, R. F. et Russell, J. R. (1994), "Forecasting transaction rates : the autoregressive conditional duration model". *NBER Working Papers* No. 4966.
- [25] Engle, R. F. et Russell, J. R. (1998), "Autoregressive conditional duration : a new model for irregularly spaced transaction data". *Econometrica* **66**, 1127-1162.
- [26] Engle, R. F., Gallo, G. M. et Velucchi, M. (2005), "An mem-based investigation of dominance and interdependence across markets". *The Computational Statistics and Data Analysis World Conference*, Cyprus.

- [27] Francq, C. et Zakoïan, J. M. (2009), *Modèles GARCH : structure, inférence statistique et applications financières*, Economica : Paris.
- [28] Glosten, L., Jaganathan, R. et Runkle, D. (1993), "On the relation between the expected value and volatility of the nominal excess return on stocks". *Journal of Finance* **48**, 1779-1801.
- [29] Gouriéroux, C. (1992), *Modèles ARCH et applications financières*, Economica : Paris.
- [30] Gouriéroux, C. (1995), *Séries temporelles et modèles dynamiques*, 2ième édition, Economica : Paris.
- [31] Hogg, R. V. et Craig, A. T. (1978), *Introduction to Mathematical Statistics*, 4ième édition, Macmillan : New York.
- [32] Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman & Hall : New York.
- [33] Johnson, N., Kotz, S. et Balakrishnan, N. (2000), *Continuous Multivariate Distributions*, John Wiley & Sons : New York.
- [34] Johnson, R. A. et Wichern, D. W. (1992), *Applied Multivariate Statistical Analysis*, 3ième édition, Prentice Hall : New Jersey.
- [35] Lee, S. et Hansen, B. (1994), "Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator". *Econometric Theory* **10**, 29-52.
- [36] Manganelli, S. (2002), "Duration, volume and volatility impact of trades". *Journal of Financial Markets* **8**, 377-399.
- [37] McNeil, A. J., Frey, R. et Embrechts, P. (2005), *Quantitative Risk Management : Concepts, Techniques, and Tools*, Princeton University Press, Princeton.
- [38] Nelsen, R. B. (1999), *An Introduction to Copulas*, Lecture Notes in Statistics **139**. Springer-Verlag : New York.
- [39] Noss, J. (2007), "The Econometrics of Optimal Execution in an order-driven market". Rapport technique, Nuffield College, Oxford University, mémoire de maîtrise.
- [40] Ramabhadran, V. (1951), "A multivariate gamma-type distribution". *Sankhyā, the Indian Journal of Statistics* **11**, 45-46.
- [41] Schmidt, T. (2006), *Coping With Copulas*, Risk Books, J. Rank, Editor.
- [42] Shumway, R. M. et Stoffer, D. S. (2006), *Time Series Analysis and Its Applications, with R Examples*, 2ième édition, Springer : New York.
- [43] Sklar, A. (1959), "Fonctions de répartition à n dimensions et leurs marges". *Publications de l'Institut de Statistique de l'Université de Paris* **8**, 229-231.
- [44] Tsay, R. S. (2010), *Analysis of Financial Time Series*, 3ième édition, Wiley : New York.

- [45] Wooldridge, J. M. (1994), "Estimation and Inference for Dependent Processes", Dans : Robert F. Engle et Daniel L. McFadden, Éditeurs, Handbook of Econometrics, Chapitre 45, Vol 4, Elsevier, New York, pp. 2639-2738.

Annexe A

Nous présentons ici le programme développé en R qui a été utilisé pour nos simulations.

```
library ( stats )
library ( MASS )
library ( rootSolve ) # pour la méthode de GMM#
library ( maxLik )    # pour la méthode du Log de vraisemblance#

#Fonction pour simuler le processus MEM.#

#Simulation des modèles 1, 2 et 3 est la suivante:

simulation <- fonction(N,phi,omega,alpha,beta0)
{
#p représente le nombre de variables#
p <- 3

#Initialisation des paramètres et des variables#
moy <- rep(0,p)
Y <- matrix( rep(0,N * p),ncol = 3 )
X <- matrix( rep(0,N * p),ncol = 3 )
mu_t <- matrix( rep(0,N * p),ncol = 3 )
phi <- c(phi[1], phi[2], phi[3])
omega <- c( omega[1], omega[2], omega[3] )
alpha <- matrix(c(0,0,0,alpha[1],alpha[2],0,0,0,alpha[3]),
               nrow =3, ncol =3)
beta0 <- c(beta0[1], beta0[2], beta0[3])
beta <- diag( beta0 )
I <- diag(3)
```

```

#Matrice de corrélation#
rho  <- matrix(c(1,0.37,0.21,0.37,1,0.18,0.21,0.18,1),3,3)

#Génère trois variables aléatoires centrés réduites iid#
for(i in 1:p)
{
Y[,i] = rnorm(N,0,1)
}

out  <- t(Y)

#Multiplier la décomposition Cholesky par le vecteur des v.a#
q    <- chol( rho ) %*% out

#Génère une copule normale#
U    <- pnorm(q,mean = 0, sd = 1)

#Extraction des innovations de la copule normale #

#L'espérance est égale à shape/rate=1#

epsilon1 <- qgamma(U[1,],shape = phi[1],rate = phi[1])
epsilon2 <- qgamma(U[2,],shape = phi[2],rate = phi[2])
epsilon3 <- qgamma(U[3,],shape = phi[3],rate = phi[3])

epsilon  <- matrix(c(epsilon1,epsilon2,epsilon3),ncol = 3)

mu0      <- solve ( I - alpha - beta ) %*% omega
mu_t[1,] <- omega + ( alpha + beta ) %*% mu0
X[1,]    <- mu_t[1,] * epsilon[1,]

for(i in 2:N)
{
mu_t[i,] <- omega + alpha %*% X[i-1,] + beta %*% mu_t[i-1,]
X[i,]    <- mu_t[i,] * epsilon[i,]
}

X

```

```
}

# Nombre d'observations n#
n    <- 1000
N    <- 2*n+1

#Nombre de variables p#
p    <- 3

#Nombre de réalisations asim#
asim <- 1000

#Méthode paramétrique en utilisant l'algorithme BHHH#

#Déclaration des paramètres et initialisations des paramètres#

#Matrice des paramètres estimés#
ML_BHHH      <- matrix(rep( NA,12*asim ),ncol = 12)

#Matrice des paramètres estimés et le nombre de convergence#
ML_BHHH1     <- matrix(rep( NA,13*asim ),ncol = 13)

#Matrice des variances échantillonnales #
varasympML   <- matrix(rep( NA ),nrow = 12 * asim,ncol = 12)

#Matrice totale des informations#
ML_BHHH_FINAL <- matrix(rep( NA,25*asim ),ncol = 25)

s <- 1
#Valeur initial du modèle 1
valeur_initial <- c( 4 , 5.5 , 3.1 , 0.0206 , 0.0772 , 0.1172 ,
                    0.1564 , 0.1283 , 0.2418 , 0.79 , 0.7795 , 0.645 )

#Valeur initial du modèle 2
valeur_initial<-c(27.9853, 28.6715, 36.3381, 0.0206,
                  0.0772, 0.1172, 0.1564, 0.1283, 0.2418, 0.79,
                  0.7795, 0.645)
```

```

#Valeur initial du modèle 3
valeur_initial<-c(1, 1, 1, 0.0206, 0.0772, 0.1172, 0.1564,
  0.1283, 0.2418, 0.79, 0.7795, 0.645)

while(s <= asim)
{

(...)
#Dans cette partie du code, disponible sur demande,
  les itérations où le code ne convergeait pas ont été retirées#

print      ( s )

set.seed ( s )

oo <- (12 * s) - 11

rr <- 12 * s

#Prendre les n dernières observations des 3 variables simulées#

simul      <- simulation(N,valeur_initial[1:3],valeur_initial[4:6]
  ,valeur_initial[7:9],valeur_initial[10:12])

maSerie <- simul[(n+2):N,]

X          <- t( maSerie )

mu_t       <- function( theta )
{
moy        <- matrix(rep( NA,n * 3 ),ncol = 3)
omega      <- c(theta[1],theta[2],theta[3])
alpha      <- matrix(c(0,0,0,theta[4],theta[5],0,0,0,theta[6])
  ,nrow = 3)
beta0      <- c(theta[7],theta[8],theta[9])
beta       <- diag( beta0 )
I          <- diag( 3 )
mu0        <- colMeans( maSerie ) #moyenne initiale de p variables#
moy[1,]    = omega + ( alpha + beta ) %*% mu0

for (t in 1:(n-1))

