

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene



Faculté de Mathématiques

Département de Recherche Opérationnelle

Mémoire

**En vue de l'obtention du Diplôme de MASTER
Modélisation Stochastique et Prévision en Recherche Opérationnelle
(MSPRO)**

Thème

Modèles de séries chronologiques de comptage multivariées :
formulation INGARCH

Présenté par : FRAHI Mehdi et SEKIOU Roufaida

Soutenu le 27 Juin 2024, devant le jury composé de :

Président : MIHOUBI Miloud, USTHB
Rapporteur : SADOUN Mohamed Djemaâ, USTHB
Examinateur : ALIAT Billel, USTHB

Code Mémoire 07/MSPRO/24

RÉSUMÉ

Ce mémoire de master explore la modélisation des séries chronologiques de comptage multivariées, en se concentrant sur le modèle bivarié Poisson INGARCH (BP-INGARCH).

Après une revue de littérature sur les modèles existants, le mémoire détaille la formulation du modèle BP-INGARCH, incluant la distribution conditionnelle, et les conditions de stationnarité. Une nouvelle méthode d'estimation basée sur les Moindres Carrés Ordinaires bloc par bloc est proposée, équivalente aux Moindres Carrés Généralisés.

Une étude de simulation approfondie démontre la fiabilité et l'efficacité de cette méthode, ainsi que son utilité pour initialiser l'estimation par Maximum de Vraisemblance.

Enfin, l'application du modèle BP-INGARCH à des données réelles d'accidents de la route met en évidence sa capacité à capturer les dynamiques et corrélations des données, offrant des résultats pertinentes.

REMERCIEMENTS

Nous tenons à exprimer notre sincère reconnaissance à notre encadrant le Dr. SADOUN Mohamed. Sa disponibilité, sa patience et ses précieux conseils ont été déterminants dans la réussite de ce travail.

Nos remerciements vont également au Pr. MIHOUBI Miloud pour nous avoir fait l'honneur de présider le jury, et au Dr. ALIAT Billel pour avoir bien voulu examiner ce mémoire.

Ce travail n'aurait pas pu voir le jour sans le continual soutien de nos familles. Nous tenons à leur exprimer notre profonde gratitude pour leur patience et leurs sacrifices. Leur présence à nos côtés et leurs encouragements constants ont été une source de motivation précieuse tout au long de ce parcours.

TABLE DES MATIÈRES

Notations	2
Introduction	3
I Revue de littérature sur la classe des modèles INGARCH multivarié	5
I.1 Modèles proposés dans la littérature et domaines d'application	7
I.1.1 Distributions de Poisson Multivariées	8
I.1.1.1 Distributions de Poisson Bivariées	8
I.1.1.2 Distribution de Poisson Trivariée	10
I.1.1.3 Distribution de Poisson Multivariée	11
I.1.2 Les modèles INGARCH multivariés	13
I.1.2.1 Modèle Bivarié BINGARCH	13
I.1.2.2 Modèle Poisson Bivarié BP-INGARCH	13
I.1.2.3 Modèle Poisson Bivarié Généralisé GBP-INGARCH	14
I.1.2.4 Modèle Multivarié V-INGARCH	14
I.1.2.5 Modèle Multivarié M-INGARCH	16
I.2 Résultats obtenus concernant la structure probabiliste	17
I.2.1 Stationnarité et érgodicité	17
I.2.2 Autocovariance et autocorrélation	19
I.3 Inférence et procédures statistiques proposées	19
I.3.1 Estimation des paramètres	19
I.3.2 Tests d'hypothèses	24
I.3.2.1 Test basé sur le QMLE	24
I.3.2.2 Test basé sur le MDPDE	25
II Modèle INGARCH bivarié Poisson BP-INGARCH	26
II.1 Formulation du modèle bivarié BP-INGARCH	27
II.2 Théorie de la stationnarité	28
II.3 Estimation des paramètres	29
II.3.1 Estimation par la méthode des Moindres Carrés Ordinaire	29
II.3.1.1 Principe de la procédure d'estimation	29
II.3.1.2 Procédure d'estimation dans le cas bivarié	31
II.3.2 Estimation par la méthode de Maximum de Vraisemblance	35

II.3.2.1	La fonction de log-Vraisemblance	35
II.3.2.2	Les propriétés asymptotiques de l'estimateur de Maximum de Vraisemblance	37
III Illustration numérique		39
III.1	Étude de simulation	40
III.2	Application sur données réelles	63
IV Conclusion, discussion et perspectives		74
IV.1	Modèle Self excited Threshold BP-INGARCH	75
IV.2	Modèle BP-INGARCH Segmenté (Change Point BP-INGARCH)	76
A Annexe		78
A.1	Preuve du Théorème II.1	78
A.2	Méthodologie de ZHU, 2011 (II.3.1.1)	80
A.3	Estimation d'un VAR par MCO/MCG :	81
A.4	Preuve de l'équivalence entre MCG et MCO bloc par bloc (A.3), (Proposition II.1)	84
A.5	Algorithme pour générer des échantillons de la distribution de Poisson bivariée	85
A.6	Ajustement du VAR(p^*)	86
A.7	Adaptation de la méthode MCO pour le cas diagonal	86
A.8	Résultats d'estimation MV en Python	89
Bibliographie		90

TABLE DES FIGURES

I.1 Schéma représentant la classe de modèles étudiés	6
I.2 Schéma résumant les différentes constructions de la loi Poisson bivariée	12
II.1 Schéma résumant les étapes de la procédure d'estimation par MCO	36
III.1 Graphes des séries temporelles Y_1 (bleu), Y_2 (rouge), composantes de la série bivariée.	41
III.2 Graphes des séries temporelles bivarié des modèles M1-M4.	41
III.3 L'évaluation de MADE et MSE des estimateurs pour les modèles M1-M2	46
III.4 L'évaluation de MADE et MSE des estimateurs pour les modèles M3-M4	47
III.5 Boîte à moustaches M1.	53
III.6 Boîte à moustaches M2.	54
III.7 Boîte à moustaches M3.	55
III.8 Boîte à moustaches M4.	56
III.9 Boîte à moustaches M1.	59
III.10Boîte à moustaches M1.	61
III.11Graphes de la série temporelle bivariée des accidents diurnes et nocturnes.	64
III.12Les corrélations croisées, autocorrélations, et autocorrélations partielles	65
III.13Graphes des prédictions (modèle 1)	68
III.14Graphes des résidus pour chaque composante de la série bivariée, et leur distribution(modèle 1)	69
III.15Graphes des prédictions(modèle 2)	70
III.16Graphes des résidus pour chaque composante de la série bivariée, et leur distribution(modèle 2)	71
III.17fonctions ACF et PACF pour le modèle 1	72
III.18fonctions ACF et PACF pour le modèle 2	72
A.1 Un échantillon de la distribution bivarié BP	86

LISTE DES TABLEAUX

III.1 Résultats d'estimation MCO pour le modèle M1	42
III.2 Résultats d'estimation MCO pour le modèle M2.	43
III.3 Résultats d'estimation MCO pour le modèle M3.	44
III.4 Résultats d'estimation MCO pour le modèle M4.	45
III.5 Résultats d'estimation MV pour le modèle M1 avec des matrices non-diagonales.	49
III.6 Résultats d'estimation MV pour le modèle M2 avec des matrices diagonales.	50
III.7 Résultats d'estimation MV pour le modèle M3 avec A matrice diagonale et B non-diagonale.	51
III.8 Résultats d'estimation MV pour le modèle M4 avec A non-diagonale et B diagonale.	52
III.9 Résultats d'estimation MV avec données initiales aléatoires pour le modèle M1.	58
III.10 Tableau de contingence pour les données réelles avec Y_1 Diurne et Y_2 Nocturne	66
III.11 Résultats des MSE/MADE des résidus du Modèle 1 et Modèle 2 des deux séries résiduelles.	73
A.1 Résultats pour les équations Y1 et Y2	86
A.2 Résultats d'estimation sur Python	89

NOTATIONS

Symbole	Description
\mathcal{F}_{t-1}	Tribu engendrée par les observations précédentes jusqu'à l'instant $t - 1$
BP^*	Distribution de Poisson bivariée construite par réduction trivariée
BP	Distribution de Poisson bivariée construite par un facteur multiplicatif
$BP\bar{G}$	Distribution de Poisson bivariée BP généralisée
I_d	Matrice identité d'ordre d
0_d	Matrice nulle d'ordre d
$\rho(A)$	Rayon spectral de la matrice A
$\ A\ _p$	Norme p de la matrice A
$Cov(X, Y)$	Covariance entre les variables X et Y
$E[X]$	Espérance de la variable X
$\Gamma_\delta(h)$	Fonction d'autocovariance à l'ordre h
$Var(X)$	Variance de la variable X
MV	Maximum de Vraisemblance
QMV	Quasi Maximum de Vraisemblance
EMV	Estimateur de Maximum de Vraisemblance
$EQMV$	Estimateur de Quasi Maximum de Vraisemblance
MBP	Algorithme de Maximisation par Partie
$MMBP$	Algorithme de Maximisation par Partie Modifié
DPD	Density Power Divergence
$MDPDE$	Minimum Density Power Divergence
$CUSUM$	Somme Cumulative
TCL	Théorème Central Limite
$p.s.$	Presque sûrement
\xrightarrow{d}	Convergence en distribution
$\min_i(y_i)$	Minimum des éléments du vecteur y
$y_1 \wedge y_2$	Minimum entre y_1 et y_2
1_d	Vecteur colonne de 1 de dimension d
$\min(y_1, y_2)$	Minimum des deux entiers y_1 et y_2
$\binom{n}{k}$	Coefficient binomial
CCF	fonction de corrélation croisée (cross correlation function)
ACF	fonction d'autocorrélation (autocorrelation function)
$PACF$	fonction d'autocorrélation partielle (partial autocorrelation function)
$BLUE$	Best Linear Unbiased Estimator

INTRODUCTION

Depuis leurs apparitions dans la littérature de l'analyse et la modélisation des séries chronologiques de comptage au milieu des années 2000, (Article de **FERLAND et al., 2006**, Article de **ZHU et al. (2010)**), les modèles autorégressifs conditionnellement hétéroscédastique à valeurs entière (Poisson-INGARCH, Négative-Binomiale-INGARCH) ont suscité un intérêt croissant et majeur durant les deux décennies qui ont suivi. Cet intérêt réside dans le fait que ces modèles ont été appliqués dans plusieurs domaines, notamment en économie (**BLUNDELL et al., 2002**), en finance (**DIOP et KENGNE, 2017**), en criminologie (**DIOP et al., 2016 ; ZHU et al., 2010**), en épidémiologie et en médecine (**KANG et LEE, 2014 ; MAO et al., 2020 ; OUZZANI et BENTARZI, 2021**), etc.

Cet engouement autour de cette classe de modèle s'illustre aussi dans l'obtention de beaucoup de résultats théoriques comme par exemple l'étude de la stationnarité (stricte et faible), l'étude de l'ergodicité et l'ergodicité géométrique, l'existence des moments, (**AKNOUCHE et DEMMOUCHE, 2019 ; DOUKHAN et al., 2021 ; FERLAND et al., 2006 ; FRANCQ et ZAKOIAN, 2022**). L'inférence statistique de ces modèles a aussi été très développé et étoffé par plusieurs chercheurs dans la littérature tant sur l'aspect théorique que sur l'aspect algorithmique. (**DIOP et KENGNE, 2017 ; FRANCQ et ZAKOIAN, 2022 ; ZHU et al., 2010**).

Cependant, pour le cas particulier des séries chronologiques de comptage multivarié, la littérature est plutôt rare et pauvre comparé à celle du cas univarié. Cela dit, malgré le fait que les données de comptage multivariées sont présentes dans plusieurs disciplines différentes comme l'épidémiologie, le marketing, la criminologie et l'ingénierie, et autres (**KARLIS (2016)**). Sans prétendre à l'exhaustivité, nous pouvons citer par exemple quelques travaux de recherche dans l'économétrie des séries de dénombrement multivarié apparus dans la littérature durant cette dernière décennie :(**CUI et ZHU, 2018 ; FOKIANOS et al., 2020 ; 2022 ; H. LIU, 2012 ; M. LIU et al., 2023 ; LIVSEY et al., 2018 ; PAUL et al., 2008 ; RAVISHANKER et al., 2016**), et autres.

Dans la plupart des exemples traités dans le cas du comptage multivarié (voir : séries du nombre de transactions financières (**CUI et ZHU (2018)**) ; séries de comptage liée à du Marketing(**RAVISHANKER et al. (2016)**), séries médicales(**KARLIS (2016)**), série du nombre d'accidents de la route (**PEDELI et KARLIS (2011)**), série du nombre d'attaques criminelles (**CUI et ZHU (2018)**)), nous observons un fait stylisé marquant est que les données collectées sont des comptages corrélés observés à différents moments. Par conséquent,

nous avons deux sources de corrélation, la corrélation en série puisque les données sont des séries chronologiques, et la corrélation croisée puisque les séries chronologiques des composants sont corrélées à chaque instant. La nécessité de prendre en compte à la fois les corrélations en série « corrélation sérielles » et croisées complique la spécification, ainsi que l'estimation et l'inférence du modèle. Ceci génère des problèmes analytiques et algorithmiques très complexes, ce qui explique pourquoi la littérature sur les modèles statistiques pour les séries chronologiques multivariées de dénombrements n'est pas très fournie.

Néanmoins, de nouveaux modèles ont été développés ces dernières années pour faciliter l'analyse et la compréhension des phénomènes de comptage multivarié, en utilisant trois différentes approches à savoir : la méthode basée sur les copules, l'opération d'amincissement multivariée, la méthode basée sur une distribution conditionnelle multivarié appelé souvent dans la littérature les modèles INGARCH multivarié, dont nous allons discuter une brève revue dans le [Chapitre I](#) de ce mémoire. Dans notre travail, nous allons nous focaliser sur cette dernière formulation dite INGARCH de distribution conditionnelle de type Poisson multivariée avec les dimensions à priori tout en spécifiant la typologie de cette dernière.

L'organisation de notre mémoire est comme suit : nous allons dans un premier temps passer en revue les modèles INGARCH multivariés introduits dans la littérature ainsi que les approches utilisées pour l'inférence statistique et enfin les champs d'application de ces modèles. Dans ce passage aussi nous discutons le volet structure probabiliste concernant cette classe de modèles. En second lieu, nous spécifions une formulation bien précise à l'intérieur de cette classe multivarié-INGARCH à deux dimensions et de type Poisson à savoir le modèle bivarié Poisson INGARCH, noté BP-INGARCH. Nous allons donc étudier à la fois : 1) la théorie de stationnarité, 2) l'estimation des paramètres avec deux méthodes d'estimations différente incluant la méthode des moindres carrées ordinaire (MCO) ainsi que la méthode du maximum de vraisemblance (MV). En dernier lieu, nous allons d'abord effectuer des illustrations numériques artificielles avec une étude de simulation intensive pour plusieurs cas de modèles en examinant et évaluant plusieurs méthodes et approches. Ensuite, nous terminons par des illustrations numériques réelles avec une application de plusieurs modèles sur des données étrangères concernant le nombre d'accident de la route par jour et par nuit dans la région de Schiphol aux Pays-Bas durant l'année 2001. Nous soulignons qu'une étude de comparaison entre les différents modèles sera réalisée ce qui permet de montrer la qualité et le gain rapporté par ce type de formulation multivariée INGARCH.

Comme perspective, nous essaierons de formuler et de proposer une extension évolutive à travers le temps permettant de définir une version à changement de régime de cette classe de modèles.

Toutes les démonstrations ainsi que certains détails de calculs sont rapportés dans la partie [annexe A](#).

REVUE DE LITTÉRATURE SUR LA CLASSE DES MODÈLES INGARCH MULTIVARIÉ

Les modèles de séries de comptage multivariées ont une large gamme d'applications, que ce soit dans les domaines économiques et financiers ou plutôt dans la médecine, épidémiologie ou écologie.

Voici quelques exemples d'application ([KARLIS, 2016](#)) :

- Épidémiologie : chaque patient avec un certain symptôme est enregistré, puis pour traquer plusieurs symptômes à la fois dans le temps, vient le besoin d'utiliser des séries (de comptage) multivariées.
- Finance : les offres et demandes d'une certaine action cotée en bourse sont modélisées à travers une série bivariée, et une série multivariée est employée pour modéliser le nombre de transactions pour plusieurs actions à la fois.
- Géophysique : plusieurs séries temporelles sont générées à partir des tremblements de terre dans plusieurs endroits, leur étude à travers la structure multivariée nous permet d'étudier la relation entre ces endroits.
- Criminologie : un type de crime est étudié à travers le temps, incorporer d'autres types dans une analyse demande l'utilisation de séries multivariées.
- Marketing : le nombre d'achats pour un produit est modélisé à travers le temps, la modélisation de plusieurs produits similaires conduit à la modélisation par séries multivariées.

La plupart du temps, on commence par suivre un événement particulier sur une période donnée, puis on examine les événements adjacents et similaires. Pour les étudier simultanément, on utilise des séries multivariées.

La similarité entre les séries se traduit par une corrélation. On distingue l'autocorrélation de chaque composante, due à la structure temporelle, et la corrélation croisée entre les différentes séries à un instant donné. Ces corrélations complexifient la spécification, l'estimation et l'inférence des modèles.

Plusieurs approches ont été développées pour modéliser les séries temporelles multiva-

riées de comptage, reposant sur des principes très différents.

Le modèle univarié INAR(1) a été proposé par **AL-OSSH et ALZAID (1987)** puis étendu à un modèle INAR(p) par **ALZAID et AL-OSSH (1990)**, les deux étant basés sur l'opérateur d'amincissement.

Ce dernier modèle peut être à son tour généralisé vers le cas multivarié pour obtenir la nouvelle classe de modèles MINAR (**LATOUE, 1997**).

Dans l'analyse des séries de comptage multivariées, l'un des obstacles est le choix de la distribution jointe, comme le note **FOKIANOS et al. (2020)**. Une manière de surmonter cet obstacle est la modélisation basée sur les copules, également proposée et développée par **FOKIANOS et al. (2020)**.

Dans notre cas, nous nous intéresserons à l'approche visant à proposer une loi jointe, et donc à des modèles basés sur la distribution conditionnelle de la série, que ce soit dans le cas bivarié, largement traité dans la littérature, ou dans la possible généralisation vers des dimensions supérieures.

Nous nous intéressons en particulier à la classe INGARCH, dont le cas univarié est bien étudié dans la littérature, contrairement au multivarié.

Nous commencerons alors par passer en revue certains modèles déjà proposés.

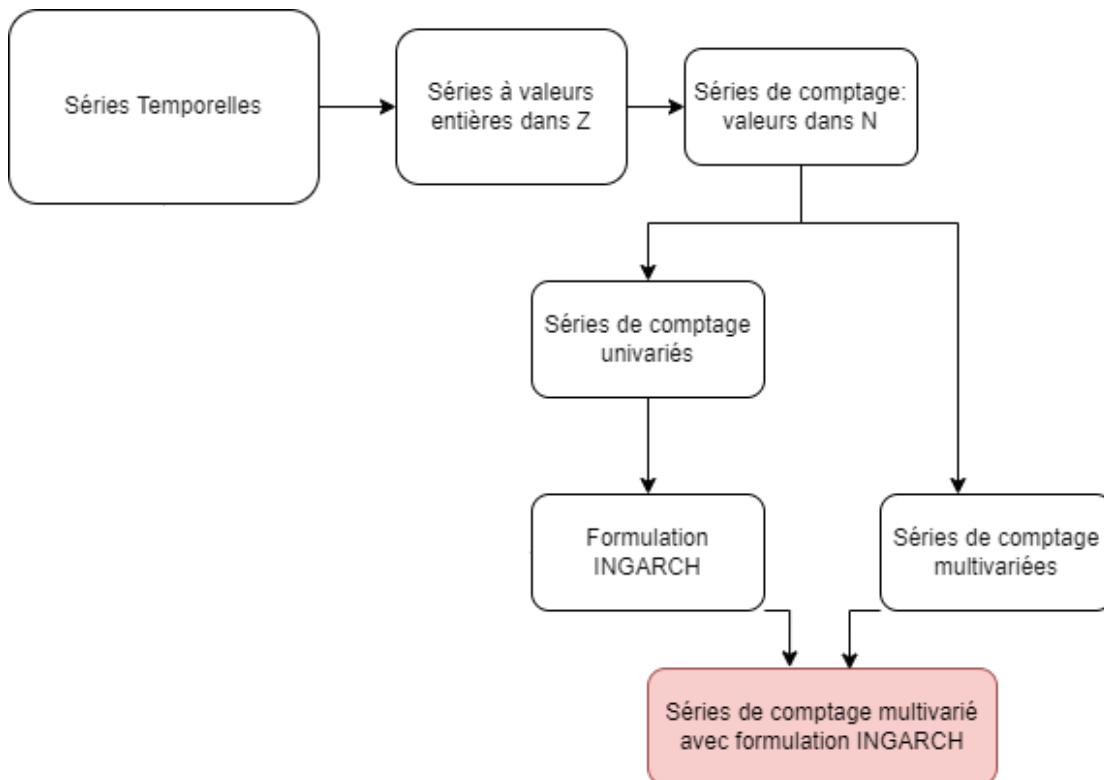


FIGURE I.1 – Schéma représentant la classe de modèles étudiés

I.1 Modèles proposés dans la littérature et domaines d'application

Il existe une variété de modèles proposés pour les séries de comptage, avec une diversité d'approches en modélisation.

Plusieurs auteurs ont contribué à la modélisation bivariée et multivariée, dont **Fokianos** (*ARMILLOTTA et FOKIANOS, 2023; FOKIANOS, 2021; FOKIANOS et al., 2020; 2022*), **Zhu et Cui** (*Cui et ZHU, 2018; Cui et al., 2020*), **Karlis** (*KARLIS, 2016*), **Weiß** (*SCOTTO et al., 2015; WEIß, 2018*), **Sangyeol Lee** (*S. LEE et al., 2023; Y. LEE et LEE, 2019; Y. LEE et al., 2018*), entre autres.

Leurs travaux ont enrichi la littérature en générant de nombreux développements et extensions récentes.

De façon générale, on peut voir le **INGARCH** comme une loi conditionnelle. Nous avons des observations : une série temporelle \mathbf{Y}_t supposée être générée par un certain processus. Conditionnellement à son passé (la tribu engendrée par les observations précédentes), l'observation à un instant t suit une certaine loi avec une certaine variance, laquelle est modélisée par un ARMA. D'où la spécification "GARCH" (combinaison linéaire des variances et des observations retardées). Puisque nous sommes dans le cas multivarié, les coefficients sont des matrices.

Le modèle **INGARCH** multivarié d'ordres p et q peut s'écrire de façon générale comme suit :

$$\begin{aligned} \mathbf{Y}_t | \mathcal{F}_{t-1} &\sim \mathcal{L}(\cdot) \\ \sigma_t &= \omega + \sum_{i=1}^p B_i \mathbf{Y}_{t-i} + \sum_{j=1}^q A_j \sigma_{t-j} \end{aligned} \tag{I.1}$$

où \mathbf{Y}_t est un vecteur d -dimensionnel d'observations à l'instant t ,

\mathcal{L} est une loi discrète quelconque et σ sa variance,

ω un vecteur d -dimensionnel,

B_i et A_j ($(i, j) \in \{1, \dots, p\} \times \{1, \dots, q\}$) des matrices carrées de taille $d \times d$.

Dans notre cas, nous nous intéressons à la distribution **Poisson**.

I.1.1 Distributions de Poisson Multivariées

Nous verrons trois types de distributions : le cas bivarié, trivarié, et plus généralement multivarié.

Parmi les distributions discrètes multivariées avec support positif (donc comptage), il en existe autres que Poisson, on peut citer notamment la distribution Multinomiale, Binomiale négative, Multinomiale négative, ou Hypergéométrique. Celles-ci ont leurs utilisations dans des cas de figure particuliers.

Néanmoins, nous nous limiterons à la distribution de Poisson en raison de son utilisation naturelle dans le cadre du comptage, de la littérature disponible, et de la portée restreinte de ce travail.

I.1.1.1 Distributions de Poisson Bivariées

La distribution de Poisson bivariée permet de modéliser conjointement deux variables aléatoires, chacune suivant une loi de Poisson. Cette extension, initiée après l'introduction de la distribution univariée au 19ème siècle, a émergé à partir du milieu du 20ème siècle ([CAMPBELL, 1934](#)), se développant ensuite grâce à des formulations mathématiques plus concrètes et offrant des méthodes statistiques pour l'analyse des données de comptage ([HOLGATE, 1964](#); [TEICHER, 1954](#)).

Parmi les différentes constructions de distributions de Poisson bivariées, nous nous intéresserons à trois cas en particulier qui ont vu une utilisations dans la modélisations des séries temporelles de comptage (spécialement le INGARCH multivarié) :

I.1.1.1.1 Distribution Poisson Bivariée BP* (réduction trivariée) :

BP* est une extension de la distribution de Poisson bivariée basée sur la méthode de réduction trivariée (*trivariate reduction*) qui consiste à utiliser trois variables aléatoires indépendantes pour construire une paire de variables aléatoires dépendantes suivant une distribution jointe bivariée ([CHESNEAU et al., 2015](#) et [SARABIA ALEGRÍA, GÓMEZ D'ENIZ et al., 2008](#)).

La distribution de Poisson bivariée BP* est utilisée par [H. LIU \(2012\)](#) et [Y. LEE et al. \(2018\)](#) pour modéliser et interpréter des séries temporelles multivariées (bivariées).

On considère X_k , $k = 1, 2, 3$ une variable aléatoire qui suit une distribution de Poisson avec paramètres $\lambda_1 - \phi$, $\lambda_2 - \phi$, ϕ respectivement :

$$\begin{aligned} X_1 &\sim \mathcal{P}(\lambda_1 - \phi) \\ X_2 &\sim \mathcal{P}(\lambda_2 - \phi) \\ X_3 &\sim \mathcal{P}(\phi) \end{aligned}$$

le schéma pour cette méthode est définit par :

$$Y_1 = X_1 + X_3$$

$$Y_2 = X_2 + X_3$$

avec Y_1, Y_2 deux variables jointes qui suivent une Poisson Bivariée $\mathbf{BP}^*(\lambda_1, \lambda_2, \phi)$:

$$(Y_1, Y_2) \sim \mathbf{BP}^*(\lambda_1, \lambda_2, \phi)$$

avec une fonction de masse définie par :

$$P(Y_1 = y_1, Y_2 = y_2) = e^{-(\lambda_1 + \lambda_2 - \phi)} \frac{(\lambda_1 - \phi)^{y_1} (\lambda_2 - \phi)^{y_2}}{y_1! y_2!} \times \sum_{k=0}^{\min(y_1, y_2)} \binom{y_1}{k} \binom{y_2}{k} k! \left(\frac{\phi}{(\lambda_1 - \phi)(\lambda_2 - \phi)} \right)^k \quad (\text{I.2})$$

où y_1, y_2 des entiers non négatifs et $\lambda_1 - \phi, \lambda_2 - \phi$ et ϕ des paramètres positifs, $\phi \in [0, \min \lambda_{t1}, \lambda_{t2}]$.

La structure de (I.2) permet de calculer la dépendance entre Y_1 et Y_2 à partir de la fonction de covariance ou de corrélation. La covariance entre Y_1 et Y_2 est donnée par :

$$\text{cov}(Y_1, Y_2) = \text{cov}(X_1 + X_3, X_2 + X_3) = \phi$$

on note que : $E(Y_i) = \text{Var}(Y_i) = \lambda_i$ pour $i = 1, 2$.

I.1.1.1.2 Distribution Poisson Bivariée BP (facteur multiplicatif) :

Cette distribution est proposée par LAKSHMINARAYANA et al. (1999) comme un produit de marginales Poisson avec un facteur multiplicatif (*multiplicative factor*) permettant capturer les corrélations négatives et positives.

La fonction de masse pour la distribution **BP** est donnée par :

$$P(Y_1 = y_1, Y_2 = y_2) = \frac{\lambda_1^{y_1} \lambda_2^{y_2}}{y_1! y_2!} \exp \{-(\lambda_1 + \lambda_2)\} [1 + \delta(e^{-y_1} - e^{-c\lambda_1})(e^{-y_2} - e^{-c\lambda_2})]$$

avec $c = 1 - e^{-1}$ et λ_1, λ_2 des paramètres positifs.

La distribution est dénotée $BP(\lambda_1, \lambda_2, \delta)$ et les marginales de Y_1 et Y_2 sont des Poisson de paramètres λ_1 et λ_2 respectivement.

Cui et al. (2020) indique que le choix de la fonction :

$$1 + \delta(e^{-y_1} - e^{-c\lambda_1})(e^{-y_2} - e^{-c\lambda_2})$$

est fait pour lier entre les distributions marginales de Y_1 et Y_2 .

$$E(Y_1) = \text{Var}(Y_1) = \lambda_1, E(Y_2) = \text{Var}(Y_2) = \lambda_2.$$

La matrice de covariance est donnée par :

$$\begin{pmatrix} \lambda_1 & \delta c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1 + \lambda_2)} \\ \delta c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1 + \lambda_2)} & \lambda_2 \end{pmatrix}$$

D'où le coefficient de corrélation est $\rho = \delta c^2 \sqrt{\lambda_1 \lambda_2} e^{-c(\lambda_1 + \lambda_2)}$, ce coefficient peut être positif, négatif ou bien nul, d'après la valeur du paramètre δ , notant que :

$$|\delta| \leq \frac{1}{(1 - e^{-c\lambda_1})(1 - e^{-c\lambda_2})}$$

I.1.1.1.3 Distribution Poisson Bivariée Généralisée *GBP* (facteur multiplicatif flexible)

Motivé par [LAKSHMINARAYANA et al. \(1999\)](#), en se basant sur le concept de facteur multiplicatif, une partie de la distribution précédente (ce facteur en question) est modifiée de trois façons différentes par [Cui et al. \(2020\)](#) pour étudier et modéliser un nouveau modèle **INGARCH** bivarié.

La fonction de masse en général s'écrit comme :

$$P(Y_1 = y_1, Y_2 = y_2) = \frac{1}{Z(\lambda_1, \lambda_2, \theta)} \frac{\lambda_1^{y_1} \lambda_2^{y_2}}{y_1! y_2!} \exp\{-(\lambda_1 + \lambda_2)\} c(F_1(y_1), F_2(y_2)) \quad (\text{I.3})$$

Avec $Z(\lambda_1, \lambda_2, \theta)$ le '*facteur de normalisation*' qui se définit par :

$$Z(\lambda_1, \lambda_2, \theta) := \sum_{y_1=0}^{\infty} \sum_{y_2=0}^{\infty} \frac{\lambda_1^{y_1} \lambda_2^{y_2}}{y_1! y_2!} \exp\{-(\lambda_1 + \lambda_2)\} c(F_1(y_1), F_2(y_2))$$

et θ le '*paramètre de dépendance*'.

$c(F_1(y_1), F_2(y_2))$ peut décrire l'un des trois facteurs :

$$c_\rho = \frac{1}{\sqrt{1 - \rho^2}} \exp\left(-\frac{\rho^2(q_1^2 + q_2^2) - 2\rho q_1 q_2}{2(1 - \rho^2)}\right) \quad (\text{I.4})$$

$$c_\gamma = \frac{-\gamma(e^{-\gamma} - 1)e^{-(u_1 + u_2)\gamma}}{[(e^{-\gamma} - 1) + (e^{-u_1\gamma} - 1)(e^{-u_2\gamma-1})]^2} \quad (\text{I.5})$$

$$c_\sigma = 1 + \sigma(1 - 2u_1)(1 - 2u_2) \quad (\text{I.6})$$

$q_i = \Phi^{-1}(u_i)$, Φ^{-1} est l'inverse de la distribution normale univariée,

et $u_i \in [0, 1]$, $i = 1, 2$. $\rho \in (-1, 1)$, $\gamma \in (-\infty, +\infty) \setminus \{0\}$, $\sigma \in [-1, 1]$

Les distributions sont alors notées respectivement : BPG (facteur Gaussien), BPF (facteur Frank), et BPFGM (facteur Farlie–Gumbel– Morgenstern).

GBP (généralisé) désigne l'une des trois distributions.

I.1.1.2 Distribution de Poisson Trivariée

Parmi les méthodes utilisées pour construire la Poisson trivariée sont : la méthode de convolution de n-distributions de Bernoulli trivariées indépendantes et identiques pour définir la distribution Binomiale trivariée, utilisée par [KAWAMURA \(1976\)](#) et la méthode "*The four parameter trivariate Poisson*" étudiée par [LOUKAS et PAPAGEORGIOU \(1991\)](#). Cette méthode est un choix pragmatique et populaire en raison de sa simplicité et de son

interprétabilité. Cependant, elle présente des limitations en termes de flexibilité pour modéliser certains types de dépendances.

On considère X' , Y' , Z' et T des variables aléatoires qui suivent indépendamment des distributions de Poisson avec paramètres $\lambda_1 - \phi$, $\lambda_2 - \phi$, $\lambda_3 - \phi$ et ϕ respectivement.