```

```

{
moy[t+1,] = omega + alpha %*% X[,t]+beta %*% moy[t,]
}
return( t( moy ) )
}

gradient_de_mu <- function( theta ) gradient( mu_t,theta )

fonction_densite <- function(i,t,phi,theta)
{
phi      <- c(phi[1],phi[2],phi[3])
mu       <- mu_t( theta )
fonction <- dgamma(X[i,t]/mu[i,t],shape = phi[i],rate = phi[i])
return( fonction )
}

#Fonction de répartition des epsilons#
repartition <- function( i,t,phi,theta )
{
phi <- c(phi[1],phi[2],phi[3])
mu  <- mu_t( theta )
r   <- pgamma(phi[i] * X[i,t] / mu[i,t],shape = phi[i],rate = 1)
return( r )
}

#Quantiles des epsilons#

vect_q_it <- function( i,t,phi,theta )
            qnorm(repartition( i,t,phi,theta ),0,1)

q_t      <- function( t,par )
            sapply(1:3,vect_q_it,t = t,phi = par[1:3],theta = par[4:12])

quantile <- function( par )
            sapply(1:3,vect_q_it,phi = par[1:3],theta = par[4:12])

Q      <- function( par )
{
quant <- quantile( par )
x     <- (t( quant ) %*% quant) / n
}

```

```
return( x )
}
```

```
D_Q <- function( par )
{
z <- diag (diag ( Q (par) ) )
return( z )
}
```

```
Rtilde <- function( par )
{
Q <- Q( par )
DQ <- solve(D_Q( par ) )
F <- DQ^(0.5) %*% Q %*% DQ^(0.5)
return( F )
}
```

```
#Calcul du LOG VRAISEMBLANCE#
```

```
somme <- function( par )
{
quant <- quantile( par )
s <- 0
for(i in ( 1:3 ) )
{
s <- s+log( t( quant[,i] ) %*% quant[,i] )
}
return( s )
}
```

```
copulacontribution <- function( par )
{
quant <- quantile( par )
som <- somme ( par )
cu <- - 0.5 * (log( det( t(quant) %*% quant ) ) - som )
return( cu )
}
```



```

marginalscontribution <- function( par )
{
  phi   <- par[1:3]
  theta <- par[4:12]
  t     <- 1:n
  l     <- matrix(rep(0,3),nrow = 3,ncol = n)
  for(i in (1:3))
  {
    l[i,] <-
      l[i,]+log(fontion_densite(i,t,phi,theta) / mu_t(theta)[i,t])
  }
  return( colSums( l ) )
}

log_vrais <- function( par )
  sum( copulacontribution(par) + marginalscontribution(par) )

#Calcul de dérivée du Log vrais.###
#Constante C#

C <- function( par )
{
  Q = Q      ( par )
  DQ = D_Q   ( par )
  R = Rtilde( par )
  c = solve( Q ) %*% DQ^(0.5) %*% Q %*% DQ^(0.5) %*% solve( Q )-
      solve( Q )+ diag(3) - solve( R ) + solve( DQ )-
      (solve( DQ ))^(0.5) %*% diag( diag( solve( Q ) %*%
      DQ^(0.5) %*% Q ) )
  return( c )
}

f_1 <- function( phi,theta )
{
  fonction <- matrix(rep(0,3),nrow=n,ncol=3)
  phi      <- c(phi[1],phi[2],phi[3])
  mu      <- mu_t(theta)
  for(i in (1:3))
  {
    fonction[,i] <- dgamma(X[i,] / mu[i,],shape = phi[i]+1,

```

```

        rate = phi[i])
    }
    return(fonction)
}

f_2 <- function( phi,theta )
{
mu <- mu_t( theta )
j <- matrix(rep(0,3),nrow=n,ncol=3)
for(i in (1:3))
{
j[,i] <- phi[i] * (X[i,]/mu[i,]-1)
}
return(j)
}

b <- function( par )
{
c <- C( par )
q <- quantile( par )
v <- matrix(rep(0,3),nrow = 3,ncol = n)
for(t in (1:n))
{
v[,t] <- -(c%%q[t,]) / dnorm(q[t,])
}
return( v )
}

a <- function( par )
{
phi <- par[1:3]
theta <- par[4:12]
f1 <- f_1( phi,theta )
f2 <- f_2( phi,theta )
a <- matrix(rep(0,3),nrow = n,ncol = 3)
b <- b( par )
mu <- mu_t( theta )
for(t in (1:n))
{
a[t,] <- (f1[t,] %% b[,t] + f2[t,]) / mu[,t]
}
}

```

```

return( a )
}

numerical.derivative <- function(phi,theta)
{
h   <- 1e-7
i   <- 1:3
t   <- 1:n
der <- ( repartition(i,t,phi[i]+h,theta) -
          repartition(i,t,phi[i],theta) ) / h
return(der)
}

p_1 <- function( par )
{
phi   <- par[1:3]
theta <- par[4:12]
p     <- matrix(rep(0,3),nrow = n,ncol = 3)
b     <- b( par )
der   <- numerical.derivative(phi,theta)
for( t in (1:n) )
{
p[t,] <- - diag( der[,t] ) %*% b[,t]
}
return( p )
}

p_2 <- function( phi,theta )
{
mu <- mu_t( theta )
p <- matrix(rep(0,3), nrow = n,ncol = 3)
for( i in (1:3))
{
u   <- X[i,] / mu[i,]
p[,i] <- log( phi[i] ) - digamma( phi[i] ) + log( u ) - u + 1
}
return( p )
}

gradient_vecteur <- function( par )

```

```

{
phi      <- par[1:3]
theta    <- par[4:12]
u        <- rep(0,12)
gradientt <- gradient_de_mu( theta )
p1       <- p_1( par )
p2       <- p_2( phi,theta )
a        <- a( par )
u[1:3]   <- colSums( p1 + p2 )
for(t in (1:n))
{
r        <- t + 2 * (t - 1)
v        <- t * 3
u[4:12] <- u[4:12] + t(gradientt[r:v,]) %*% a[t,]
}
return( u )
}

#OPG#
hessien_inverse<-function( par )
{
phi      <- par[1:3]
theta    <- par[4:12]
u        <- matrix(rep(0,12),nrow = n,ncol = 12)
gradientt<- gradient_de_mu( theta
p1       <- p_1( par )
p2       <- p_2( phi,theta )
a        <- a( par )
u[,1:3] <- ( p1+p2 )
for(t in (1:n))
{
r        <- t+2*(t-1)
v        <- t*3
u[t,4:12] <- t( gradientt[r:v,] ) %*% a[t,]
}
return( solve( t( u ) %*% u ) )
}

grvrais<-function( par )
{
phi      <- par[1:3]

```

```

theta      <- par[4:12]
u          <- matrix(rep(0,12),nrow = n,ncol = 12)
gradientt <- gradient_de_mu(theta)
p1        <-p_1( par )
p2        <-p_2( phi,theta )
a         <-a( par )
u[,1:3]   <-( p1+p2 )
for(t in (1:n))
{
r         <- t + 2 * ( t-1 )
v         <- t * 3
u[t,4:12] <- t( gradientt[r:v,]) %*% a[t,]
}
return( u )
}

hessien<-function( par )
{
phi       <- par[1:3]
theta     <- par[4:12]
u         <- matrix(rep(0,12),nrow = n,ncol = 12)
gradientt <- gradient_de_mu( theta )
p1        <-p_1( par )
p2        <-p_2( phi,theta )
a         <-a( par )
u[,1:3]   <-( p1+p2 )
for(t in (1:n))
{
r         <- t + 2 * ( t-1 )
v         <- t * 3
u[t,4:12] <- t( gradientt[r:v,]) %*% a[t,]
}
return( - ( t( u ) %*% u ) )
}

bhhh <- maxBHHH(log_vrais,grad = grvrais,
               hess = hessien,start = valeur_initial)

summary( bhhh )
hessien_inverse( summary(bhhh)$estimate[,1] )
ML_BHHH[s,]      <- summary( bhhh )$estimate[,1]

```

```

ML_BHHH1[s,]      <- c(ML_BHHH[s,],summary( bhhh )$code)
varasympML[oo:rr,] <- hessian_inverse(ML_BHHH[s,])
ML_BHHH_FINAL[s,] <- c( ML_BHHH1[s,],diag( varasympML[oo:rr,] ) )
print(ML_BHHH1[s,])

s = s+1

}

ML_BHHH

ML_BHHH1

varasympML

ML_BHHH_FINAL

#Méthode GMM#

#Matrice des paramètres #
GMM      <- matrix( rep(0,9*asim),ncol = 9 )

#Matrice des variances échantillonnales#
varasympGMM <- matrix( rep(0),nrow = 9*asim,ncol = 9 )

#Matrice totale des résultats#
GMMFINAL   <- matrix( rep(0,18*asim),ncol = 18 )

s <- 1

while(s <= asim)
{

(...)

Dans cette partie du code, disponible sur demande,
les itérations où le code ne convergeait pas ont été retirées

```

```

print      ( s )

set.seed ( s )

o <- ( 9 * s ) - 8
r <-  9 * s

#Prendre les n dernières observations des 3 variables simulées#

simul      <- simulation(N,valeur_initial[1:3],valeur_initial[4:6]
                        ,valeur_initial[7:9],valeur_initial[10:12])

maSerie <- simul[(n+2):N,]

X          <- t( maSerie )

mu_t       <- function( theta )
{
moy        <- matrix(rep( NA,n * 3 ),ncol = 3)
omega      <- c(theta[1],theta[2],theta[3])
alpha      <- matrix(c(0,0,0,theta[4],theta[5],0,0,0,theta[6])
                    ,nrow = 3)
beta0      <- c(theta[7],theta[8],theta[9])
beta       <- diag( beta0 )
I          <- diag( 3 )
mu0        <- colMeans( maSerie ) #moyenne initiale de p variables#
moy[1,]    = omega + ( alpha + beta ) %*% mu0

for ( t in 1:(n-1))
{
moy[t+1,] = omega + alpha %*% X[,t]+beta %*% moy[t,]
}
return( t( moy ) )
}

gradient_de_mu <- function( theta ) gradient( mu_t,theta )

#deux étapes#

#si sigma = identité #

```

```

g <- function( theta )
{
  b          <- matrix( c( 1,1,1 ),nrow = 1 )
  output     <- matrix( rep( 0 ),nrow = 9,ncol = 1 )
  gradient   <- gradient_de_mu( theta )
  mu         <- mu_t( theta )
  for(t in 1:n)
  {
    r        <- t + 2 * ( t - 1 )
    v        <- t * 3
    output <- output + t( gradient[r:v,] ) %*%
      solve( diag( mu[,t] ) ) %*% t( X[,t] / mu[,t] - b )
  }
  return( as.vector( output / n ) )
}

#Résoudre l'équation g#

ss          <- multiroot( g,valeur_initial[4:12] )

parametre <- ss$root

#Calcul de sigma pour passer à la deuxième étape#

sigma <- function( theta )
{
  u <- matrix( rep(0),nrow =3 ,ncol = 3 )
  mu <- mu_t( theta )
  for(t in 1:n)
  {
    a <- X[,t]/mu[,t]-rep( 1,3 )
    u <- u+( a %*% t( a ) )
  }
  return ( u / n )
}

#Si sigma = Sigma( parametre ) #

#Deuxième étape#

g1 <- function( theta )

```



```

{
b          <- matrix( c(1,1,1),nrow = 1 )
output     <- matrix( rep( 0 ),nrow = 9,ncol = 1 )
Sigma      <- sigma( parametre )
gradient   <- gradient_de_mu( theta )
mu         <- mu_t( theta )
for(t in 1:n)
{
r          <- t + 2 * ( t - 1 )
v          <- t * 3
output <- output + t ( gradient[r:v,] ) %*%
solve( diag( mu[,t] ) %*% Sigma ) %*% t( X[,t] / mu[,t] - b )
}
return( as.vector( output / n ) )
}

GMM[s,] <- multiroot( g1,valeur_initial [4:12] )$root

#Calcul de la variance asymptotique#

avarGMM <- fonction( theta )
{
output     <- matrix( rep( 0 ),nrow = 9,ncol = 9 )
gradient   <- gradient_de_mu( theta )
mu         <- mu_t ( theta )
Sigma      <- sigma( theta )
for(t in 1:n)
{
r          <- t + 2 * ( t - 1 )
v          <- t * 3
output <- output + t( gradient[r:v,] ) %*% solve( diag( mu[,t] )
          %*% Sigma %*% diag( mu[,t] ) ) %*% gradient[r:v,]
}
return( solve( output ) )
}

varasympGMM[o:r,] <- avarGMM( GMM[s,] )

GMMFINAL[s,]      <- c( GMM[s,],diag( varasympGMM[o:r,] ) )

print( GMMFINAL[s,] )

```

```

s = s+1

}

GMM
varasympGMM
GMMFINAL
#FIN METHODE GMM

#Méthode WNLS #
#Déclaration des paramètres et initialisations des paramètres #

WNLSparam <- matrix( rep( 0, 9 * asim ),ncol = 9 )
WNLSparam1 <- matrix( rep( 0, 11 * asim ),ncol = 11 )
varasyWNLS <- matrix( rep( 0 ),nrow = 9 * asim,ncol = 9 )
WNLSFINAL <- matrix( rep( 0, 20 * asim ),ncol = 20 )

s <- 1

while(s <= asim)
{
print      ( s )
set.seed ( s )

o <- ( 9 * s ) - 8
r <- 9 * s

oo <- (12*s)-11
rr <- 12*s

simul      <- simulation( N,valeur_initial[1:3],valeur_initial[4:6]
                        ,valeur_initial[7:9],valeur_initial[10:12] )

maSerie <- simul[ (n+2):N, ]

X          <- t( maSerie )

```

```

mu_t <- function( theta )
{
  omega <- c( theta[1],theta[2],theta[3] )
  alpha <- matrix( c(0,0,0,theta[4],theta[5],0,0,0,theta[6] )
                  ,nrow = 3,ncol = 3 )
  beta0 <- c(theta[7],theta[8],theta[9])
  beta <- diag( beta0 )
  I <- diag( 3 )
  #moyennes initiales des 3 variables #
  mu0 <- colMeans( maSerie )
  moy <- matrix( rep( 0,n * 3),ncol = 3 )
  moy[1,] <- omega+(alpha + beta) %*% mu0
  for (t in 1:(n-1))
  {
    moy[t+1,] = omega + alpha %*% X[,t] + beta %*% moy[t,]
  }
  return( t( moy ) )
}

gradient_de_mu<-function(theta) gradient( mu_t,theta )

#methode WNLS#
#Étape 1#
#Sigma=identité#

WNLS <- function( theta )
{
  mu <- mu_t( theta )
  d <- 0
  for(t in 1:n)
  {
    d <- d + t( X[,t]-mu[,t] ) %*% solve( diag( mu[,t] )
      %*% diag( mu[,t] ) ) %*% ( X[,t] - mu[,t] )
  }
  return( d/( 2 * n ) )
}

S <- 1e-6

min <- nlminb( valeur_initial[4:12],WNLS,lower = rep(S^2),
              upper = rep(1-S),control = list(maxit = 100))

```

```

#Étape 2#
#Sigma estimé#

sigma <- function( theta )
{
  u <- matrix( rep(0),nrow = 3,ncol = 3 )
  mu <- mu_t( theta )
  for(t in 1:n)
  {
    a <- X[,t] / mu[,t] - rep( 1,3 )
    u <- u + ( a %*% t( a ) )
  }
  return( u / n )
}

theta0 <- min$par

WNLS1 <- function( theta )
{
  mu <- mu_t( theta )
  Sigma <- sigma( theta0 )
  d <- 0
  for(t in 1:n)
  {
    d <- d + t( X[,t] - mu[,t] ) %*% solve( diag ( mu[,t] ) %*% Sigma
      %*% diag( mu[,t] ) ) %*% ( X[,t] - mu[,t] )
  }
  return( d / ( 2 * n ) )
}

min1 <- nlminb( valeur_initial[4:12],WNLS1,lower = rep(S^2),
  upper = rep(1-S),control = list(maxit = 100))

avarWNLSa <- function( theta )
{
  outputa <- matrix( rep(0),nrow = 9,ncol = 9 )
  gradient <- gradient_de_mu( theta )
  mu <- mu_t( theta )
  Sigma <- sigma( theta )
  for(t in 1:n)

```

```

{
r<-t + 2 * ( t-1 )
v<-t * 3
outputa<-outputa + t( gradient[r:v,] ) %*% solve( diag(mu[,t] )
          %*% Sigma %*% diag( mu[,t] ) ) %*% gradient[r:v,]
}
return( solve( outputa ) )
}

```

```

avarWNLSb <- function( theta )
{
outputb <- matrix( rep(0),nrow = 9,ncol = 9 )
gradient <- gradient_de_mu( theta )
mu      <- mu_t( theta )
Sigma   <- sigma( theta )
for(t in 1:n)
{
r <- t + 2 * ( t-1 )
v <- t * 3
outputb <- outputb + as.vector(t ( X[,t] - mu[,t] )
          %*% solve( diag(mu[,t]) %*% Sigma %*% diag(mu[,t]))
          %*% ( X[,t] - mu[,t] ) ) * t( gradient[r:v,] )
          %*% solve( diag(mu[,t] ) %*% Sigma
          %*% diag(mu[,t])) %*% gradient[r:v,]
}
return( outputb )
}