On pose :

$$X = X' + T$$

$$Y = Y' + T$$

$$Z = Z' + T$$

Le vecteur aléatoire (X, Y, Z) suit alors une distribution Poisson trivariée avec λ_1 , λ_2 , λ_3 et ϕ comme paramètres.

La fonction génératrice des probabilités pour la distribution de Poisson trivariée est définie par :

$$\begin{aligned} G_{X,Y,Z}(u, v, w) &= E(u^X v^Y w^Z) = E(u^{X'})E(v^{Y'})E(w^{Z'})E\{(uvw)^T\} \\ &= \exp\{(\lambda_1 - \phi)(u - 1) + (\lambda_2 - \phi)(v - 1) + (\lambda_3 - \phi)(w - 1) + \phi(uvw - 1)\} \end{aligned} \quad (\text{I.7})$$

avec $\lambda_1, \lambda_2, \lambda_3 > 0$ et $0 < \phi \leq \min(\lambda_1, \lambda_2, \lambda_3)$.

Les distributions marginales sont des Poisson avec $E(X) = Var(X) = \lambda_1$, $E(Y) = Var(Y) = \lambda_2$ et $E(Z) = Var(Z) = \lambda_3$.

La fonction de covariance est donnée par :

$$Cov(X, Y) = Cov(Y, Z) = Cov(X, Z) = \phi$$

I.1.1.3 Distribution de Poisson Multivariée

Des constructions de distribution Poisson multivariées ont été proposées dans la littérature, et des auteurs qui ont travaillé plus en détail sur les distributions bivariées avancent que la généralisation vers un cas multivarié est possible ([H. Liu \(2012\)](#) par exemple).

Néanmoins, généraliser la distribution et obtenir la fonction de masse devient de plus en plus difficile à mesurer lorsque la dimension augmente. En général, la fonction de masse jointe d'un vecteur de variables aléatoires discrètes a une forme fonctionnelle complexe, ce qui la rend peu pratique pour développer une inférence basée sur la vraisemblance ([FOKIANOS, 2021](#)). Ces distributions sont rarement, voire jamais, utilisées pour les modèles de séries temporelles de comptage.

Suivant une construction similaire encore une fois à celle de la réduction trivariée, on peut introduire une distribution de poisson multivariées ([JOHNSON et al., 1997](#); [KOCHERLAKOTA et KOCHERLAKOTA, 2017](#)).

[FOKIANOS et al. \(2022\)](#) par exemple pour l'utilisation dans les séries de comptage multivariées, inspiré par le vecteur d'observations $\mathbf{Y} = (Y_1, \dots, Y_d)^T$ qui est un

vecteur de comptes d -dimensionnel, dont les composantes ne sont pas nécessairement indépendantes.

Une extension naturelle de la distribution Poisson univariée vers le multivarié est ensuite faite (JOHNSON et al., 1997; KOCHERLAKOTA et KOCHERLAKOTA, 2017) :

On dénote $Y_i = W_i + W$, $i \in \{1, \dots, d\}$ avec $W_i \sim \text{Poisson}(\lambda_i)$ et $W \sim \text{Poisson}(\lambda_0)$. Les variables aléatoires $\{W, W_1, \dots, W_d\}$ sont mutuellement indépendantes.

La fonction de masse jointe du vecteur \mathbf{Y} est donnée par :

$$P[\mathbf{Y} = \mathbf{y}] = \exp \left(- \sum_{i=0}^d \lambda_i \right) \left(\prod_{i=1}^d \frac{\lambda_i^{y_i}}{y_i!} \right) \sum_{k=0}^{\min_i y_i} \left(\prod_{i=1}^d \binom{y_i}{k} \right) k! \left(\frac{\lambda_0}{\prod_{i=1}^d \lambda_i} \right)^k, \quad \mathbf{y} = (y_1, \dots, y_d)^T. \quad (\text{I.8})$$

$Y_i \sim \mathcal{P}(\lambda_i + \lambda_0)$ avec $i \in \{1, \dots, d\}$ et $\text{Cov}(Y_i, Y_j) = \lambda_0 > 0$.

Le paramètre λ_0 détermine toutes les corrélations en paires possibles.

La forme fonctionnelle de (I.8) par contre rend la tâche d'estimation difficile, les modèles se basant dessus seront alors limités aux basses dimensions.

Dans la suite par exemple, pour des modèles de séries de comptage multivariées avec 3 ou plus de variables, nous prendrons une approche différente, ou des auteurs affirment que la distribution explicite n'est pas nécessaire mais juste l'expression de la moyenne (S. LEE et al., 2023). celle-ci sera développée dans la section suivante.

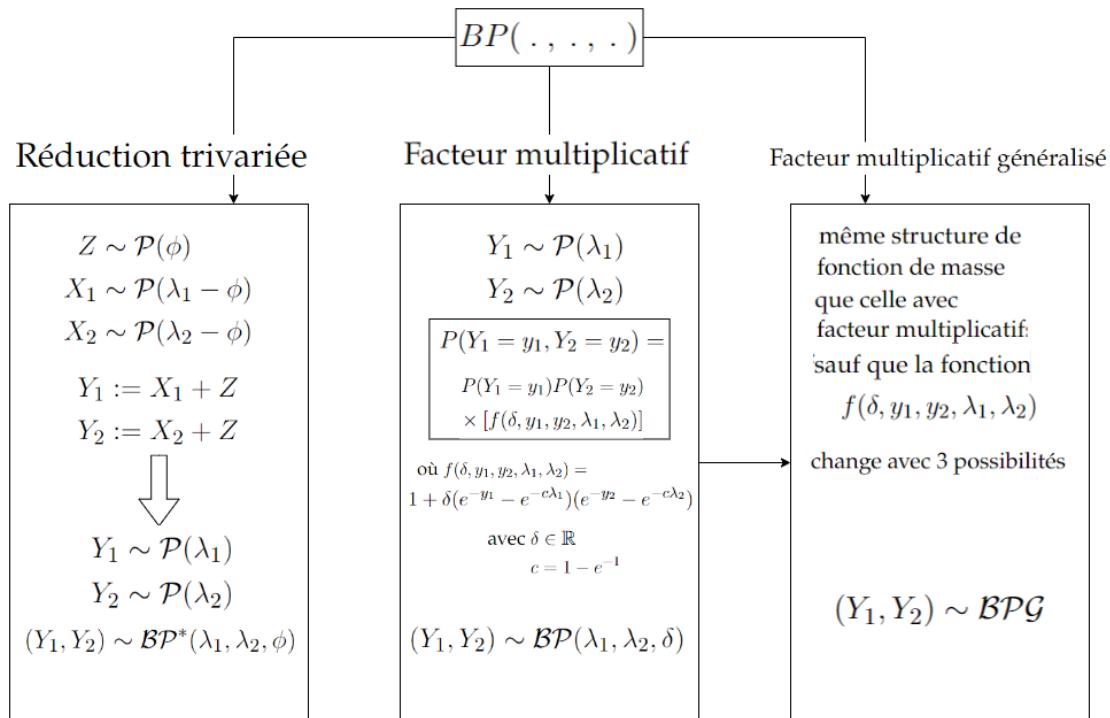


FIGURE I.2 – Schéma résumant les différentes constructions de la loi Poisson bivariée

I.1.2 Les modèles INGARCH multivariés

Nous avons noté précédemment que le **INGARCH** n'est en fait qu'une loi conditionnelle à support naturel, et nous avons vu sa forme générale (I.1) :

$$\begin{aligned} \mathbf{Y}_t | \mathcal{F}_{t-1} &\sim \mathcal{L}(\sigma_t) \\ \sigma_t &= \omega + \sum_{i=1}^p B_i \mathbf{Y}_{t-i} + \sum_{j=1}^q A_j \sigma_{t-j} \end{aligned}$$

Il ne reste maintenant plus qu'à introduire les différentes distributions dans cette formulation pour définir les différents modèles (qui seront donnés sans grand détails dans cette section de revue).

Commençant par les modèles qui traitent le cas bivarié, dans ce qui suit \mathbf{Y}_t dénotera un vecteur bi-dimensionnel d'observations.

I.1.2.1 Modèle Bivarié BINGARCH

H. Liu (2012) présente un modèle bivarié **INGARCH** nommé "**BINGARCH**" et avance que l'extension au cas multivarié est possible en utilisant une formulation similaire.

Considérant la construction par réduction trivariée (I.1.1.1), le modèle **BINGARCH(1,1)** se formule ainsi :

$$\begin{aligned} Y_t | \mathcal{F}_{t-1} &\sim BP^*(\lambda_{t,1}, \lambda_{t,2}, \phi) \\ \lambda_t &= (\lambda_{t,1}, \lambda_{t,2})^T = \delta + A\lambda_{t-1} + BY_{t-1} \end{aligned} \tag{I.9}$$

La généralisation de ce modèle a été faite par **H. Liu (2012)** vers le **BINGARCH(m,n)** avec $m, n \in \mathbb{N}$. Celui-ci est défini par :

$$\begin{aligned} Y_t | \mathcal{F}_{t-1} &\sim BP^*(\lambda_{t,1}, \lambda_{t,2}, \phi) \\ \lambda_t &= (\lambda_{t,1}, \lambda_{t,2})^T = \delta + \sum_{i=1}^m A_i \lambda_{t-i} + \sum_{j=1}^n B_j Y_{t-j} \end{aligned} \tag{I.10}$$

I.1.2.2 Modèle Poisson Bivarié BP-INGARCH

Il est avancé par **Cui et Zhu (2018)** que l'inconvénient principal dans l'utilisation de la distribution précédente (par réduction trivariée) est qu'elle ne permet pas des valeurs de corrélation négatives. Une nouvelle distribution est alors choisie, celle construite par facteur multiplicatif (I.1.1.2).

Le modèle se formule ainsi :

$$\begin{aligned} Y_t | \mathcal{F}_{t-1} &\sim BP(\alpha_{t,1}, \alpha_{t,2}, \phi) \\ \alpha_t &= (\alpha_{t,1}, \alpha_{t,2})^T = \omega + A\alpha_{t-1} + BY_{t-1} \end{aligned} \tag{I.11}$$

(Ce modèle et ses propriétés seront discutés en plus de détail par la suite (Chapitre II))

I.1.2.3 Modèle Poisson Bivarié Généralisé GBP-INGARCH

Cui et al. (2020) affirme que les modèles précédents ont une structure de dépendance limitée et n'offrant pas assez de flexibilité pour la modélisation, d'où la recherche d'alternatives.

Traité ici est une classe de modèles qui sont inspirés de LAKSHMINARAYANA et al. (1999), et se baseront sur le facteur multiplicatif flexible, donc distribution Poisson bivariée généralisée 'GBP'(I.1.1.3).

Le modèle se formule ainsi :

$$\begin{aligned} Y_t | F_{t-1} &\sim GBP(\lambda_t^T) \\ \lambda_t = (\lambda_{t,1}, \lambda_{t,2}) &= \omega + A\lambda_{t-1} + BY_{t-1} \end{aligned} \quad (I.12)$$

où GBP (généralisé) désigne l'une des trois distributions correspondante à cette classe (I.1.1.3).

I.1.2.4 Modèle Multivarié V-INGARCH

FOKIANOS (2021) utilise, comme discuté dans (I.1.1.3), la classe de distributions Poisson multivariées (JOHNSON et al., 1997; KOCHERLAKOTA et KOCHERLAKOTA, 2017) qui généralise les modèles poisson univariés.

On rappelle sa construction similaire à la réduction trivariée :

nous aurons plusieurs variables aléatoires discrètes suivant des distributions Poisson distinctes, avec une variable particulière en commun qui sera utilisée pour faire le lien entre elles.

Soit $Y = (Y_1, \dots, Y_d)$ et $Y_i = W_i + W$, $i = 1, 2, \dots, d$

où $W_i \sim \text{Poisson}(\lambda_i)$ et $W \sim \text{Poisson}(\lambda_0)$ avec tous les W indépendants.

La fonction de masse jointe est donnée par (I.8) :

$$P[Y = y] = \exp\left(-\sum_{i=0}^d \lambda_i\right) \left(\prod_{i=1}^d \frac{\lambda_i^{y_i}}{y_i!}\right) \sum_{k=0}^{\min_i(y_i)} \left(\prod_{i=1}^d \binom{y_i}{k}\right) k! \left(\frac{\lambda_0}{\prod_{i=1}^d \lambda_i}\right)^k$$

En raison de la forme compliquée de cette fonction de masse, ce modèle est approprié pour des analyses de basse dimension où les marginales sont positivement corrélées (FOKIANOS, 2021).

Pour initier la construction du modèle, le modèle standard VAR(1) est considéré, mais dans le contexte plus précis d'une autorégression de Poisson multivariée.

On dénote \mathcal{F}_t la σ -algèbre engendrée par les observations précédentes à l'instant t du processus : $Y_{ss \leq t}$, et on suppose $\lambda_t = (\lambda_{i,t})$, $i = 1, 2, \dots, d$, $t \in \mathbb{Z}$ le processus d'intensité d -dimensionnel correspondant avec $E[Y_t | \mathcal{F}_t] = \lambda_t$.

On définit alors :

$$\begin{aligned} \mathbf{Y}_{i,t} | \mathcal{F}_{t-1} &\sim \text{Poisson}(\lambda_{i,t}) \text{(marginale)} \\ \lambda_t &= \omega + B_1 Y_{t-1} \end{aligned} \quad (\text{I.13})$$

avec ω un vecteur d -dimensionnel, et B_1 une matrice $d \times d$ inconnue.

Les éléments de ω et B_1 sont considérés positifs de telle sorte que $\lambda_{i,t} > 0$ pour tout i et t .

Générer des données en utilisant (I.13) se fait en imposant (I.1.1.3), mais le fitting/ajustement fait face à des difficultés sous condition d'une distribution de poisson multivariée. Ces difficultés peuvent être adressées en faisant recours aux copules, bien que d'autres fonctions de masse soient applicables.

Le modèle (I.13) sera maintenant noté **V-INARCH(1)**.

Il est généralisé par un **V-INARCH(p)** :

$$\lambda_t = \omega + \sum_{i=1}^p B_i Y_{t-i} \quad (\text{I.14})$$

Ou encore un **V-INGARCH(p,q)** :

$$\begin{aligned} \mathbf{Y}_{i,t} | F_{t-1} &\sim \text{Poisson}(\lambda_{i,t}) \\ \lambda_t &= \omega + \sum_{i=1}^p B_i Y_{t-i} + \sum_{j=1}^q A_j \lambda_{t-j} \end{aligned} \quad (\text{I.15})$$

Avec $(A_j)_{j=1}^q, (B_i)_{i=1}^p$ deux matrices inconnues de dimensions $d \times d$, $\lambda_t = (\lambda_{1,t}, \dots, \lambda_{d,t})^T$ et tous les éléments de ω , $(A_j)_{j=1}^q, (B_i)_{i=1}^p$ sont positifs de sorte que $\lambda_{i,t} > 0$ pour tout $i = \{1, \dots, d\}$ et t .

Remarque I.1 Il est aussi mentionné par [FOKIANOS \(2021\)](#) que vu la distribution de chaque élément du processus $\{\mathbf{Y}_t\}_t$ étant une Poisson marginale, l'abréviation INGARCH ne fait référence qu'à la similitude structurelle au **GARCH** ordinaire de [BOLLERSLEV \(1986\)](#), mais étant donné que la variance d'une variable Poisson est égale à sa moyenne, la structure ressemble à celle d'un **GARCH** multivarié.

Et bien que dans ce cas le terme "V-INGARCH" (Vector INGARCH) ne reflète pas précisément le processus générateur des données, il sera tout de même utilisé comme généralisation de la terminologie introduite pour le cas univarié par [FERLAND et al., 2006](#).

[FOKIANOS, 2021](#) considère aussi le modèle Log-Linéaire, analogue au cas univarié introduit par [FOKIANOS et TJØSTHEIM, 2011](#) :

$$\begin{aligned} Y_{i,t} | F_{t-1} &\sim \text{Poisson}(\lambda_{i,t}) \text{(marginale)} \\ v_t &= \omega + \sum_{i=1}^p B_i \log(Y_{t-i} + 1_d) + \sum_{j=1}^q A_j v_{t-j} \end{aligned} \quad (\text{I.16})$$

I.1.2.5 Modèle Multivarié M-INGARCH

Il est argumenté par [S. LEE et al. \(2023\)](#) que dans les études pour les modèles multivariés tel que le **V-INGARCH** ([I.1.2.4](#)), les auteurs ont fourni plus d'effort pour la spécification de la distribution jointe de la série temporelle et la distribution marginale des composantes, ce qui intensifie la complexité de la modélisation par rapport au cas INGARCH univarié.