```

```

avarWNLS <- function( theta )
{
output <- avarWNLSa( theta ) %*% (avarWNLSb( theta ))
          %*% avarWNLSa( theta )
return( output )
}

```

```

WNLSparam[s,] <- min1$par

```

```

WNLSparam1[s,] <- c(min1$par,min$message,min1$message)

```

```

varasyWNLS[o:r,] <- avarWNLS( WNLSparam[s,] )

```

```

WNLSFINAL[s,]      <- c(WNLSparam1[s,],diag(varasyWNLS[o:r,]))

print( WNLSFINAL[s,] )

s = s+1

}

#FIN METHODE WNLS

WNLSparam
WNLSparam1
varasyWNLS
WNLSFINAL

# Pour les modèles 4,5 et 6, la forme de la matrice alpha
a été changé donc nous présentons juste la simulation,
la moyenne mu_t et les valeurs initiales

simulation <- fonction(N,phi,omega,alpha,beta0)
{
#p représente le nombre de variables#
p <- 3

#Initialisation des paramètres et des variables#
moy   <- rep(0,p)
Y     <- matrix( rep(0,N * p),ncol = 3 )
X     <- matrix( rep(0,N * p),ncol = 3 )
mu_t  <- matrix( rep(0,N * p),ncol = 3 )
phi   <- c(phi[1], phi[2], phi[3])
omega <- c( omega[1], omega[2], omega[3] )
alpha <- matrix( c( alpha[1], alpha[2], alpha[3], alpha[4],
alpha[5], 0,0,0, alpha[6] ), nrow = 3, ncol = 3)
beta0 <- c(beta0[1], beta0[2], beta0[3])
beta  <- diag( beta0 )
I     <- diag(3)

#Matrice de corrélation#
rho   <- matrix(c(1,0.37,0.21,0.37,1,0.18,0.21,0.18,1),3,3)

```

```

#Génère trois variables aléatoires centrés réduites iid#
for(i in 1:p)
{
Y[,i] = rnorm(N,0,1)
}

out <- t(Y)

#Multiplier la décomposition Cholesky par le vecteur des v.a#
q <- chol( rho ) %*% out

#Génère une copule normale#
U <- pnorm(q,mean = 0, sd = 1)

#Extraction des innovations de la copule normale #

#L'espérance est égale à shape/rate=1#

epsilon1 <- qgamma(U[1,],shape = phi[1],rate = phi[1])
epsilon2 <- qgamma(U[2,],shape = phi[2],rate = phi[2])
epsilon3 <- qgamma(U[3,],shape = phi[3],rate = phi[3])

epsilon <- matrix(c(epsilon1,epsilon2,epsilon3),ncol = 3)

mu0 <- solve ( I - alpha - beta ) %*% omega
mu_t[1,] <- omega + ( alpha + beta ) %*% mu0
X[1,] <- mu_t[1,] * epsilon[1,]

for(i in 2:N)
{
mu_t[i,] <- omega + alpha %*% X[i-1,] + beta %*% mu_t[i-1,]
X[i,] <- mu_t[i,] * epsilon[i,]
}

X

}

```

```

mu_t <- function( theta )
{
  omega <- c( theta[1],theta[2],theta[3] )
  alpha <- matrix( c(theta[4], theta[5], theta[6], theta[7],
theta[8], 0,0,0, theta[9] ), nrow = 3, ncol = 3)
  beta0 <- c(theta[10],theta[11],theta[12])
  beta <- diag( beta0 )
  I <- diag( 3 )
  #moyennes initiales des 3 variables #
  mu0 <- colMeans( maSerie )
  moy <- matrix( rep( 0,n * 3),ncol = 3 )
  moy[1,] <- omega+(alpha + beta) %*% mu0
  for (t in 1:(n-1))
  {
    moy[t+1,] = omega + alpha %*% X[,t] + beta %*% moy[t,]
  }
  return( t( moy ) )
}

#Valeur initial du modèle 4
valeur_initial<-c(27.0821, 28.1324, 35.8882, 0.0206, 0.0406,
0.0982, 0.1961, 0.1864, 0.0775, 0.0529, 0.0694, 0.1770,
0.6675, 0.6715, 0.6618)

#Valeur initial du modèle 5
valeur_initial<-c(24.5, 35.63, 12.98, 0.0206, 0.0406 ,0.0982,
0.1961, 0.1864, 0.0775, 0.0529, 0.0694, 0.1770, 0.6675, 0.6715,
0.6618)

#Valeur initial du modèle 6
valeur_initial<-c(1, 1, 1, 0.0206, 0.0406, 0.0982, 0.1961, 0.1864,
0.0775, 0.0529, 0.0694, 0.1770, 0.6675, 0.6715, 0.6618)

# Pour le modèle 7, les matrices alpha et beta sont diagonales
donc nous présentons juste la simulation,
la moyenne mu_t et les valeurs initiales
simulation <- fonction(N,phi,omega,alpha0,beta0)
{
  #p représente le nombre de variables#
  p <- 3

```



```

#Initialisation des paramètres et des variables#
moy      <- rep(0,p)
Y        <- matrix( rep(0,N * p),ncol = 3 )
X        <- matrix( rep(0,N * p),ncol = 3 )
mu_t     <- matrix( rep(0,N * p),ncol = 3 )
phi      <- c(phi[1], phi[2], phi[3])
omega    <- c( omega[1], omega[2], omega[3] )
alpha0   <- c( alpha0[1], alpha0[2], alpha0[3])
alpha    <-diag( alpha0 )
beta0    <- c(beta0[1], beta0[2], beta0[3])
beta     <- diag( beta0 )
I        <- diag(3)

#Matrice de corrélation#
rho      <- matrix(c(1,0.37,0.21,0.37,1,0.18,0.21,0.18,1),3,3)

#Génère trois variables aléatoires centrés réduites iid#
for(i in 1:p)
{
Y[,i] = rnorm(N,0,1)
}

out      <- t(Y)

#Multiplier la décomposition Cholesky par le vecteur des v.a#
q        <- chol( rho ) %*% out

#Génère une copule normale#
U        <- pnorm(q,mean = 0, sd = 1)

#Extraction des innovations de la copule normale #

#L'espérance est égale à shape/rate=1#

epsilon1 <- qgamma(U[1,], shape = phi[1], rate = phi[1])
epsilon2 <- qgamma(U[2,], shape = phi[2], rate = phi[2])
epsilon3 <- qgamma(U[3,], shape = phi[3], rate = phi[3])

epsilon  <- matrix(c(epsilon1,epsilon2,epsilon3),ncol = 3)

```

```

mu0      <- solve ( I - alpha - beta ) %**% omega
mu_t[1,] <- omega + ( alpha + beta ) %**% mu0
X[1,]    <- mu_t[1,] * epsilon[1,]

for(i in 2:N)
{
mu_t[i,] <- omega + alpha %**% X[i-1,] + beta %**% mu_t[i-1,]
X[i,]    <- mu_t[i,] * epsilon[i,]
}

X

}

mu_t <- function( theta )
{
omega   <- c( theta[1],theta[2],theta[3] )
alpha0  <- c( theta[4], theta[5], theta[6])
alpha   <-diag( alpha0 )
beta0   <- c(theta[7],theta[8],theta[9])
beta    <- diag( beta0 )
I       <- diag( 3 )
#moyennes initiales des 3 variables #
mu0     <- colMeans( maSerie )
moy     <- matrix( rep( 0,n * 3),ncol = 3 )
moy[1,] <- omega+(alpha + beta) %**% mu0
for ( t in 1:(n-1))
{
moy[t+1,] = omega + alpha %**% X[,t] + beta %**% moy[t,]
}
return( t( moy ) )
}

#Valeur initial du modèle 7
valeur_initial<-c(27.9853, 28.6715, 36.3381, 0.02, 0.077,
0.12, 0.61, 0.54, 0.4, 0.2, 0.1, 0.25)

```

Nous présentons ici le programme développé en R qui a été utilisé pour nos applications.

```

library(stats)
library(MASS)
#calculer le gradient et Résoudre un système non-linéaire#
library(rootSolve)
# Algorithme BHHH#
library(maxLik)
#test de Ljung-Box#
library(portes)
#Résoudre le système non-linéaire#
library(nleqslv)

SPCOMP<-read.table("C:\\ ...\\SP500.txt",header=T, sep="\t")
SP400 <-read.table("C:\\ ...\\SP400.txt",header=T, sep="\t")
FTSE100<-read.table("C:\\ ...\\FTSE.txt",header=T, sep="\t")

SP500_date          <- as.character(SPCOMP$date)
SP500_dates         <- as.Date(SP500_date,format='%Y%m%d')
SP500_return        <- SPCOMP$ret
SP500_absolute_return <- SPCOMP$ret.abs
SP500_realized_kernel_volatility <- SPCOMP$rk.vol
SP500_daily_range   <- SPCOMP$h1

SP400_date          <- as.character(SP400$date)
SP400_dates         <- as.Date(SP400_date,format='%Y%m%d')
SP400_absolute_return <- SP400$ret.abs
SP400_realized_kernel_volatility <- SP400$rk.vol
SP400_daily_range   <-SP400$h1
SP400_return        <-SP400$ret

FTSE100_date        <- as.character(FTSE100$date)
FTSE100_dates       <- as.Date(FTSE100_date,format='%Y%m%d')
FTSE100_return       <- FTSE100$ret
FTSE100_absolute_return <- FTSE100$ret.abs
FTSE100_realized_volatility<- FTSE100$rk.vol
FTSE100_daily_range <- FTSE100$h1

```

```

#Construire les graphiques#
par(mfrow=c(3,1))
plot(SP500_dates,SP500_absolute_return,type ="l")
plot(SP500_dates,SP500_realized_kernel_volatility,type ="l")
plot(SP500_dates,SP500_daily_range,type ="l")

par(mfrow=c(3,1))
plot(SP400_dates,SP400_absolute_return,type="l")
plot(SP400_dates,SP400_realized_kernel_volatility,type="l")
plot(SP400_dates,SP400_daily_range,type ="l")

par(mfrow=c(3,1))
plot(FTSE100_dates,FTSE100_absolute_return,type ="l")
plot(FTSE100_dates,FTSE100_realized_volatility,type ="l")
plot(FTSE100_dates,FTSE100_daily_range,type ="l")

Serie_SP500    <- as.matrix(cbind(SP500_absolute_return,
                                SP500_realized_kernel_volatility,SP500_daily_range),
                            length(SP500_return),3)
Serie_SP400    <- as.matrix(cbind(SP400_absolute_return,
                                SP400_realized_kernel_volatility,SP400_daily_range),
                            length(SP400_return),3)
Serie_FTSE100 <- as.matrix(cbind(FTSE100_absolute_return,
                                FTSE100_realized_volatility,FTSE100_daily_range),
                            length(FTSE100_return),3)

#Variable indicatrice pour étudier l'asymétrie des données#
EFFET    <- fonction(i,return)
{
OUT      <- c(0)
if(return[i] < 0)
{
OUT[i] <- 1
}
else
{
OUT[i] <- 0
}
}

```

```

    }
  OUT[i]
}

#Première analyse en utilisant vMEM(1,1)#

#Moyenne conditionnelle utilisée
pour les trois indices boursiers #

mu_t <- fonction( theta, return, maSerie )
{
  alpha      <- c(theta[1],theta[2],theta[3])
  alpha1     <- diag(alpha)
  gama      <- c(theta[4],theta[5],theta[6])
  gamma     <- diag(gama)
  bet       <- c(theta[7],theta[8],theta[9])
  betaetoile <- diag(bet)
  betaa     <- c(theta[7] - theta[1] - theta[4]/2,
                theta[8] - theta[2] - theta[5]/2,
                theta[9] - theta[3] - theta[6]/2)
  beta      <- diag(betaa)
  I         <- diag(3)
  betaetoile <- beta + alpha1 + gamma/2
  X         <- t(maSerie)
  mu0       <- rowMeans(X) #moyenne initiale#
  moy       <- matrix(rep(0,length(return)*3),ncol = 3)
  moy[1,]   <- mu0

  for (t in 2:length(return))
  {
    moy[t,] = mu0 + betaetoile %*% ( moy[t-1,] - mu0 )+
              alpha1 %*% ( X[,t-1] - moy[t-1,] ) + gamma %*%
              ( X[,t-1] * c(EFFET(t-1,return),EFFET(t-1,return),
                            EFFET(t-1, return) ) - moy[ t-1, ]/2 )
  }
  return( t( moy ) )
}

#Méthode semi-paramétrique GMM#

gradient_de_mu <- fonction(theta, return , maSerie)

```

```

        gradient(mu_t,theta, return=return, maSerie=maSerie )

g      <- function( theta, return, maSerie )
{
b      <- matrix(c(1,1,1),nrow = 1)
output <- matrix(rep(0),nrow = length(valeur_initial),
                ncol = 1)
X      <- t(maSerie)
gradient <- gradient_de_mu(theta, return , maSerie )
mu     <- mu_t(theta,return, maSerie )

for(t in 1:length(return))
{
r <- t+2*(t-1)
v <- t*3
output<- output + t( gradient[r:v,] )%% solve( diag(mu[,t]) )
                %% t( X[,t]/mu[,t]-b )
}
return( as.vector( output/length(return) ) )
}

sss <- multiroot( g,valeur_initial,return = return,
                maSerie = maSerie )

parametre <- sss$root

sigma <- function( theta, return, maSerie )
{
u <- matrix( rep(0), nrow = 3, ncol = 3 )
mu <- mu_t( theta, return, maSerie )
X <- t(maSerie)
for(t in 1:length(return))
{
a <- X[,t] / mu[,t] - rep(1,3)
u <- u + ( a %% t(a) )
}
return( u/length(return) )
}

```

```

}

g1 <- function(theta, return, maSerie)
{
  b      <- matrix(c(1,1,1),nrow = 1)
  output <- matrix(rep(0),nrow = length(valeur_initial)
                  ,ncol = 1)
  Sigma  <- sigma( parametre, return = return,
                  maSerie = maSerie )
  gradient <- gradient_de_mu( theta,return , maSerie )
  X        <- t(maSerie)
  mu       <- mu_t( theta,return, maSerie )
  for(t in 1:length(return))
  {
    r      <- t + 2 * (t-1)
    v      <- t*3
    output <- output + t( gradient[r:v,] ) %*%
                  solve( diag( mu[,t] ) %*% Sigma )%*%
                  t( X[,t]/mu[,t] - b )
  }
  return( as.vector( output/length( return) ) )
}

GMM <- multiroot(g1,valeur_initial,return = return,
                 maSerie = maSerie )

# Test de Ljung-Box #

# Calculons les résidus estimés
epsilon <- t(maSerie)/mu_t(GMM$root,return = return ,
                          maSerie = maSerie )

#Tester l'indépendance des résidus estimés
portest( t(epsilon), lag = c(1,2,6,12,22,32,46),
         test = "LjungBox", MonteCarlo = FALSE )

```

```

#Deuxième analyse en utilisant vMEM(2,1)#

#Les moyennes conditionnelles utilisées
pour les trois indices boursiers #

mu_t_SP500 <- fonction( theta, return = SP500_return,
                        maSerie = Serie_SP500 )
{
  alpha1 <- matrix(c(theta[1],theta[2],0,theta[3],theta[4],
                    theta[5],theta[6],theta[7],theta[8]),nrow=3,ncol=3)
  alpha2 <- matrix(c(theta[9],theta[10],0,theta[11],theta[12],
                    theta[13],0,theta[14],theta[15]),nrow=3,ncol=3)
  gama <- c(theta[16],theta[17],theta[18])
  gamma <- diag(gama)
  beta <- matrix(c(theta[19],theta[20],0,0,theta[21],0,0,0,
                    theta[22]),nrow=3,ncol=3)
  I <- diag(3)
  X <- t(maSerie)
  #moyenne des 3 variables initiale#
  mu0 <- rowMeans(X)
  omega <- ( I- ( beta + gamma/2 + alpha1 + alpha2 ) )
           %*% mu0
  moy <- matrix(rep(0,length(return)*3),ncol = 3)
  moy[1,] <- mu0
  moy[2,] <- omega + beta %*% moy[1,] + alpha1 %*% X[,1] +
            alpha2 %*% mu0 + gamma %*% (X[,1] *
            c(EFFET(1,return),EFFET(1,return),EFFET(1,return)))
  for (t in 3:length(return))
  {
    moy[t,] = omega + beta %*% moy[t-1,] + alpha1 %*% X[,t-1] +
              alpha2 %*% X[,t-2] + gamma %*% ( X[,t-1] *
              c(EFFET(t-1,return),EFFET(t-1,return),
                EFFET(t-1,return)))
  }
  return( t( moy ) )
}