Par contre, en introduisant une nouvelle structure plus simple se basant sur le modèle univarié, s'inspirant du point de vue qu'une distribution jointe explicite dans le modèle n'est pas l'objectif le plus pressant, et l'expression de l'équation de la moyenne conditionnelle forme l'essentiel de la modélisation, la modélisation se trouve simplifiée à son tour.

Cette approche est encore plus pertinente lorsque l'on est concerné plus tard par l'estimation des paramètres et prédictions dans nos inférences.

En effet, en supposant que les marginales suivent une distribution de la famille exponentielle, et où la moyenne conditionnelle est récursivement produite par une fonction récurrente de toutes les composantes, notre modèle se trouve plus malléable. De plus, vu que les distributions marginales peuvent différer les unes des autres, et que l'équation de la moyenne conditionnelle n'est pas nécessairement linéaire, la flexibilité est améliorée.

[S. LEE et al. \(2023\)](#) a défini le modèle sans expliciter la distribution jointe, $Y_t = (Y_{t1}, \dots, Y_{tm})^T$, $t > 1$ étant la série temporelle de comptage avec $Y_t \in \mathbb{N}^m$.

Le modèle **INGARCH** univarié est pris comme base pour construire un modèle multivarié tel que chaque composante de Y_t suit conditionnellement une loi appartenant à la famille de lois exponentielle avec un seul paramètre :

$$p_i(y|\eta) = \exp\{\eta y - A_i(\eta)\} h_i(y), y \in \mathbb{N} \quad (\text{I.17})$$

qui représente la fonction de masse avec η le paramètre naturel, $A_i(\eta)$, $h_i(y)$ sont des fonctions connues, avec A_i et sa dérivée $B_i = A'_i$ strictement croissantes.

On considère le modèle suivant :

$$\begin{aligned} Y_{ti} | \mathcal{F}_{t-1} &\sim p_i(y|\eta_{ti}), i = 1, \dots, m \\ X_t := E(Y_t | F_{t-1}) &= f_\theta(X_{t-1}, Y_{t-1}) \end{aligned} \quad (\text{I.18})$$

Avec \mathcal{F}_{t-1} la σ -algèbre engendré par les observations précédentes l'instant t , $B_i(\eta_{ti}) = X_{ti}$ et f_θ une fonction non-négative définit en $[0, \infty)^m \times \mathbb{N}^m$ et qui dépend du paramètre $\theta \in \Theta \subset \mathbb{R}^d$, avec $d = 1, 2, \dots$

Bien que chaque composante de Y_t soit modélisée par un INGARCH univarié, la structure de dépendance est imposée par le processus de moyenne conditionnelle X_t . A noter que $B'_i(\eta_{ti})$ est la variance conditionnelle de Y_{ti} .

Cette modélisation nous permet alors principalement de faciliter la tâche vu l'utilisation de distributions marginales des composantes uniquement ainsi qu'une flexibilité résultante de la possibilité d'avoir des distributions distinctes.

Un exemple de la fonction f_θ est une fonction linéaire de forme :

$$f_\theta(x, y) = W + Ax + By$$

Où $\theta = \text{vec}(W, A, B)$ dénote une *vectorisation* de W un vecteur m -dimensionnel de composantes positives, et des matrices $m \times m$ $A = (a_{ij})$ et $B = (b_{ij})$ de composantes non-négatives qui satisfont soit :

$$\sup_{\theta \in \Theta} \left\{ \sum_{j=1}^m (a_{ij} + b_{ij}) \right\} < 1, \quad i = 1, \dots, m \quad (\text{I.19})$$

pour Θ un compact de \mathbb{R}^{m+2m^2} , ou bien :

$$\sup_{\theta \in \Theta} \left\{ \max_{1 \leq j \leq m} \left(\sum_{i=1}^m a_{ij} \right) \max_{1 \leq j \leq m} \left(\sum_{i=1}^m b_{ij} \right) \right\} < 1 \quad (\text{I.20})$$

qui sont toutes deux des conditions suffisantes pour l'existence d'une solution strictement stationnaire et ergodique de $\{Y_t\}$ satisfaisant (I.18).

Et le modèle (I.18) devient enfin un **MINGARCH** linéaire stationnaire :

$$\begin{aligned} Y_{ti} | \mathcal{F}_{t-1} &\sim p_i(y | \eta_{ti}) \\ X_t &= W + AX_{t-1} + BY_{t-1} \end{aligned} \quad (\text{I.21})$$

Quant à la distribution p_i , on peut adopter une distribution de poisson avec $Y_{ti} | \mathcal{F}_{t-1} \sim \text{Poisson}(X_{ti})$

Le modèle (I.21) est applicable de façon universelle en pratique lorsque soit A ou B sont diagonales, vu que cette condition réduit le nombres de paramètres. En pratique, un grand nombre de paramètres devient nuisible à la précision de l'estimation, cette condition/mise en place est recommandée par [S. LEE et al. \(2023\)](#).

I.2 Résultats obtenus concernant la structure probabiliste

Ici seront présentés quelques résultats disponibles dans la littérature concernant les différents modèles.

I.2.1 Stationnarité et érgodicité

- [H. LIU, 2012](#) a étudié les conditions de stabilité pour le modèle **BINGARCH(1,1)** (I.9). basé sur les propriétés des chaînes de Markov bivariées $\{\lambda_t, t \geq 1\}$.

Proposition I.1 *H. Liu (2012)*

Pour le modèle (I.9) avec δ, A, B des entiers non-négatifs :

- Si $\rho(A + B) < 1$, alors il existe au moins une distribution stationnaire de $\{\lambda_t\}$, et si de plus $\|A\|_p < 1$, pour $1 \leq p \leq \infty$ alors la distribution stationnaire est unique.
- Si $\|A\|_p + 2^{(1-1/p)}\|B\|_p < 1$, pour $1 \leq p \leq \infty$, alors $\{\lambda_t\}$ est une chaîne de Markov à contraction géométrique des moments avec une unique distribution stationnaire et ergodique.

Après la généralisation du modèle au **BINGARCH(m,n)**, On définit :

$$Z_t = \begin{cases} (\lambda_t, \dots, \lambda_{t-m+1}, n = 1) \\ (\lambda_t, \dots, \lambda_{t-m+1}, Y_{t-1}, \dots, Y_{t-n+1}), n \geq 1 \end{cases} \quad (\text{I.22})$$

$\{Z_t\}$ est une chaîne de Markov d'après la dynamique du modèle.

Théorème I.1 *H. Liu (2012)*

considérons le modèle (I.10), et suppose que

$$\sum_{i=1}^m \|A_i\|_p + 2^{(1-1/p)} \sum_{j=1}^n \|B_j\|_p < 1$$

pour $p \in [1, \dots, \infty]$.

Alors $\{Z_t\}$ est un moment géométrique à contraction et τ -faiblement dépendant. D'où $\{(Y_t, Z_t)\}$ stationnaire et ergodique.

- Les modèle **BP-INGARCH** (I.11) (**Cui et Zhu, 2018**) et **GBP-INGARCH** (I.12) (**Cui et al., 2020**) ont les mêmes conditions que celles vues dans (**H. Liu, 2012**) pour la stabilité car ils sont eux aussi des modèles bivariés et possèdent des propriétés markoviennes.
- Les conditions de stabilité pour le modèle **V-INGARCH(p,q)** (I.15) (**FOKIANOS et al., 2022**) ont été étudiées sous le cadre de distributions de Poisson multivariées par **H. Liu (2012)**, ainsi que par **FOKIANOS et al. (2020)** pour le cas $p = q = 1$ sous la construction par copules et récemment améliorés par **DEBALY et TRUQUET (2019)**.

La condition

$$\rho(\sum_i A_i + \sum_j B_j) < 1$$

garantie la stabilité du processus.

Le modèle Log-linéaire (I.16) a des conditions plus compliquées, **FOKIANOS et al. (2020)** considère le cas $p = q = 1$ et prouve que soit :

$$|||A_1|||_2 + |||B_1|||_2 < 1$$

$$\text{ou } |||A_1|||_1 + |||B_1|||_1 < 1$$

avec $\|A\|_d = \max_{\|x\|_d=1} \|A_x\|_d$, garantissent l'**érgodicité** du processus. D'autres conditions de stabilité sont discutées par **DEBALY et TRUQUET (2019)**.

L'étude des conditions de stabilité pour le modèle **M-INGARCH(I,21)** a été faite par **S. LEE et al. (2023)**, elles se résument par le théorème suivant :

Théorème I.2 (*S. LEE et al., 2023*)

Si les conditions de contraction suivantes sont valables pour tout $x, x' \in [0, \infty)^m$ et $y, y' \in \mathbb{N}_0^m$:

$$\max_{1 \leq i \leq m} \sup_{\theta \in \Theta} |f_{i,\theta}(x, y) - f_{i,\theta}(x', y')| \leq \lambda_1 \|x - x'\|_1 + \lambda_2 \|y - y'\| \quad (\text{I.23})$$

avec les constantes $\lambda_1, \lambda_2 \geq 0$ et $\lambda_1 + \lambda_2 \leq 1$, où $f_{i,\theta}$ représente la i -ème composante de f_θ et $\|x\|_1 = \sum_{i=1}^m |x_i|$ pour $x = (x_1, \dots, x_m)^T$.

Alors il existe une solution strictement stationnaire et érgodique de $\{Y_t\}$ satisfaisant le modèle (I.18), l'expression de $X_t(\theta) = f_\infty^\theta(y_{t-1}, y_{t-2}, \dots)$ est autorisée pour une certaine fonction mesurable non-négative f_∞^θ définie sur $(\mathbb{N}_0^m)^\infty$.

I.2.2 Autocovariance et autocorrélation

En supposant la stationnarité de 1er ordre du modèle **V-INGARCH** défini par (I.15), il est prouvé que (I.15) est un processus $VARMA(max(p, q), q)$. **FOKIANOS (2021)** a déterminé la fonction de covariance. On a $\xi_t = Y_t - \lambda_t$, $E[Y_t] = (I_d - \sum_{i=1}^p A_i - \sum_{j=1}^q B_j)^{-1}\omega$, et si $\rho(\sum_{i=1}^p A_i - \sum_{j=1}^q B_j) < 1$.

$$(Y_t - E[Y_t]) = \sum_{i=1}^{\max(p,q)} (A_i + B_i)(Y_{t-i} - E[Y_t]) + \xi_t - \sum_{j=1}^q A_j \xi_{t-j} \quad (\text{I.24})$$

avec $A_i = 0_d$ si $q < p$ pour $i = q+1, \dots, p$ ou $B_i = 0_d$ si $p < q$ pour $i = p+1, \dots, q$.

Dans le cas de $p = q = 1$ le $MA(\infty)$ est donné par $Y_t = \mu + \sum_{j=0}^{\infty} \Phi_j \xi_{t-j}$, $\Phi_0 = I_d$ et $\Phi_j = (A_1 + B_1)^{j-1} B_1$, $j \geq 1$.

Pour $h > 0$, et en utilisant les propriétés des processus multivariés :

$$\Gamma_\delta(h) = \sum_{j=0}^{\infty} (A_1 + B_1)^{j+h-1} B_1 E[\Sigma_t] B_1^T (A_1^T + B_1^T)^{j+h-1} \quad (\text{I.25})$$

I.3 Inférence et procédures statistiques proposées

I.3.1 Estimation des paramètres

- **H. LIU (2012)** utilise la méthode de Maximum de Vraisemblance **MV** pour estimer les paramètres du modèle **BINGARCH(1,1)**.

Y_1, Y_2, \dots, Y_n des observations du modèle (I.9), A est supposé une matrice $A = diag(\alpha_1, \alpha_2)$ pour ne pas perdre l'aspect de généralité. $B = (\beta_{ij})_{i,j=1,2}$. Le vecteur des paramètres à estimer est :

$$\theta = (\delta_1, \delta_2, \alpha_1, \alpha_2, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \phi)^T$$

La fonction de Vraisemblance conditionnellement à $\lambda_1 = (\lambda_{1,1}, \lambda_{1,2})^T$ et basée sur les observations Y_1, Y_2, \dots, Y_n est définie par :

$$\begin{aligned} \mathbb{L}(\theta|Y_1, Y_2, \dots, Y_n, \lambda_1) &= f(Y_1|\theta, \lambda_1) \prod_{t=2}^n f(Y_t|Y_1, \dots, Y_{t-1}, \lambda_1, \theta) \\ &= \prod_{t=1}^n e^{-(\lambda_{t,1} + \lambda_{t,2} - \phi)} \frac{(\lambda_{t,1} - \phi)^{Y_{t,1}} (\lambda_{t,2} - \phi)^{Y_{t,2}}}{Y_{t,1}! Y_{t,2}!} \\ &\times \sum_{s=0}^{\min(Y_{t,1}, Y_{t,2})} \binom{Y_{t,1}}{s} \binom{Y_{t,2}}{s} s! \left(\frac{\phi}{(\lambda_{t,1} - \phi)(\lambda_{t,2} - \phi)} \right)^s \end{aligned}$$

et la fonction log-Vraisemblance est donnée par :

$$\begin{aligned} \mathcal{L}(\theta) &= - \sum_{t=1}^n (\lambda_{t,1} + \lambda_{t,2} - \phi) + \sum_{t=1}^n Y_{t,1} \log(\lambda_{t,1} - \phi) + \sum_{t=1}^n Y_{t,2} \log(\lambda_{t,2} - \phi) \\ &+ \sum_{t=1}^n \log \left\{ \sum_{s=0}^{\min(Y_{t,1}, Y_{t,2})} \binom{Y_{t,1}}{s} \binom{Y_{t,2}}{s} s! \left(\frac{\phi}{(\lambda_{t,1} - \phi)(\lambda_{t,2} - \phi)} \right)^s \right\} \end{aligned}$$

La fonction de $\mathcal{L}(\theta)$ est à maximiser sous la contrainte $\phi \leq \min\{\frac{\delta_1}{(1-\alpha_1)}, \frac{\delta_2}{(1-\alpha_2)}\}$, l'estimateur de Maximum de Vraisemblance (EMV) est une solution pour ce problème d'optimisation contrainte.

cette contrainte assure que $\phi \leq \min\{\lambda_{t,1}, \lambda_{t,2}\}$, $\forall t$. Les estimations obtenues devraient satisfaire les conditions de (I.9) pour assurer la stabilité du modèle ajusté.

- La méthode de Maximum de Vraisemblance a été utilisé aussi pour l'estimation du modèle **BP-INGARCH** (I.11) par **Cui et ZHU (2018)** sous l'hypothèse que A, B sont des matrices non négatives. Les fonction de Vraisemblance conditionnellement à $\lambda_1 = (\lambda_{1,1}, \lambda_{1,2})^T$, log-Vraisemblance et score sont données, la dernière étant définie comme suit :

$$S_n(\theta) = \frac{\partial \mathcal{L}(\theta)}{\partial \theta} = \sum_{t=2}^n \frac{\partial \mathcal{L}_t(\theta)}{\partial \theta}$$

On note par $\hat{\theta}$ l'estimateur de Maximum de Vraisemblance de l'équation $S_n(\theta) = 0$.

Cui et ZHU (2018) étudie aussi les propriétés asymptotiques de l'EMV $\hat{\theta}$.

- L'inférence du modèle **BPG-INGARCH** (I.12) est étudiée par **Cui et al. (2020)**. $Y_i, i=1,..,n$ sont des observations du modèle (I.12). A et B sont des matrices diagonales et non négatives, l'extension pour le cas dans les matrices A et B non diagonales et non négatives est directe (voir **Cui et al., 2020** pour plus de détails).

- Le vecteur des vraies valeurs des paramètres : $\theta^0 = (\theta_1^0, \theta_2^0)^T$.
- Le vecteurs des paramètres à estimer : $\theta = (\theta_1, \theta_2)^T$ où $\theta_1 = (\omega_1, \alpha_1, \beta_1, \omega_2, \alpha_2, \beta_2)^T$ et $\theta_2 = \rho, \delta$, où σ dépend du choix du facteur multiplicatif $c(., .)$. Z^{-1} est considéré comme un poids constant.

alors, la fonction log-Vraisemblance est définie par :

$$\mathcal{L}(\theta) = \sum_{t=2}^n \mathcal{L}_t(\theta) = \sum_{t=2}^n [Y_{t,1} \ln \lambda_{t,1}(\theta_1) + Y_{t,2} \ln \lambda_{t,2}(\theta_1) - \lambda_{t,1}(\theta_1) - \lambda_{t,2}(\theta_1)] \quad (\text{I.26})$$

$$+ \sum_{t=2}^n [\ln c(F_1(Y_{t,1}; \lambda_{t,1}(\theta_1)), F_2(Y_{t,2}; \lambda_{t,2}(\theta_1)))] \quad (\text{I.27})$$

avec $F_i(Y_{t,i}; \lambda_{t,i})$ la fonction de masse cumulée conjointe marginale avec paramètres $\lambda_{t,i}$ et $Y_{t,i}$.