```



```

}

mu_t_SP400 <- function(theta, return = SP400_return,
                      maSerie = Serie_SP400 )
{
alpha1 <- matrix(c(theta[1],0,0,theta[2],theta[3],theta[4],0,
                  theta[5],theta[6]),nrow = 3,ncol = 3)
alpha2 <- matrix(c(0,0,0,0,theta[7],theta[8],theta[9],
                  theta[10],0),nrow = 3,ncol = 3)
beta <- matrix(c(theta[11],theta[12],0,theta[13],theta[14],
                 theta[15],0,0,theta[16]),nrow = 3,ncol = 3)
I <- diag(3)
X <- t(maSerie)
#moyenne des 3 variables initiale#
mu0 <- rowMeans(X)
omega <- (I - ( alpha1 + alpha2 + beta ) ) %*% mu0
moy <- matrix(rep(0,length(return)*3),ncol = 3)
moy[1,] <- mu0
moy[2,] <- omega + alpha1 %*% X[,1] + alpha2 %*% mu0 +
           beta %*% moy[1,]
for (t in 2:(length(return)-1))
{
moy[t+1,]=omega + alpha1 %*% X[,t] + alpha2 %*% X[,t-1] +
           beta %*% moy[t,]
}
return( t( moy ) )
}

mu_t_FTSE100 <- function(theta, return = FTSE100_return,
                        maSerie = Serie_FTSE100 )
{
alpha1 <- matrix(c(theta[1],theta[2],0,theta[3],theta[4],
                  theta[5],0,theta[6],theta[7]),nrow = 3,ncol = 3)
alpha2 <- matrix(c(theta[8],0,theta[9],0,theta[10],theta[11],
                  0,theta[12],0),nrow = 3,ncol = 3)
beta <- matrix(c(theta[13],0,0,theta[14],theta[15],0,0,
                 theta[16],theta[17]),nrow = 3,ncol = 3)
I <- diag(3)
X <- t(maSerie)

```

```

#moyenne des 3 variables initiale#
mu0      <- rowMeans(X)
omega    <- (I - ( alpha1 + alpha2 + beta ) ) %*% mu0
moy      <- matrix(rep(0,length(return)*3),ncol = 3)
moy[1,]  <- mu0
moy[2,]  <- omega + alpha1 %*% X[,1] + alpha2 %*% mu0 +
           beta %*% moy[1,]
for (t in 2:(length(return)-1))
{
moy[t+1,] = omega + alpha1 %*% X[,t] + alpha2 %*% X[,t-1] +
           beta %*% moy[t,]
}
return( t( moy ) )
}

gradient_de_mu <- function(theta, mu_t)  gradient(mu_t,theta)

g      <- function( theta, maSerie, return, mu_t )
{
b      <- matrix(c(1,1,1),nrow = 1)
output <- matrix(rep(0),nrow = length(valeur_initial),
                 ncol = 1)
X      <- t(maSerie)
gradient <- gradient_de_mu(theta, mu_t)
mu     <- mu_t(theta)

for(t in 1:length(return))
{
r <- t+2*(t-1)
v <- t*3
output<- output + t( gradient[r:v,] )%*% ginv(diag(mu[,t]))
           %*% t( X[,t]/mu[,t]-b )
}
return( as.vector( output/length(return) ) )
}

ss<-nleqslv(valeur_initial,return = return,maSerie = maSerie,
            mu_t= mu_t, g, method = c("Broyden", "Newton"),

```

```

control = list(trace=1))

parametre<-ss$x

sigma <- function( theta, return, maSerie ,mu_t)
{
u <- matrix( rep(0), nrow = 3, ncol = 3 )
mu <- mu_t( theta )
X <- t(maSerie)
for(t in 1:length(return))
{
a <- X[,t] / mu[,t] - rep(1,3)
u <- u + ( a %*% t(a) )
}
return( u/length(return) )
}

g1 <- function(theta, return, maSerie, mu_t)
{
b <- matrix(c(1,1,1),nrow = 1)
output <- matrix(rep(0),nrow = length(valeur_initial),
ncol = 1)
Sigma <- sigma( parametre, return, maSerie ,mu_t)
gradient <- gradient_de_mu( theta,mu_t )
X <- t(maSerie)
mu <- mu_t( theta )
for(t in 1:length(return) )
{
r <- t + 2 * (t-1)
v <- t*3
output <- output + t( gradient[r:v,] ) %*% ginv( diag( mu[,t] )
%*% Sigma )%*% t( X[,t]/mu[,t] - b )
}
return( as.vector( output/length( return) ) )
}

sss<-nleqslv( valeur_initial, return = return,
maSerie = maSerie, mu_t = mu_t, g1,

```

```

method = c("Broyden", "Newton"),
control = list(trace = 1) )

# Test de Ljung-Box #

# Calculons les résidus estimés

epsilon <- t(maSerie)/mu_t( sss$x )

#Tester l'indépendance des résidus estimés
portest( t( epsilon ), lag = c(1,2,6,12,22,32,46),
        test = "LjungBox", MonteCarlo = FALSE )

# Méthode paramétrique utilisée dans les deux analyses #

fontion_densite <-
  function( i, phi, theta, return ,maSerie, mu_t )
  {
phi      <- c( phi[1],phi[2],phi[3] )
mu       <- mu_t( theta )
X        <- t(maSerie)
fonction <- dgamma( X[i,] / mu[i,], shape=phi[i],rate=phi[i] )
for( s in ( 1:length( return ) ) )
  {
if( fonction[s] == Inf)
  {
fonction[s] <- 10^(-12)
}
}
return( fonction )
  }

##fonction de répartition des épsilons##

repartition <- function( i, t, phi, theta, return ,maSerie, mu_t )
  {

```

```

phi <- c(phi[1],phi[2],phi[3])
mu <- mu_t( theta )
X <- t(maSerie)
r <- c(0)
r <-pgamma( phi[i]*X[i,t]/mu[i,t],shape = phi[i], rate = 1 )
if( r >= 1 )
{
r == 1-10^(-6)
}
if(r <= 0)
{
r == 10^(-6)
}
return( r )
}

repartitioni <- function(i,phi,theta, return ,maSerie, mu_t )
{
r <- rep(0,n)
phi <- c(phi[1],phi[2],phi[3])
mu <- mu_t(theta)
X <- t(maSerie)
r <- pgamma(phi[i]*X[i,]/mu[i,],shape = phi[i],rate = 1)

for(s in (1:length( return )))
{
if( r[s] >= 1 )
{
r[s] = 1 - 10^(-6)
}
if( r[s] <= 0 )
{
r[s] = 10^(-6)
}
}
return ( r )
}

##vecteurs des quantiles##

```

```

quantilei <- function( i, par, return , maSerie , mu_t )
{
  phi   <- par[1:3]
  theta <- par[4:length(valeur_initial)]
  t     <- qnorm(repartitioni(i,phi,theta,return ,maSerie, mu_t )
                ,0,1)
  return( t )
}

```

```

quantile<-function(par, return , maSerie , mu_t )
  sapply(1:3, quantilei , par = par, return = return,
         maSerie = maSerie, mu_t = mu_t)

```

```

Q <- function( par, return , maSerie , mu_t )
{
  quant <- quantile( par, return , maSerie , mu_t )
  x     <- (t(quant) %*% quant)/length( return )
  return(x)
}

```

```

D_Q <- function(par, return , maSerie , mu_t )
{
  z <- diag( diag( Q(par,return , maSerie , mu_t ) ) )
  return ( z )
}

```

```

Rtilde <- function ( par, return , maSerie , mu_t )
{
  Q <- Q(par, return , maSerie , mu_t )
  DQ <- solve( D_Q(par, return , maSerie , mu_t ) )
  F <- DQ^(0.5) %*% Q %*% DQ^(0.5)
  return(F)
}

```

```
#LOG VRAISEMBLANCE#
```

```
#Décomposition de contribution de copule
```

```

somme <- function(par, return , maSerie , mu_t )
{

```

```

quant <- quantile(par, return , maSerie , mu_t )
s      <- 0
for(i in (1:3))
{
s <- s + log( t( quant[,i] ) %*% quant[,i] )
}
return ( s )
}

copulacontribution <-
      function(par, return , maSerie , mu_t )
{
quant <- quantile(par, return , maSerie , mu_t )
som   <- somme(par , return , maSerie , mu_t )
cu    <- -0.5 * ( log( det( t(quant) %*% quant ) ) - som )
return(cu)
}

marginalscontributioni <-
      function(i,par, return , maSerie , mu_t )
{
phi   <- par[1:3]
theta <- par[4:length(valeur_initial)]
l     <- matrix(rep(0,3),nrow = 3, ncol = length(return))
for( t in ( 1:length(return) ) )
{
r <- fontion_densite(i,phi,theta, return , maSerie , mu_t )
if(r[t] <= 0)
{
r[t] = 10^(-6)
}
if(r[t] == Inf)
{
r[t] = 10^(-12)
}
return(r)
l[i,] <- l[i,] + log( r/mu_t(theta)[i,] )
}
return ( l )
}