Le problème de cette méthode est que la fonction objective est très compliquée et les dérivées du second ordre sont difficiles à dériver analytiquement. Alors, Cui et al. (2020) a fait appel aux approches alternatives suivantes :

- **MBP** : La maximisation par partie est un algorithme proposé par SONG et al. (2005), il consiste à réécrire la fonction de log-Vraisemblance :

$$\mathcal{L}(\theta) = \mathcal{L}_m(\theta_1) + \mathcal{L}_c(\theta_1, \theta_2) = \sum_{t=2}^n \mathcal{L}_{m,t}(\theta_1) + \sum_{t=2}^n \mathcal{L}_{c,t}(\theta_1, \theta_2)$$

dans laquelle la partie de la densité marginale de $\mathcal{L}_m(\theta_1)$ représente le premier terme (I.26), et $\mathcal{L}_c(\theta_1, \theta_2)$ est le terme (I.27) du facteur multiplicatif.

La fonction score est donnée par :

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = \frac{\partial \mathcal{L}_m(\theta_1)}{\partial \theta} + \frac{\partial \mathcal{L}_c(\theta_1, \theta_2)}{\partial \theta}$$

L'algorithme **MBP** se résume comme suit :

- **étape 1** : résoudre $\frac{\partial \mathcal{L}_m(\theta_1)}{\partial \theta_1} = 0$ pour $\hat{\theta}_{1,n}^1$ et $\frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^1, \theta_2)}{\partial \theta_2} = 0$ pour $\hat{\theta}_{2,n}^1$.
- **étape k** : résoudre $\frac{\partial \mathcal{L}_m(\theta_1)}{\partial \theta_1} = -\frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^{k-1}, \hat{\theta}_{2,n}^{k-1})}{\partial \theta_1}$ pour $\hat{\theta}_{1,n}^k$ et $\frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^{k-1}, \theta_2)}{\partial \theta_2} = 0$ pour $\hat{\theta}_{2,n}^k$, $k=2,3,\dots$

- **MMBP** : L'algorithme de maximisation par partie modifié proposé par Y. LIU et LUGER (2009) en raison de la difficulté de convergence et pour améliorer l'efficacité de l'estimation.

La fonction log-Vraisemblance est donnée par :

$$\mathcal{L}^*(\theta) = \mathcal{L}_m^*(\theta_1) + \mathcal{L}_c^*(\theta_1, \theta_2) = [\mathcal{L}_m^*(\theta_1) + \mathcal{L}_c^*(\theta_1, \theta_{2,n}^1)] + [\mathcal{L}_c^*(\theta_1, \theta_2) - \mathcal{L}_c^*(\theta_1, \theta_{2,n}^1)]$$

les étapes de l'algorithme **MMBP** :

- **étape 1** : résoudre $\frac{\partial \mathcal{L}_m(\theta_1)}{\partial \theta_1} = 0$ pour $\hat{\theta}_{1,n}^1$ et $\frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^1, \theta_2)}{\partial \theta_2} = 0$ pour $\hat{\theta}_{2,n}^1$.
- **étape 2** : résoudre $\frac{\partial \mathcal{L}_m(\theta_1)}{\partial \theta_1} + \frac{\partial \mathcal{L}_c(\theta_1, \hat{\theta}_{2,n}^1)}{\partial \theta_1} = 0$ pour $\hat{\theta}_{1,n}^2$ et $\frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^2, \theta_2)}{\partial \theta_2} = 0$ pour $\hat{\theta}_{2,n}^2$.
- **étape k** : résoudre $\frac{\partial \mathcal{L}_m(\theta_1)}{\partial \theta_1} + \frac{\partial \mathcal{L}_c(\theta_1, \hat{\theta}_{2,n}^k)}{\partial \theta_1} = -\left(\frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^{k-1}, \hat{\theta}_{2,n}^{k-1})}{\partial \theta_1} - \frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^{k-1}, \hat{\theta}_{2,n}^k)}{\partial \theta_1}\right)$ pour $\hat{\theta}_{1,n}^k$ et $\frac{\partial \mathcal{L}_c(\hat{\theta}_{1,n}^k, \theta_2)}{\partial \theta_2} = 0$ pour $\hat{\theta}_{2,n}^k$, $k=3,4,\dots$

- **TMB** : *Template Model Builder* Cui et al. (2020) est une méthode plus flexible et rapide basé sur les packages du langage R.

- L'estimation par la méthode de Quasi Maximum de Vraisemblance **QMV** a été utilisé aussi pour le modèle **V-INGARCH** (I.1.2.4), par **FOKIANOS** (2021) suivant la même théorie d'estimation des fonctions utilisée dans **FOKIANOS et al.**, 2020.

Supposant que $Y_t, t = 1, \dots, n$ des observations du modèle (I.1.2.4) pour $p = q = 1$, le vecteur des paramètres à estimer est : $\theta = (d^T, vec^T(A_1), vec^T(B_1))$ avec $dim(\theta) = \kappa = d(1 + 2d)$.

En donnant une valeur initiale λ_0 pour θ , la fonction de quasi-vraisemblance conditionnelle est :

$$L(\theta) = \prod_{t=1}^n \prod_{i=1}^d \left\{ \frac{\exp(-\lambda_{i,t}(\theta))\lambda_{i,t}^{y_{i,t}}(\theta)}{y_{i,t}!} \right\}$$

(voir **FOKIANOS et al.**, 2020 pour plus de détails)

La fonction de quasi log-vraisemblance est alors :

$$\mathcal{L}(\theta) = \sum_{t=1}^n \sum_{i=1}^d (y_{i,t} \log \lambda_{i,t}(\theta) - \lambda_{i,t}(\theta))$$

avec $\hat{\theta} \equiv \arg \max_{\theta} \mathcal{L}(\theta)$, l'estimateur de Quasi Maximum de Vraisemblance (**EQMV**) de θ .

La fonction Score est donnée par :

$$S_n(\theta) = \sum_{t=1}^n \sum_{i=1}^d \left(\frac{y_{i,t}}{\lambda_{i,t}(\theta)} - 1 \right) \frac{\partial \lambda_{i,t}(\theta)}{\partial \theta} = \sum_{t=1}^n \frac{\partial \lambda_t^T(\theta)}{\partial \theta} D_t^{-1}(\theta) (Y_t - \lambda_t(\theta))$$

avec $\frac{\partial \lambda_t}{\partial \theta^T}$ est une matrice $(d \times \kappa)$, et D_t est une matrice diagonale $(d \times d)$ avec le i-ème élément de la diagonale $\lambda_{i,t}(\theta)$, $i = 1, \dots, p$.

$$\begin{aligned} \frac{\partial \lambda_t}{\partial \theta^T} &= I_d + A_1 \frac{\partial \lambda_{t-1}}{\partial d^T} \\ \frac{\partial \lambda_t}{\partial vec^T(A_1)} &= (\lambda_{t-1} \otimes I_d)^T + A_1 \frac{\partial \lambda_{t-1}}{\partial vec^T(A_1)} \\ \frac{\partial \lambda_t}{\partial vec^T(B_1)} &= (Y_{t-1} \otimes I_d)^T + A_1 \frac{\partial \lambda_{t-1}}{\partial vec^T(B_1)} \end{aligned}$$

La matrice Hessienne est donnée par :

$$H_n(\theta) = \sum_{t=1}^n \sum_{i=1}^p \frac{y_{i,t}}{\lambda_{i,t}^2(\theta)} \frac{\partial \lambda_{i,t}(\theta)}{\partial \theta} \frac{\partial \lambda_{i,t}(\theta)}{\partial \theta^T} - \sum_{t=1}^n \sum_{i=1}^p \left(\frac{y_{i,t}}{\lambda_{i,t}(\theta)} - 1 \right) \frac{\partial^2 \lambda_{i,t}(\theta)}{\partial \theta \partial \theta^T}$$

et la matrice d'information conditionnelle est égale à :

$$G_n(\theta) = \sum_{t=1}^n \frac{\partial \lambda_t^T(\theta)}{\partial \theta} D_t^{-1}(\theta) \Sigma_t(\theta) D_t^{-1}(\theta) \frac{\partial \lambda_t(\theta)}{\partial \theta^T}$$

où la matrice $\Sigma_t(\cdot)$ dénote la vraie matrice de variance-covariance du vecteur Y_t .

- **S. Lee et al. (2023)** a utilisé deux méthodes d'estimation pour le modèle **M-INGARCH** (I.21), la méthode Quasi Maximum de Vraisemblance (**QMV**) et **MDPD** (minimum density power divergence).

Pour l'estimation **QMV** on note que $\theta = Vec(W, A, B)$ et la vraie valeur θ_0 est l'intérieur d'un voisinage compact :

$$\Theta_1 = \{\theta : \omega_s \leq W_i \leq \omega^*, 0 \leq \sum_{i=1}^m (a_{ij} + b_{ij}) \leq 1 - \epsilon, i = 1, \dots, m\}$$

ou

$$\Theta_2 = \{\theta : \omega_s \leq W_i \leq \omega^*, 0 \leq \max_{1 \leq j \leq m} \left(\sum_{i=1}^m a_{ij} \right) + \max_{1 \leq j \leq m} \left(\sum_{i=1}^m b_{ij} \right) \leq 1 - \epsilon\}$$

$\epsilon > 0$ et $0 < \omega_s < \omega^*$.

La fonction de pseudo-vraisemblance est donnée par :

$$\mathcal{L}(\theta) = \prod_{t=1}^n \prod_{i=1}^m \exp \{ \tilde{l}(\theta) + \log h_i(Y_{ti}) \}$$

avec $\tilde{l}_{ti}(\theta) = \tilde{\eta}_{ti}(\theta)Y_{ti} - A_i(\tilde{\eta}_{ti}(\theta))$ et $\tilde{\eta}_{ti}(\theta) = B_i^{-1}(\tilde{X}_{ti}(\theta))$

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmax}} \tilde{\mathcal{L}}_n(\theta) := \underset{\theta \in \Theta}{\operatorname{argmax}} \frac{1}{n} \sum_{t=1}^n \sum_{i=1}^m \tilde{l}_{ti}(\theta)$$

Pour étudier les propriétés asymptotique de $\hat{\theta}$, **S. Lee et al. (2023)** propose les hypothèses suivantes :

- **A0** : La condition (I.23) est valable et $f(\theta) = (f_{1,\theta}, \dots, f_{k,\theta})^T$ satisfait pour $c_* > 0$:

$$\inf_{1 \leq i \leq m, \theta \in \Theta} \inf_{x \in [0, \infty)^m, y \in \mathbb{N}_0^m} f_{i,\theta}(x, y) \leq c_*,$$

- **A1** : θ_0 est un point intérieur d'un espace paramétrique compact $\Theta \subset \mathbb{R}^d$.
- **A2** : pour tout i , il est vérifié que :
 $E(Y_{1i}^4) < \infty$, $E(\sup_{\theta \in \Theta} X_{1i}(\theta))^4 < \infty$, et $E\{\sup_{\theta \in \Theta} (Y_{1i}|\eta_{1i}(\theta)| + |A_i(\eta_{1i}(\theta))|)\} < \infty$.
- **A3** : Il existe $\underline{c} > 0$, tel que pour tout i, t :

$$\inf_{\theta \in \Theta} \inf_{0 \leq \delta \leq 1} B'_i((1 - \delta)\eta_{ti}(\theta) + \delta\tilde{\eta}_{ti}(\theta)) \geq \underline{c},$$

avec B'_i est la dérivée de B_i .

- **A4** : S'il existe $t > 1$ tel que $X_t(\theta) = X_t(\theta_0)$ p.s., alors $\theta = \theta_0$.
- **A5** : Pour tout i, t , $\partial^2 X_{ti}(\theta)/\partial \theta \partial \theta^T$ est continue en Θ ,
 $E(\sup_{\theta \in \Theta} \|\frac{\partial X_{1i}(\theta)}{\partial \theta}\|^4) < \infty$ et $E(\sup_{\theta \in \Theta} \|\frac{\partial^2 X_{1i}(\theta)}{\partial \theta \partial \theta^T}\|^2) < \infty$.
- **A6** : pour tout i, t ,

$$\sup_{\theta \in \Theta} \left\| \frac{\partial \tilde{X}_{ti}(\theta)}{\partial \theta} - \frac{\partial X_{ti}(\theta)}{\partial \theta} \right\| \leq V\rho^t, \text{ p.s.}$$

- **A7** : Pour chaque i , $v^T \partial X_{ti}(\theta_0)/\partial \theta = 0$ p.s. si et seulement si $v = 0$.

- A8 : pour tout i,t,

$$\sup_{\theta \in \Theta} \sup_{0 \leq \delta \leq 1} \left| \frac{B_i''((1-\delta)\eta_{ti}(\theta) + \delta\tilde{\eta}_{ti}(\theta))}{B_i' B_i''((1-\delta)\eta_{ti}(\theta) + \delta\tilde{\eta}_{ti}(\theta))^{5/2}} \right| \leq K, \text{ pour } K > 0.$$

avec B_i'' désigne la dérivée deuxième de B_i .

Théorème I.3 *S. LEE et al., 2023* Supposant que les hypothèses (A0)-(A4) sont valables, $n \rightarrow \infty$:

$$\hat{\theta}_n \rightarrow \theta_0 \text{ p.s.}$$

Théorème I.4 *S. LEE et al., 2023* Sous les hypothèses (A0)-(A8), $\hat{\theta}_n$ est asymptotiquement normal, comme $n \rightarrow \infty$:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, J^{-1} I J^{-1})$$

avec $I = E \left(\sum_{i=1}^m \sum_{j=1}^m \frac{\partial \ell_{ti}(\theta_0)}{\partial \theta} \frac{\partial \ell_{tj}(\theta_0)}{\partial \theta^T} \right)$, $J = -E \left(\sum_{i=1}^m \frac{\partial^2 \ell_{ti}(\theta_0)}{\partial \theta \partial \theta^T} \right)$ et $\ell_{ti}(\theta) = \eta_{ti}(\theta) Y_{ti} - A_i(\eta_{ti}(\theta))$.

Pour la 2ème méthode, la **DPD** (density power divergence) noté d_α entre deux fonctions de densité g et h introduite par *Basu et al, 1998*, et en se basant sur d_α , ils ont définit le **MDPDE** pour les Y_i , $i = 1, \dots, n$ générer par le modèle (I.21).

Le **MDPDE** pour θ_0 est définit par :

$$\hat{\theta}_{\alpha,n} = \underset{\theta \in \Theta}{\operatorname{argmin}} \tilde{\mathcal{L}}_{\alpha,n}(\theta) = \underset{\theta \in \Theta}{\operatorname{argmin}} \frac{1}{n} \sum_{t=1}^n \sum_{i=1}^m \tilde{\ell}_{\alpha,t,i}(\theta),$$

I.3.2 Tests d'hypothèses

S. LEE et al. (2023) a appliqué les tests d'un point de changement sur le modèle **M-INGARCH** définit par l'équation (I.21).

Deux types de tests sont utilisés, l'un basé sur le QMLE et l'autre sur le **MDPDE** pour tester les hypothèses suivantes :

$H_0 : \theta$ ne change pas sur Y_1, \dots, Y_n vs $H_1 : \theta$ pas H_0 .

I.3.2.1 Test basé sur le QMLE

En utilisant pour ce type les tests **CUSUM** (somme cumulative) basés sur des résidus standardisés pondérés par les moyennes conditionnelles et les vecteurs de score pour évaluer les hypothèses mentionnées précédemment.

S. LEE et al. (2023) propose une version modifiée du test **CUSUM** basé sur les résidus $\epsilon_t = (\epsilon_{t1}, \dots, \epsilon_{tm})^T$ avec :

$$\epsilon_{ti} = (Y_{ti} - X_{ti}(\theta_0)) / \sqrt{X_{ti}(\theta_0)}.$$

$\{\epsilon_t, \mathcal{F}_t\}$ une suite de différences de martingale stationnaire et ergodique intégrable au carré sous les hypothèses (A₀) et (A₂), d'après le théorème central limite (**TCL**) on peut calculer la statistique du test T_n^{res} .