```

```

}

marginalscontribution <- function(par, return , maSerie , mu_t )
  rowSums( sapply( 1:3, marginalscontributioni, par = par,
    return=return , maSerie=maSerie , mu_t=mu_t ))

log_vrais<-function(par, return , maSerie , mu_t )
  sum(copulacontribution( par, return , maSerie , mu_t ) +
    marginalscontribution( par, return , maSerie , mu_t ))

#Calcul du gradient###

#constante C#

C <- function(par, return , maSerie , mu_t )
{
  Q = Q(par, return , maSerie , mu_t )
  DQ = D_Q(par, return , maSerie , mu_t )
  R = Rtilde(par, return , maSerie , mu_t )
  c = solve( Q ) %*% DQ^(0.5) %*% Q %*% DQ^(0.5) %*% solve( Q )
    - solve( Q ) + diag(3) - solve( R ) + solve( DQ )
    - (solve( DQ ))^(0.5) %*% diag( diag( solve( Q )
    %*% DQ^(0.5) %*% Q ) )
  return( c )
}

gradient_de_mu <- function(theta,mu_t) gradient(mu_t,theta)

f_1 <- function (phi,theta, return , maSerie , mu_t )
{
  phi      <- c(phi[1],phi[2],phi[3])
  mu       <- mu_t( theta )
  X        <- t(maSerie)
  fonction <- matrix(rep(0,3),nrow = length(return), ncol = 3)
  for(i in (1:3))

```



```

{
fonction[,i] = dgamma(X[i,]/mu[i,],shape = phi[i]+1,
                    rate = phi[i])
}
return( fonction )
}

f_2 <- function( phi,theta , return , maSerie , mu_t )
{
mu <- mu_t(theta)
j <- matrix(rep(0,3), nrow = length(return), ncol = 3)
X<-t(maSerie)
for(i in (1:3))
{
j[,i] <- phi[i] * (X[i,]/mu[i,] - 1)
}
return( j )
}

b <- function( par, return , maSerie , mu_t )
{
c <- C(par, return , maSerie , mu_t )
q <- quantile(par, return , maSerie , mu_t )
v <- matrix(rep(0,3), nrow = 3, ncol = length(return))
for(t in (1:length(return)))
{
v[,t] = -( c %*% q[t,] )/dnorm( q[t,] )
}
return( v )
}

a <- function(par, return , maSerie , mu_t )
{
phi <- par[1:3]
theta <- par[4:length(valeur_initial)]
f1 <- f_1(phi,theta, return , maSerie , mu_t )
f2 <- f_2(phi,theta, return , maSerie , mu_t )
}

```

```

a      <- matrix(rep(0,3), nrow = length(return), ncol = 3)
b      <- b( par , return , maSerie , mu_t )
mu     <- mu_t( theta )
for(t in (1:length(return)))
{
a[t,] <- ( f1[t,] %*% b[,t] + f2[t,] )/mu[,t]
}
return( a )
}

numerical.derivative <-
      function(phi,theta,return ,maSerie ,mu_t )
{
der <- matrix( rep(0,3), nrow = 3, ncol = length(return) )
h   <- 1e-7
for(i in (1:3))
{
der[i,] <- ( repartitioni(i, phi + h, theta, return ,
      maSerie , mu_t )
      - repartitioni(i, phi, theta, return ,
      maSerie , mu_t ) ) / h
}
return( der )
}

p_1 <- function(par, return , maSerie , mu_t )
{
phi   <- par[1:3]
theta <- par[4:length(valeur_initial)]
p     <- matrix( rep(0,3), nrow = length(return), ncol = 3 )
b     <- b( par , return , maSerie , mu_t )
der   <- numerical.derivative(phi,theta,return ,maSerie ,mu_t )
for( t in (1:length(return)) )
{
p[t,]<- - diag(der[,t]) %*% b[,t]
}
return( p )
}

```

```

}

p_2 <- function(phi,theta, return , maSerie , mu_t )
{
mu <- mu_t( theta )
p <- matrix(rep(0,3), nrow = length(return), ncol = 3)
X<-t(maSerie)
for( i in (1:3))
{
u <- X[i,]/mu[i,]
for(s in(1:length(return)))
{
if(u[s] <= 0)
{
u[s] <- 10^(-6)
}
}
p[,i] = log( phi[i] ) - digamma( phi[i] ) +
log( u ) - u + 1
}
return( p )
}

gradient_vecteur <- function(par, return , maSerie , mu_t )
{
phi <- par[1:3]
theta <- par[4:length(valeur_initial)]
u <- rep(0,length(valeur_initial))
gradientt <- gradient_de_mu( theta,mu_t )
p1 <- p_1( par , return , maSerie , mu_t )
p2 <- p_2(phi,theta, return , maSerie , mu_t )
a <- a( par , return , maSerie , mu_t )
u[1:3] <- colSums( p1 + p2 )
for(t in (1:length(return)))
{
r <- t + 2 * ( t - 1 )
v <- t * 3
u[4:length(valeur_initial)] <- u[4:length(valeur_initial)]
+ t(gradientt[r:v,]) %*% a[t,]
}
}

```

```

}
return( u )
}

#OPG#

grvrais <- function(par, return , maSerie , mu_t )
{
phi      <- par[1:3]
theta    <- par[4:length(valeur_initial)]
u        <- matrix(rep(0,length(valeur_initial)),
                  nrow=length(return),ncol=length(valeur_initial))
gradientt <- gradient_de_mu(theta,mu_t)
p1       <- p_1( par, return , maSerie , mu_t )
p2       <- p_2(phi, theta, return , maSerie , mu_t )
a        <- a( par , return , maSerie , mu_t )
u[,1:3]  <- ( p1 + p2 )
for(t in (1:length(return)))
{
r <- t + 2 * ( t - 1 )
v <- t * 3
u[t,4:length(valeur_initial)] <- t( gradientt[r:v,] )
                                %*% a[t,]
}
return( u )
}

hessien<-function(par, return , maSerie , mu_t )
{
phi      <- par[1:3]
theta    <- par[4:length(valeur_initial)]
u        <- matrix(rep(0,length(valeur_initial)),
                  nrow=length(return),ncol=length(valeur_initial))
gradientt <- gradient_de_mu(theta,mu_t)
p1       <- p_1(par, return , maSerie , mu_t )
p2       <- p_2(phi,theta, return , maSerie , mu_t )
a        <- a(par, return , maSerie , mu_t )
u[,1:3]  <- (p1+p2)
for(t in (1:length(return)))

```

```

{
r <- t+2*(t-1)
v <- t*3
u[t,4:length(valeur_initial)] <- t(gradientt[r:v,])%*%a[t,]
}
return( - ( t( u ) %*% u ) )
}

bhhh <- maxBHHH(log_vrais, return = return , maSerie = maSerie ,
mu_t = mu_t,grad = grvrais, hess = hessien,
start = valeur_initial, print.level = 2)

summary( bhhh )

# Test de Ljung-Box #

# Calculer les résidus estimés

epsilon_parametrique <- X /
mu_t(summary( bhhh )$estimate[,1][4:length( valeur_initial )])

#Tester l'indépendance des résidus estimés

portest( t( epsilon_parametrique ),lag = c(1,2,6,12,22,32,46),
test = "LjungBox", MonteCarlo = TRUE )

```

Annexe B

Nous présentons dans cette partie les graphiques représentant les séries des trois variables relatives aux indices boursiers S&P500, S&P400 et FTSE100 : Le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas.