Comme le ϵ_t n'est pas observable on passe d'abord par le calcul d'une estimation pour ϵ_t notée par $\hat{\epsilon}_t$ avec :

$$\hat{\epsilon}_{ti} = (Y_{ti} - \hat{X}_{ti}) / \sqrt{\hat{X}_{ti}}.$$

Alors, on considère une nouvelle statistique pour ce test notée par \hat{T}_n^{res} , ou bien la statistique du test **CUSUM** basée sur le vecteur score \hat{T}_n^{score} .

I.3.2.2 Test basé sur le MDPDE

Ce test est fondé sur la fonction objective du **MDPDE** "*minimum density power divergence*" pour définir la statistique du test notée \hat{T}^α .

Notant que si $\alpha = 0$ on a l'égalité suivante :

$$\hat{T}^\alpha = \hat{T}_n^{score}.$$

Remarque Pour les tests \hat{T}_n^{res} , \hat{T}_n^{score} et \hat{T}^α , quand l'hypothèse nulle est rejetée, l'emplacement de test d'un point de changement est estimé comme un point de temps qui maximise la statistique du test.

II

MODÈLE INGARCH BIVARIÉ Poisson BP-INGARCH

En raison de leur large éventail d'applications dans de nombreux domaines pratiques comme vu au [Chapitre I](#), les modèles de séries chronologiques de comptages ont fait l'objet d'études intensives ces dernières années. Le principal obstacle dans la modélisation de ce type des séries chronologiques est le choix d'une distribution discrète appropriée.

Notre objectif est d'étudier la classe de modèles bivariés basé sur la distribution de Poisson, spécifiquement le modèle **BP-INGARCH** ([CUI et ZHU, 2018](#)) basé sur le *facteur multiplicatif*.

La section [II.1](#) est consacrée à la définition et à la formulation du modèle. On examine dans la section [II.2](#) les propriétés de stationnarité et d'ergodicité de ce dernier. Enfin, la section [II.3](#) est dédiée à l'inférence et l'estimation, en s'appuyant sur la représentation ARMA du modèle et la vraisemblance.

II.1 Formulation du modèle bivarié BP-INGARCH

Comme discuté précédemment, le modèle introduit par [Cui et Zhu \(2018\)](#) est inspiré de celui avancé par [H. Liu \(2012\)](#), et bien qu'il garde la même approche générale, utilise une différente distribution basé sur un facteur multiplicatif afin de prendre en compte les corrélations négatives, améliorant l'applicabilité en général.

La distribution bivariée dans ce cas possèdera des marginales Poisson avec paramètres qu'on nommera λ_1, λ_2 respectivement, et sa fonction de masse jointe est comme donnée précédemment([I.1.1.1.2](#)) par :

$$P(Y_1 = y_1, Y_2 = y_2) = \frac{\lambda_1^{y_1} \lambda_2^{y_2}}{y_1! y_2!} \exp\{-(\lambda_1 + \lambda_2)\} [1 + \delta(e^{-y_1} - e^{-c\lambda_1})(e^{-y_2} - e^{-c\lambda_2})], \quad (\text{II.1})$$

Avec $c = 1 - e^{-1}$.

La distribution est dénotée $BP(\lambda_1, \lambda_2, \delta)$, tel que λ_1 et λ_2 sont des paramètres positifs.

Notons que : $E(Y_1) = Var(Y_1) = \lambda_1$, et $E(Y_2) = Var(Y_2) = \lambda_2$

La matrice variance-covariance de la distribution est donnée par :

$$\begin{pmatrix} \lambda_1 & \delta c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)} \\ \delta c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)} & \lambda_2 \end{pmatrix}.$$

Avec un coefficient de corrélation $\rho = \delta c^2 \sqrt{\lambda_1 \lambda_2} e^{-c(\lambda_1+\lambda_2)}$.

On a :

$$|\delta| \leq \frac{1}{(1 - e^{-c\lambda_1})(1 - e^{-c\lambda_2})},$$

D'où

$$|\rho| \leq \min\left\{\frac{c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)}}{(1 - e^{-c\lambda_1})(1 - e^{-c\lambda_2})}, 1\right\}.$$

Ce coefficient peut être positif, négatif ou bien nul dépendant de la valeur de paramètre δ .

On note par $\mathbf{Y}_t = (Y_{t,1}, Y_{t,2})^T$ les observations bivariées à l'instant t , qui suivent la distribution bivariée $\mathbf{BP}(\lambda_{t,1}, \lambda_{t,2}, \delta)$.

Le modèle d'intérêt, dénommé BP-INGARCH, est défini dans le cas d'ordre (1,1) par la formulation mathématique suivante :

$$\begin{aligned} \mathbf{Y}_t | \mathcal{F}_{t-1} &\sim \mathbf{BP}(\lambda_{t,1}, \lambda_{t,2}, \delta), \\ \lambda_t &= (\lambda_{t,1}, \lambda_{t,2})^T = \omega + \mathbf{A}\lambda_{t-1} + \mathbf{B}\mathbf{Y}_{t-1}, \end{aligned} \quad (\text{II.2})$$

où $\mathcal{F}_t = \sigma\{\lambda_1, \mathbf{Y}_1, \dots, \mathbf{Y}_t\}$ représente la filtration naturelle du processus,

$\omega = (\omega_1, \omega_2)^T \in \mathbb{R}^+{}^2$ est un vecteur de paramètres constants positifs,

et \mathbf{A} et \mathbf{B} sont des matrices carrées d'ordre 2 à coefficients non négatifs :

$$\mathbf{A} = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix} \text{ et } \mathbf{B} = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{pmatrix}.$$

l'écriture matricielle de la fonction lien (celle de δ) peut être donnée comme suit :

$$\begin{pmatrix} \lambda_{t,1} \\ \lambda_{t,2} \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} + \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix} \begin{pmatrix} \lambda_{t-1,1} \\ \lambda_{t-1,2} \end{pmatrix} + \begin{pmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{pmatrix} \begin{pmatrix} Y_{t-1,1} \\ Y_{t-1,2} \end{pmatrix},$$

D'où on peut expliciter deux équations distinctes pour les composantes de δ_t

$$\begin{cases} \lambda_{t,1} = \omega_1 + \alpha_{1,1}\lambda_{t-1,1} + \alpha_{1,2}\lambda_{t-1,2} + \beta_{1,1}Y_{t-1,1} + \beta_{1,2}Y_{t-1,2}, \\ \lambda_{t,2} = \omega_2 + \alpha_{2,1}\lambda_{t-1,1} + \alpha_{2,2}\lambda_{t-1,2} + \beta_{2,1}Y_{t-1,1} + \beta_{2,2}Y_{t-1,2}, \end{cases}$$

D'où :

$$\lambda_{t,i} = \omega_i + \sum_{j=1}^2 \alpha_{i,j}\lambda_{t-1,j} + \sum_{j=1}^2 \beta_{i,j}Y_{t-1,j}, \quad i = 1, 2. \quad (\text{II.3})$$

Ce type d'écriture nous sera utile plus tard lors de l'estimation. ((II.12), (II.15))

II.2 Théorie de la stationnarité

Cette section explore les conditions assurant la stationnarité du processus λ_t défini par le modèle BP-INGARCH introduit précédemment.

La stationnarité, propriété essentielle pour l'inférence statistique, est étudiée à travers le prisme de la théorie des chaînes de Markov et des contractions.

Le processus $\{\lambda_t, t \geq 1\}$ représente une chaîne de Markov, par récursion on a :

$$\lambda_t = (I + \mathbf{A} + \cdots + \mathbf{A}^{s-1})\omega + \mathbf{A}^s\lambda_{t-s} + \sum_{k=0}^{s-1} \mathbf{A}^k \mathbf{B} \mathbf{Y}_{t-k-1} \quad (\text{II.4})$$

avec \mathbf{I} la matrice identité.

En supposant que $\rho(\mathbf{A}) < 1$ pour $1 \leq p \leq \infty$, on a la formule suivante :

$$\lambda_t = (I - \mathbf{A})^{-1}\omega + \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{B} \mathbf{Y}_{t-k-1}. \quad (\text{II.5})$$

Ainsi, sous la condition $\rho(\mathbf{A}) < 1$, nous avons $\lambda_t \geq (I - \mathbf{A})^{-1}\omega$ pour tout t .

Le théorème suivant, issu des travaux de Cui et Zhu (2018), établit des conditions suffisantes pour l'existence et l'unicité d'une distribution stationnaire de λ_t .

Théorème II.1 (*Cui et ZHU, 2018, Théorème 1, pg6*)

Supposons que $\{Y_t, t \geq 1\}$ suive (II.2), ω , A et B ont des composantes non négatives.

- Si $\rho(A + B) < 1$, alors il existe au moins une distribution stationnaire pour $\{\lambda_t\}$. De plus, si $\|A\|_p < 1$ pour un certain $1 \leq p \leq \infty$, alors la distribution stationnaire est unique.
- Si $\|A\|_p + 2(1 - 1/p)\|B\|_p < 1$ pour un certain $1 \leq p \leq \infty$, alors $\{\lambda_t\}$ est une chaîne de Markov à contraction géométrique des moments avec une distribution stationnaire et ergodique unique, notée π .

Preuve :

La preuve complète est donnée en annexe (Section A.1) et peut être résumée en quelques étapes :

- On note que $\{\lambda_t\}$ possède au moins une distribution stationnaire (*H. LIU, 2012*)
- Observer de (II.2) que $(I - A)^{-1}$ est un état atteignable sous certaines conditions.
- Vérifier qu'une mesure de probabilité invariante unique existe (*théorème 18.8.4 MEYN et TWEEDIE, 2012*).
- Montrer que $\{\lambda_t\}$ est une *e-chain* (utilisation du *lemme 6.4* dans *WANG et al. (2014)*) qui prouve qu'il existe une unique distribution stationnaire.
- Proposition 4.2.1 dans *H. LIU (2012)* pour prouver la deuxième partie du théorème.

Comme conséquence de ce théorème l'espérance non-conditionnelle du processus $\{Y_t\}$ est peut être calculée par :

$$\begin{aligned}
 \mathbb{E}(\mathbf{Y}_t) &= E(E(Y_t | \mathcal{F}_{t-1})) \\
 &= E(\lambda_t) \\
 &= \omega + A\lambda_{t-1} + B Y_{t-1} \\
 &= \omega + (A + B)E(Y_{t-1})
 \end{aligned} \tag{II.6}$$

Et donc $\mathbb{E}(\mathbf{Y}_t) = \mathbb{E}(\lambda_t) = (\mathbf{I} - \mathbf{A} - \mathbf{B})^{-1}\omega$.

avec \mathbf{I} la matrice identité.

II.3 Estimation des paramètres

II.3.1 Estimation par la méthode des Moindres Carrés Ordinaire

II.3.1.1 Principe de la procédure d'estimation

En s'inspirant des travaux précédents de *FERLAND et al. (2006)* où un modèle univarié **INGARCH** est représenté par un **ARMA**, et de *ZHU (2011)* où une estimation par Moindres carrés Conditionnels **MCC** est faite sur la représentation **ARMA** d'un **NBINGARCH** (negative binomial), on considère la méthodologie suivante :

- On commence par notre modèle **BP-INGARCH** en question.

- On propose une représentation **ARMA** bivariée de ce modèle fondé sur un théorème et corollaire proposés pour le **INGARCH** univarié (**FERLAND et al.**, 2006 corollaire 2 p.935).
- On suit la méthodologie utilisée lors de l'estimation d'un **INGARCH** par (**MCC**).
- Une fois l'estimation par **MCC** finie, cette estimation elle-même pourra être utilisée dans la suite comme valeur initiale pour l'estimation par Maximum de Vraisemblance (**MV**).

On commence par poser un Lemme et Corollaire qui nous seront essentiels par la suite :

Lemme II.1 (**FERLAND et al.**, 2006 Lemme 2, pg13)

Supposons $\{X_t\}$ et $\{Y_t\}$ deux processus stationnaires centrés avec $\gamma(\cdot)$ comme fonction d'autocovariance et $\{Y_t\}$ un processus ARMA(p,q). Alors $\{X_t\}$ aussi est un processus ARMA(p,q).

Corollaire II.1 (**FERLAND et al.**, 2006 Corollaire 2, pg13)

*Supposons $\{X_t\}_{t \in \mathbb{Z}}$ un **INGARCH(1,1)** (univarié) satisfaisant l'équation :*

$$\begin{aligned} X_t | \mathcal{F}_{t-1} &\sim \mathcal{P}(\lambda_t) \\ \lambda_t &= \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1}. \end{aligned} \tag{II.7}$$

*Alors $\{X_t\}_{t \in \mathbb{Z}}$ est un **ARMA(1,1)** pouvant s'écrire comme :*

$$(X_t - \mu) - (\gamma_1 + \delta_1)(X_{t-1} - \mu) = e_t - \delta_1 e_{t-1}, \tag{II.8}$$

avec $\{e_t\}_{t \in \mathbb{Z}}$ est un bruit blanc de variance $\sigma^2 = \mu = \gamma_0(1 - \gamma_1 - \delta_1)^{-1}$.

Ou bien :

$$Y_t - \phi Y_{t-1} = \varepsilon_t + \theta \varepsilon_{t-1}, \tag{II.9}$$

avec $\{\varepsilon_t\}_t$ est une suite i.i.d de $\mathcal{N}(0, \sigma^2)$. Notant que :

$$\phi = (\gamma_1 + \delta_1), \theta = -\gamma_1 \text{ et } \sigma^2 = \mu.$$

La procédure qui suit est une adaptation de celle utilisée par **ZHU (2011)** dans le cas d'un **NBINGARCH(1,1)** (donnée en plus de détails en annexe A.2) pour un **INGARCH(1,1)** (univarié) : Le **INGARCH(1,1)** est défini par **FERLAND et al. (2006)** :

$$\begin{aligned} X_t | \mathcal{F}_{t-1} &\sim \mathcal{P}(\lambda_t) \\ \lambda_t &= \alpha_0 + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}. \end{aligned} \tag{II.10}$$

(X_t la série d'observations, $|\mathcal{F}_{t-1}$ la tribu générée par l'information précédente, et $\alpha_0, \alpha_1, \beta_1$ des paramètres positifs < 1)

En utilisant le Lemme II.1 et Corollaire II.1 une écriture **ARMA(1,1)** équivalente est tirée :

$$(X_t - \mu) - a(X_{t-1} - \mu) = e_t + b e_{t-1}, \tag{II.11}$$

avec $a = \alpha_1 + \beta_1$, $b = -\beta_1$ et e_t un bruit blanc de variance μ où $\mu = \alpha_0(1 - \alpha_1 - \beta_1)^{-1}$

D'où on tire l'écriture suivante :

$$\begin{aligned} X_t - \mu - aX_{t-1} + a\mu &= e_t + be_{t-1} \\ X_t &= \mu + aX_{t-1} - a\mu + e_t + be_{t-1} \\ X_t &= \mu(1 - a) + aX_{t-1} + e_t + be_{t-1} \\ X_t &= \frac{\alpha_0}{(1 - a)}(1 - a) + aX_{t-1} + e_t + be_{t-1} \\ X_t &= \alpha_0 + aX_{t-1} + e_t + be_{t-1} \end{aligned}$$

qui nous arrange lors de l'estimation car elle explicite le paramètre de la **constante** (α_0).

Et on a donc :

$$e_t = X_t - (\alpha_0 + aX_{t-1} + be_{t-1})$$

Reste-t-il que l'erreur n'est pas observable, il nous faut donc l'estimer à son tour.

On pose : $Y_t = X_t - \bar{X}_t$ puis ajuste par rapport à un autorégressif d'ordre supérieur **AR(p*)** pour avoir une estimation des erreurs.

$$\left(\bar{X} = \sum_{i=1}^n \frac{X_i}{n}, n \text{ le nombre d'observations disponibles} \right)$$

Une fois les erreurs estimées, On procède à la minimisation de la fonction objective :

$$\sum_{t=1}^n (X_t - \alpha_0 - aX_{t-1} - b\hat{e}_{t-1})^2$$

qui nous permet d'obtenir les estimations $\hat{\alpha}_0$, \hat{a} , \hat{b} et celles des paramètres d'intérêt par :

$$\begin{aligned} \hat{\alpha}_1 &= \hat{a} + \hat{b} \\ \hat{\beta}_1 &= -\hat{b} \end{aligned}$$

II.3.1.2 Procédure d'estimation dans le cas bivarié

Nous procédons par identification afin de produire une représentation équivalente dans le cas **bivarié**.

Pour un processus $\{Y_t\}$ satisfaisant l'équation (II.2), nous voulons arriver à une écriture **ARMA** bivariée.