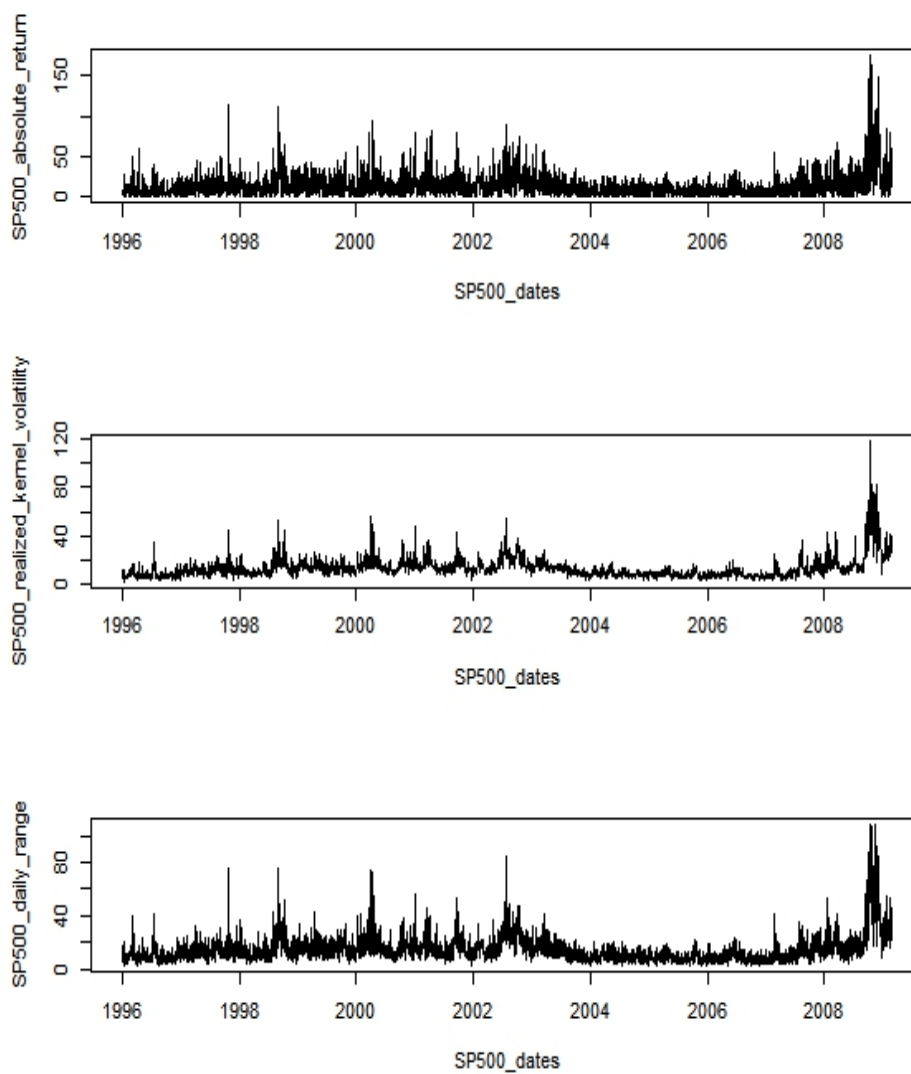


FIGURE B.1. Graphiques représentant les séries des trois variables relatives à S&P500 : Le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas.

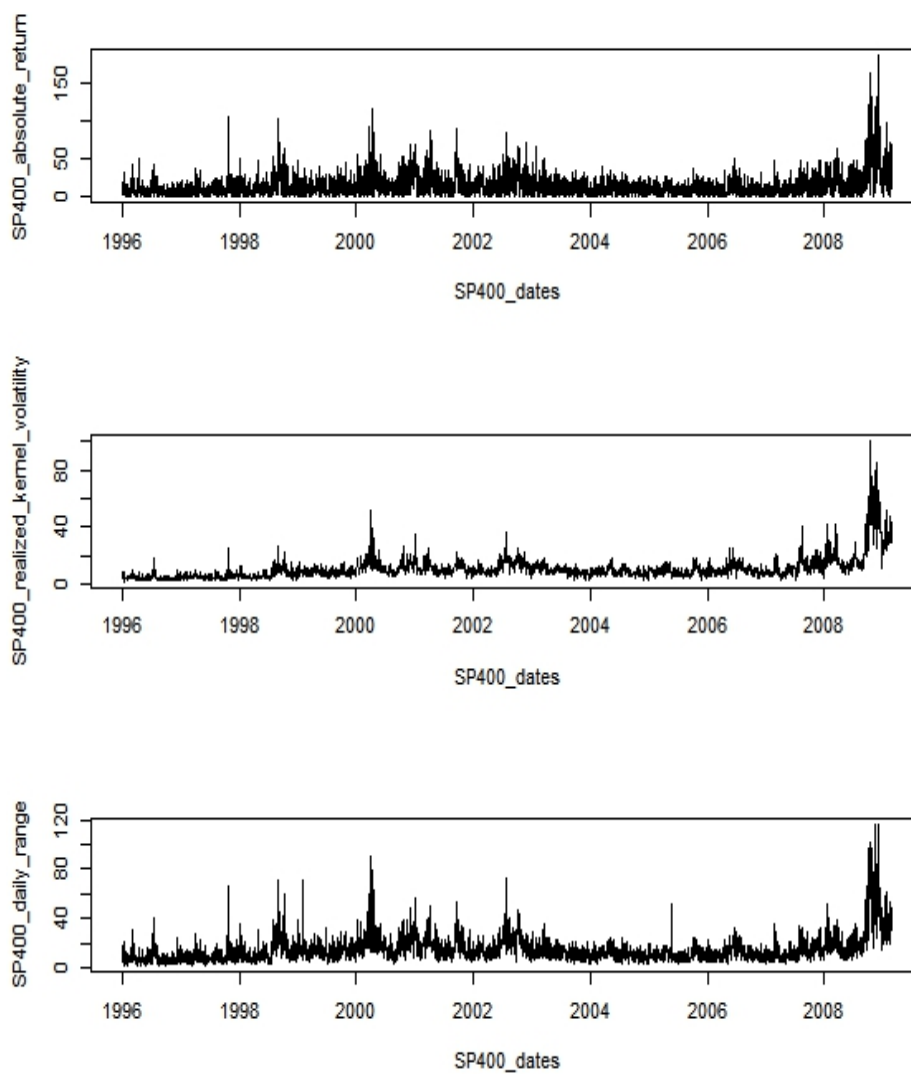


FIGURE B.2. Graphiques représentant les séries des trois variables relatives à S&P400 : Le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas.

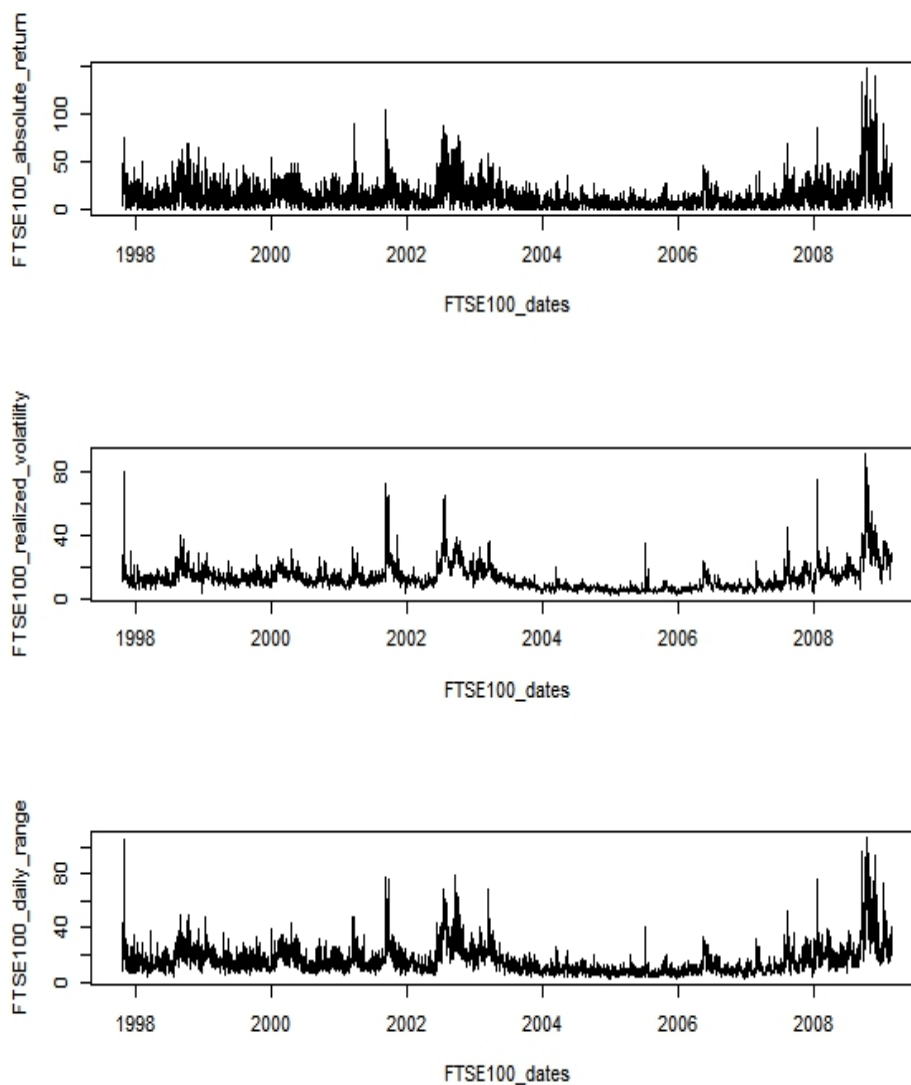


FIGURE B.3. Graphiques représentant les séries des trois variables relatives à FTSE100 : Le rendement absolu, la volatilité réalisée selon la méthode du noyau et la différence entre le prix le plus élevé et le prix le plus bas.