En centrant le processus en question et en identifiant de façon analogue au cas univarié nous avons :

$$\begin{pmatrix} Y_{t,1} \\ Y_{t,2} \end{pmatrix} - \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} - (B + A) \begin{pmatrix} Y_{t-1,1} \\ Y_{t-1,2} \end{pmatrix} = \begin{pmatrix} e_{t,1} \\ e_{t,2} \end{pmatrix} - A \begin{pmatrix} e_{t-1,1} \\ e_{t-1,2} \end{pmatrix} \quad (\text{II.12})$$

D'où

$$Z_t - \Phi Z_{t-1} = \varepsilon_t + \Theta \varepsilon_{t-1}$$

Avec $Z_t = (X_t - \mu)$, $\Phi = B + A$, $\Theta = -A$,

et $\{\epsilon_t\}$, une suite i.i.d bivariée, où la variance de chaque composante $\epsilon_{t,1}$ et $\epsilon_{t,2}$ est λ_1 , λ_2 respectivement.

$$\text{où } \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = (\mathbf{I} - \mathbf{A} - \mathbf{B})^{-1} \omega.$$

Nous ne passons pas directement à l'estimation de Φ et Θ mais explicitons d'abord la constante.

On a de (II.12) :

$$\begin{pmatrix} Y_{t,1} \\ Y_{t,2} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \Phi \begin{pmatrix} Y_{t-1,1} \\ Y_{t-1,2} \end{pmatrix} - \Phi \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} e_{t,1} \\ e_{t,2} \end{pmatrix} + \Theta \begin{pmatrix} e_{t-1,1} \\ e_{t-1,2} \end{pmatrix}$$

Pour simplifier l'écriture, nous utilisons Y_t , μ , e_t pour dénoter les vecteurs.

$$\begin{aligned} Y_t &= \mu + \Phi Y_{t-1} - \Phi \mu + e_t + \Theta e_{t-1} \\ &= (I - \Phi)\mu + \Phi Y_{t-1} + e_t + \Theta e_{t-1} \\ &= (I - A - B)((I - A - B)^{-1} \omega) + \Phi Y_{t-1} + e_t + \Theta e_{t-1} \\ &= (I - A - B)(I - A - B)^{-1} \omega + \Phi Y_{t-1} + e_t + \Theta e_{t-1} \end{aligned}$$

L'écriture :

$$Y_t = \omega + \Phi Y_{t-1} + \Theta e_{t-1} + e_t \quad (\text{II.13})$$

avec $\Phi = B + A$, $\Theta = -A$, $\Phi = (\phi_{ij})$, $\Theta = (\theta_{ij})$,

correspond bien par identification à celle qu'on retrouve dans le cas univarié.

Nous voudrions maintenant introduire une proposition concernant une nouvelle méthode d'estimation pour le modèle **BP-INGARCH** (par MCO/MCG).

On décrit alors les étapes de cette dernière (la méthode plus générale comme posée par **LARDIC et MIGNON (2002)** est détaillée en annexe [Section A.3](#))

- **étape 1 :** On pose $X_t := Y_t - \bar{Y}$ le processus centré de Y_t , où $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$ et n est le nombre d'observations disponibles.

La représentation $VAR(p^*)$ bivariée de $X_t = (X_{t,1}, X_{t,2})^T$, donnée par :

$$X_{t,j} = \sum_{i=1}^{p^*} a_{i,j} X_{t-i} + e_{t,j}, j = 1, 2. \quad (\text{II.14})$$

(où $a_{i,j}$, $i = 1, \dots, p^*$, $j = 1, 2$ la j -ème ligne de la matrice a_i de taille (2×2) et $e_{t,j}$ aussi la j -ème composante du vecteur bivarié e_t .)

Nous procédons à un ajustement du VAR pour obtenir les estimations \hat{e}_t qui sont définis par :

$$\hat{e}_{t,j} = X_{t,j} - \sum_{i=1}^{p^*} \hat{a}_{i,j} X_{t-i}, j = 1, 2.$$

- **étape 2 :** D'après l'équation (II.13), on a la représentation suivante :

$$\begin{aligned} Y_{t,1} &= \omega_1 + \phi_{1,1}Y_{t-1,1} + \phi_{1,2}Y_{t-1,2} + \theta_{1,1}e_{t-1,1} + \theta_{1,2}e_{t-1,2} + e_{t,1}, \\ Y_{t,2} &= \omega_2 + \phi_{2,1}Y_{t-1,1} + \phi_{2,2}Y_{t-1,2} + \theta_{2,1}e_{t-1,1} + \theta_{2,2}e_{t-1,2} + e_{t,2}, \end{aligned} \quad (\text{II.15})$$

deux équations décrivant les composantes de notre série bivariée Y_t .

On pose :

$$Y'_t := (1, Y_{t-1,1}, Y_{t-1,2}, e_{t-1,1}, e_{t-1,2}) \quad (\text{II.16})$$

, un vecteur de taille (1×5) composantes contenant les *régresseurs*.

On pose aussi :

$$\psi_1 = \begin{pmatrix} \omega_1 \\ \phi_{1,1} \\ \phi_{1,2} \\ \theta_{1,1} \\ \theta_{1,2} \end{pmatrix}, \psi_2 = \begin{pmatrix} \omega_2 \\ \phi_{2,1} \\ \phi_{2,2} \\ \theta_{2,1} \\ \theta_{2,2} \end{pmatrix}.$$

deux vecteurs de taille (5×1) contenant les coefficients de ces régresseurs pour chaque équation, qui sont les paramètres qu'on voudrait estimer.

d'où

$$Y_{t,1} = Y'_t \psi_1 + e_{t,1},$$

$$Y_{t,2} = Y'_t \psi_2 + e_{t,2},$$

On pose maintenant :

$$\underline{Y}_1 = \begin{pmatrix} Y_{1,1} \\ Y_{2,1} \\ \vdots \\ Y_{T,1} \end{pmatrix}, \underline{Y}_2 = \begin{pmatrix} Y_{1,2} \\ Y_{2,2} \\ \vdots \\ Y_{T,2} \end{pmatrix} \quad (\text{II.17})$$

deux vecteurs de taille $(T \times 1)$ contenant les observations de chaque composante.

et

$$\underline{Y} = \begin{pmatrix} Y'_1 \\ Y'_2 \\ \vdots \\ Y'_T \end{pmatrix} = \begin{pmatrix} 1 & Y_{0,1} & Y_{0,2} & e_{0,1} & e_{0,2} \\ 1 & Y_{1,1} & Y_{1,2} & e_{1,1} & e_{1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & Y_{T-1,1} & Y_{T-1,2} & e_{T-1,1} & e_{T-1,2} \end{pmatrix} \quad (\text{II.18})$$

ainsi que

$$e_j = \begin{pmatrix} e_{1,j} \\ e_{2,j} \\ \vdots \\ e_{T,j} \end{pmatrix} \quad (\text{II.19})$$

avec \underline{Y} une matrice de taille $(T \times 5)$ et e_j un vecteur de taille T .

On a alors :

$$\begin{pmatrix} \underline{Y}_1 \\ \underline{Y}_2 \end{pmatrix} = \begin{pmatrix} \underline{Y} & 0 \\ 0 & \underline{Y} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

On pose : $\Psi := \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$

En remplaçant e_j par son estimation \hat{e}_j obtenue dans l'étape 1, l'estimateur par MCO bloc par bloc, qui est équivalent à l'estimateur MCG d'après le théorème de Zellner, pour ce cas est donné par :

$$\hat{\Psi}_{MCO} = \begin{pmatrix} (\underline{Y}^T \underline{Y})^{-1} \underline{Y}^T \underline{Y}_1 \\ (\underline{Y}^T \underline{Y})^{-1} \underline{Y}^T \underline{Y}_2 \end{pmatrix} \quad (\text{II.20})$$

avec $\hat{\Psi}_{MCO}$ un vecteur de taille (10×1) .

où $\hat{\psi}_i = (\hat{\omega}_1, \hat{\phi}_{i,1}, \hat{\phi}_{i,2}, \hat{\theta}_{i,1}, \hat{\theta}_{i,2})^T$, $i = 1, 2$.

à la fin de cette procédure nous aurons estimé les paramètres du modèle ARMA bivarié (II.13), et par transformations simples ne restera qu'à obtenir les paramètre d'intérêt du modèle II.2 :

$$\hat{A} = -\hat{\Theta}, \quad \hat{B} = \hat{\Phi} + \hat{\Theta}$$

Nous avançons alors la proposition suivante :

Proposition II.1 (*estimateur MCO*)

Pour un modèle BP-INGARCH comme défini par : (II.2). Et comme défini dans l'écriture (II.13)

Soit $\Psi = (\psi_1^T, \psi_2^T)^T \in \mathbb{R}_+^{10}$, avec $\psi_i = (\omega_i, \phi_{i,1}, \phi_{i,2}, \theta_{i,1}, \theta_{i,2})^T$, pour $i = 1, 2$. L'estimateur de Ψ par la méthode de Moindres Carrés Ordinaires $\hat{\Psi}_{MCO}$ est donnée par les deux étapes suivantes :

- **étape 1** : Estimation des erreurs e_t : on définit $X_t := Y_t - \sum_{i=1}^n Y_i/n$ puis ajustons X_t par un VAR bivarié d'ordre p^* : $X_t = \sum_{i=1}^{p^*} a_i \hat{X}_{t-i} + e_t$. pour obtenir les estimateurs des coefficients \hat{a}_i .

On définit ensuite : $\hat{e}_{t,j} = X_{t,j} - \sum_{i=1}^{p^*} \hat{a}_{i,j} X_{t-i,j}$, $j = 1, 2..$

- **étape 2** : On remplace les erreurs par leurs estimations $\hat{e}_{t,j}$ obtenues par la 1ère étape dans (II.13). L'estimateur est obtenu par un MCO bloc par bloc (LARDIC et MIGNON (2002)) :

On pose les mêmes vecteurs et matrices définis en : (II.16), (II.17), (II.18), (II.19), (II.20). et on a :

$$\hat{\Psi}_{MCO} = \begin{pmatrix} \hat{\psi}_{1MCO} \\ \hat{\psi}_{2MCO} \end{pmatrix}$$

$$\hat{\psi}_{1MCO} = (\hat{\omega}_1, \hat{\phi}_{1,1}, \hat{\phi}_{1,2}, \hat{\theta}_{1,1}, \hat{\theta}_{1,2})^T,$$

$$\hat{\psi}_{2MCO} = (\hat{\omega}_2, \hat{\phi}_{2,1}, \hat{\phi}_{2,2}, \hat{\theta}_{2,1}, \hat{\theta}_{2,2})^T.$$

Enfin :

$$\begin{cases} \hat{A}_{MCO} = -\hat{\Theta}, \\ \hat{B}_{MCO} = \hat{\Phi} + \hat{\Theta}. \end{cases}$$

D'après le théorème de Zellner ([LARDIC et MIGNON \(2002\)](#)) cet estimateur $\hat{\Psi}_{MCO}$ est équivalent à un estimateur **MCG** résultat de la forme de la matrice de variables explicatives.

(La preuve du théorème de Zellner est en annexe : [Section A.4](#), ainsi que la procédure qui peut être suivie pour ajuster le VAR à l'étape 1 [Section A.3](#))

Remarque II.1

Cette méthode d'estimation est une méthode **directe** qui, issu de produits matriciels, nous fournit un résultat fixe pour l'estimation des paramètres $\psi_i = (\omega_i, \phi_{i,1}, \phi_{i,2}, \theta_{i,1}, \theta_{i,2})^T$, ($i = 1, 2$) (sans le δ qui est paramètre de la distribution conditionnelle).

Lorsque certaines spécifications du modèle supposent à priori certains paramètres comme étant nuls (par exemple, des matrices de paramètres diagonales), le résultat reste inchangé. Une méthode plus flexible, capable de s'adapter à ces changements, pourrait donc être plus judicieuse. La section suivante ([Section II.3.2](#)) traitera justement d'une telle méthode.

Néanmoins, l'annexe [Section A.7](#) montrera qu'une adaptation au cas particulier où les deux matrices de paramètres sont diagonales est possible

II.3.2 Estimation par la méthode de Maximum de Vraisemblance

Cette méthode a été utilisé par [CUI et ZHU \(2018\)](#) pour estimer les paramètres du modèle **BP-INGARCH**.

II.3.2.1 La fonction de log-Vraisemblance

On a (Y_1, Y_2, \dots, Y_n) des observations d'un processus suivant le modèle ([II.2](#)), et sous l'hypothèse que A, B sont des matrices non négatives alors le vecteur des paramètres est :

$$\theta = (\theta_1, \theta_2, \dots, \theta_{11}) = (\omega_1, \alpha_{11}, \alpha_{12}, \beta_{11}, \beta_{12}, \omega_2, \alpha_{21}, \alpha_{22}, \beta_{21}, \beta_{22}, \delta)^T$$

et le vecteur des vraies valeurs est :

$$\theta^0 = (\omega_1^0, \alpha_{11}^0, \alpha_{12}^0, \beta_{11}^0, \beta_{12}^0, \omega_2^0, \alpha_{21}^0, \alpha_{22}^0, \beta_{21}^0, \beta_{22}^0, \delta^0)$$

La fonction de Vraisemblance conditionnellement à $\lambda_1 = (\lambda_{1,1}, \lambda_{1,2})^T$:

$$\mathbb{L}(\theta | Y_1, \dots, Y_n, \lambda_1) = \prod_{t=2}^n \frac{\lambda_{t,1}^{Y_{t,1}} \lambda_{t,2}^{Y_{t,2}}}{Y_{t,1}! Y_{t,2}!} \exp\{-(\lambda_{t,1} + \lambda_{t,2})\} \varphi_t \quad (\text{II.21})$$

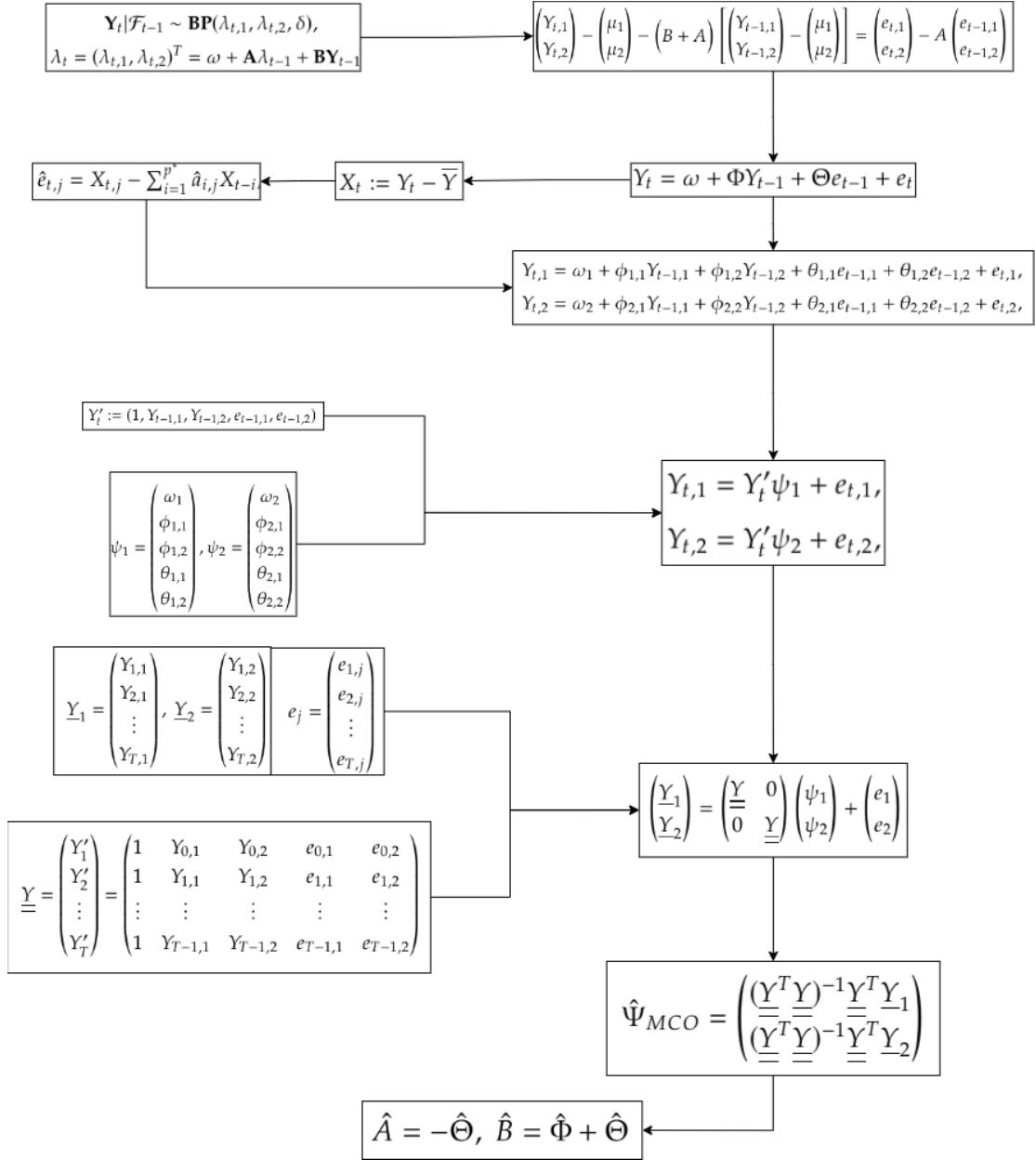


FIGURE II.1 – Schéma résumant les étapes de la procédure d'estimation par MCO

Avec $\varphi = \varphi_t(\theta) = 1 + \delta(e^{-Y_{t,1}} - e^{-cY_{t,1}})(e^{-Y_{t,2}} - e^{-cY_{t,2}})$.

La fonction log-Vraisemblance (à une constante près) est donnée par :

$$\mathcal{L}(\theta) = \sum_{t=2}^n [Y_{t,1} \ln \lambda_{t,1}(\theta) + Y_{t,2} \ln \lambda_{t,2}(\theta) - \lambda_{t,1}(\theta) - \lambda_{t,2}(\theta) + \ln \phi_t(\theta)] \quad (\text{II.22})$$

La fonction score est :

$$S_n(\theta) = \frac{\partial \mathcal{L}(\theta)}{\partial \theta} = \sum_{t=2}^n \frac{\partial \mathcal{L}_t(\theta)}{\partial \theta}$$

$\hat{\theta}$ est l'estimateur de Maximum de Vraisemblance est solution de l'équation $S_n(\theta) = 0$.

La matrice Hessienne est définie par :

$$H_n(\theta) = - \sum_{t=2}^n \frac{\partial^2 \mathcal{L}_t(\theta)}{\partial \theta \partial \theta^T}$$

II.3.2.2 Les propriétés asymptotiques de l'estimateur de Maximum de Vraisemblance

Pour étudier les propriétés asymptotiques de l'EMV $\hat{\theta}$ **CUI et ZHU (2018)** propose d'introduire des bornes supérieures et inférieures pour chaque composante de θ^0 :

$$0 < \omega_L \leq \omega_i^0 \leq \omega_U, 0 < \alpha_L \leq \alpha_{ij}^0 \leq \alpha_U, 0 < \beta_L \leq \beta_{ij}^0 \leq \beta_U, i, j = 1, 2.$$

Pour formuler les résultats, on pose les hypothèses suivantes :

- **H1** : Un voisinage compact de θ^0 est défini par $\Theta(\theta^0) = \{\theta | 0 < \omega_L \leq \omega_i \leq \omega_U, 0 < \alpha_L \leq \alpha_{ij} \leq \alpha_U, 0 < \beta_L \leq \beta_{ij} \leq \beta_U, \text{ et } |\delta| \leq \delta_U, i, j = 1, 2\}$.
- **H2** : $0 < \varphi_L \leq |\varphi_t| \leq \varphi_U$.
- **H3** : Il existe $p \in [1, \infty]$ tel que $\|A\|_p + 2^{(1-1/p)} \|B\|_p < 1$.

Nous commençons par introduire quelques lemmes afin d'établir les propriétés asymptotiques de l'estimateur $\hat{\theta}$.

Lemmes page 6 Cui et Zhu (2018)

- **Lemme 1** : La fonction score évalué à la vraie valeur de $\theta = \theta^0$ satisfait :

$$\frac{1}{\sqrt{n}} S_n(\theta) \xrightarrow{d} N(0, G)$$

avec G est une matrice tel que :

$$G(\theta) = E \left(\frac{\partial \mathcal{L}_t(\theta)}{\partial \theta} \frac{\partial \mathcal{L}_t(\theta)}{\partial \theta^T} \right)$$

Un estimateur consistant de G est donné par :

$$G_t(\hat{\theta}), \text{ où } G_t(\theta) = \text{Var} \left(\frac{\partial \mathcal{L}_t}{\partial \theta} | F_{t-1} \right).$$

- **Lemme 2 :** La matrice Hessienne évalué à la vraie valeur de $\theta = \theta^0$ satisfait :

$$\frac{1}{n} H_n(\theta) \xrightarrow{p} G$$

- **Lemme 3 :** avec le voisinage $\Theta(\theta^0)$, et sous **H1 – H3** on a :

$$\max_{i,j,k=1,\dots,11} \sup_{\theta \in \Theta(\theta^0)} \left| \frac{1}{n} \sum_{t=1}^n \frac{\partial \mathcal{L}_t(\theta)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| \leq M_n := \frac{1}{n} \sum_{t=1}^n m_t$$

de plus :

$$\begin{aligned} m_t &= C g(Y_{t,i}) h(\mu_{ts}), \\ \mu_{ts} &= \beta_U \sum_{j=1}^{t-s} k_{j,s} \alpha_U^{j-1} Y_{t-s-j,i}, s = 0, 1, 2, 3, \\ k_{j,0} &= 0, k_{j,1} = j, k_{j,3} = j(j+1)(j+2) \end{aligned}$$

Avec $g(\cdot)$ et $h(\cdot)$ des fonctions de $Y_{t,i}$ et μ_{ts} respectivement. $0 \leq M_n \longrightarrow M$, M et C sont deux constantes.

On a alors le théorème suivant, qui porte sur les propriétés de l'estimateur du maximum de vraisemblance $\hat{\theta}$.

Théorème II.2 (*Cui et ZHU, 2018, Théorème 2, page 8*)

Considérons le modèle (II.2) d'ordre (1,1), et supposant que à la vraie valeur θ^0 , les hypothèses **H1 – H3** sont vérifiées. Alors il existe un voisinage fixé $\eta(\theta^0)$ tel que la fonction de log-Vraisemblance a un unique maximum $\hat{\theta}$ avec une probabilité qui tend vers 1.

De plus $\hat{\theta}$ est \sqrt{n} consistant et asymptotiquement normal :

$$\sqrt{n}(\hat{\theta} - \theta^0) \xrightarrow{d} N(0, G^{-1})$$

En ce qui concerne la prévision, nous utilisons la méthode de l'espérance conditionnelle pour prédire les taux de Poisson. En utilisant la formule de récurrence (II.4), une prévision naturelle à un pas de λ_t est donnée par :

$$\hat{\lambda}_{t+1} = E(\lambda_{t+1}|F_t) = \omega + A\lambda_t + BY_t = \sum_{k=0}^{t-1} (A^k \omega + A^k BY_{t-k}) + A^t \lambda_1,$$

où les matrices de paramètres sont remplacées par leurs estimations correspondantes. Plus généralement, une prévision à h pas de λ_t , pour un certain $h > 1$, peut être obtenue par le même schéma récursif.

III

ILLUSTRATION NUMÉRIQUE

Ce chapitre est consacré à l’illustration numérique du modèle **BP-INGARCH**.

Les deux sections principales portent respectivement sur une étude de simulation et une application à des données réelles.

La première section, présentée en [Section III.1](#), est dédiée à une étude de simulation. Elle utilise les langages Python et Matlab pour explorer les propriétés et le comportement du modèle. L’objectif est de générer des réalisations du modèle et des trajectoires possibles, permettant ainsi de visualiser son fonctionnement dans un environnement contrôlé et d’évaluer les méthodes d’estimation développées dans ce travail. Elle offre ainsi une validation des résultats théoriques.

La seconde section, présentée en [Section III.2](#), explore l’application du modèle à un ensemble de données réelles afin d’en démontrer l’utilité pratique.

III.1 Étude de simulation

Une étude de simulation du BP-INGARCH(1,1) est menée pour évaluer les performances des estimateurs par la méthode MCO et MV, appliquée sur quatres différents modèles (M1, M2, M3, et M4).

L'algorithme d'estimation par MCO a été programmé sur "Python", le package "statsmodels" est aussi utilisé pour comparer les résultats d'ajustement du $VAR(p^*)$ (illustré en annexe Section A.6).

La fonction "fmincon" sur Matlab est utilisé pour la maximisation de la fonction de log-Vraisemblance, où les estimations obtenues par MCO peuvent être insérées comme valeurs initiales pour permettre à la routine d'optimisation d'arriver aux résultats plus vite.

Des trajectoires de tailles différentes sont simulées : n=200, 500 et 1000, avec r=500 réplications pour chaque modèle. L'erreur moyenne de déviation absolue (MADE : *Mean Absolute Deviation Error*) et l'erreur quadratique moyenne (MSE : *Mean Squared Error*) sont calculées pour évaluer les performances des estimateurs selon les formules suivantes :

$$\text{MADE} = \frac{1}{m} \sum_{j=1}^m |\hat{\theta}_j - \theta_0|, \quad \text{MSE} = \frac{1}{m} \sum_{j=1}^m (\hat{\theta}_j - \theta_0)^2,$$

où $\hat{\theta}_j$ est l'estimateur de θ_0 dans la j -ème réplication.

Considérons les configurations de paramètres suivantes :

- (M1) : **A** et **B** sont non-diagonales :

$$\begin{aligned}\theta &= (\omega_1, \alpha_{11}, \alpha_{12}, \beta_{11}, \beta_{12}, \omega_2, \alpha_{21}, \alpha_{22}, \beta_{21}, \beta_{22}, \delta)^T \\ &= (0.5, 0.3, 0.2, 0.1, 0.2, 0.3, 0.1, 0.3, 0.3, 0.2, 0.4)^T.\end{aligned}$$

- (M2) : **A** et **B** sont diagonales :

$$\begin{aligned}\theta &= (\omega_1, \alpha_{11}, \beta_{11}, \omega_2, \alpha_{22}, \beta_{22}, \delta)^T \\ &= (0.3, 0.2, 0.5, 0.5, 0.4, 0.3, 0.7)^T.\end{aligned}$$

- (M3) : **A** diagonale et **B** non-diagonale :

$$\begin{aligned}\theta &= (\omega_1, \alpha_{11}, \beta_{11}, \beta_{12}, \omega_2, \alpha_{22}, \beta_{21}, \beta_{22}, \delta)^T \\ &= (0.3, 0.2, 0.5, 0.1, 0.5, 0.4, 0.3, 0.2, 0.7)^T.\end{aligned}$$

- (M4) : **A** non-diagonale et **B** diagonale :

$$\begin{aligned}\theta &= (\omega_1, \alpha_{11}, \alpha_{12}, \beta_{11}, \omega_2, \alpha_{21}, \alpha_{22}, \beta_{22}, \delta)^T \\ &= (0.5, 0.2, 0.3, 0.2, 0.3, 0.4, 0.2, 0.3, -0.5)^T.\end{aligned}$$

En utilisant l'algorithme proposé par **Cui et ZHU (2018)** (détailé en annexe **Section A.5**), pour générer un échantillon de la distribution bivarié $BP(\lambda_1, \lambda_2, \delta)$, on peut simuler notre modèle d'intérêt : **BP-INGARCH(1,1)**.

Les figures III.1 et III.2 illustrent le comportement de la série temporelle bivariée $Y_t = (Y_{t,1}, Y_{t,2})$ des modèles M1, M2, M3 et M4 d'abord par des graphes distincts pour chaque composante, et ensuite combinés.

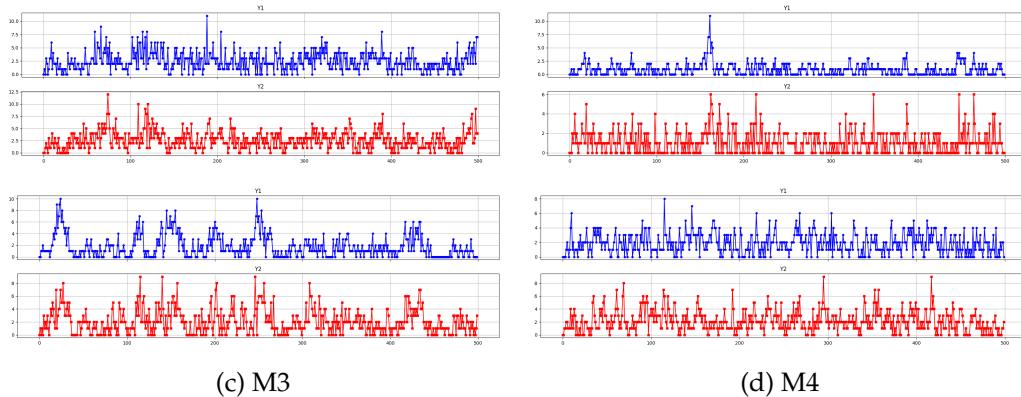


FIGURE III.1 – Graphes des séries temporelles Y_1 (bleu), Y_2 (rouge), composantes de la série bivariée.

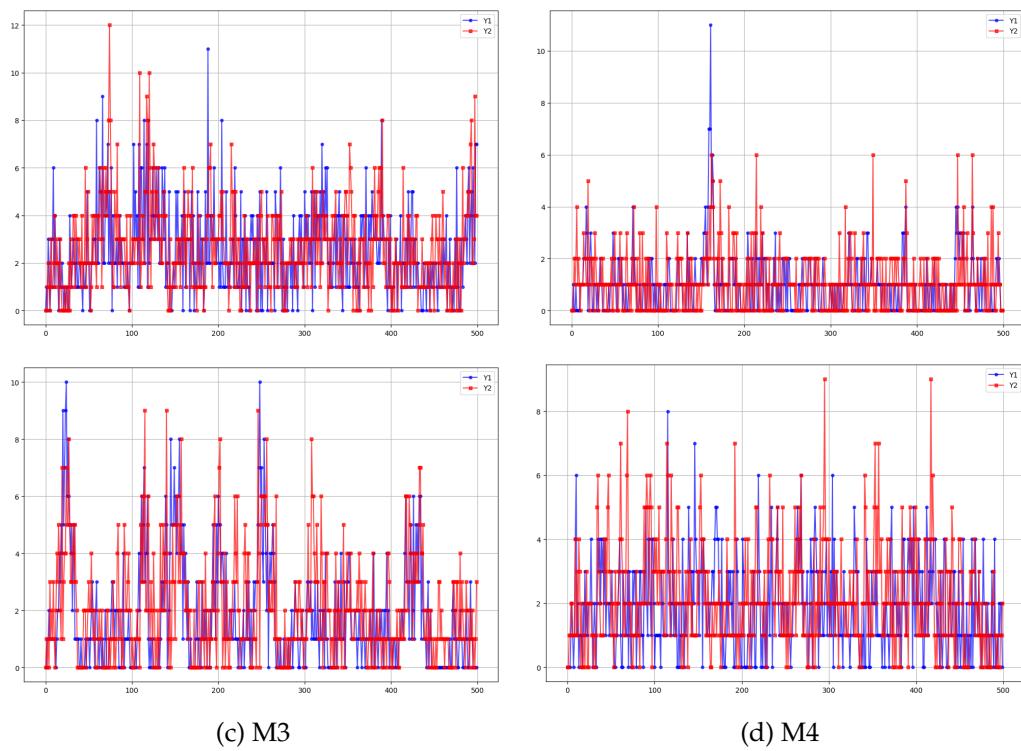


FIGURE III.2 – Graphes des séries temporelles bivarié des modèles M1-M4.

Modèle	n	Statistique	ω_1	α_{11}	α_{12}	β_{11}	β_{12}	ω_2	α_{21}	α_{22}	β_{21}	β_{22}
200	Moyenne	0.6597	0.2391	0.2020	0.0954	0.2497	0.4019	0.1502	0.2243	0.2550	0.1899	
	Biais	0.1597	-0.0609	0.0020	-0.0046	0.0497	0.1019	0.0502	-0.0757	-0.0450	-0.0101	
	Écart type	0.3820	0.3356	0.2813	0.0757	0.4940	0.3865	0.3719	0.2969	0.4959	0.0799	
	MADE	0.3027	0.2657	0.2194	0.0598	0.3807	0.2946	0.2843	0.2323	0.3757	0.06367	
	MSE	0.1714	0.1163	0.0791	0.0057	0.2464	0.1597	0.1408	0.0938	0.2479	0.0064	
	Moyenne	0.5791	0.2423	0.2341	0.0974	0.2037	0.3437	0.1321	0.2472	0.2991	0.1995	
500	Biais	0.0791	-0.0577	0.0341	-0.0026	0.0037	0.0437	0.0321	-0.0528	-0.0009	-0.0005	
	Écart type	0.2691	0.2684	0.2179	0.0479	0.3865	0.2527	0.2656	0.2241	0.3733	0.0481	
	MADE	0.2003	0.2117	0.1741	0.0387	0.3047	0.2015	0.2079	0.1781	0.2939	0.0381	
	MSE	0.0786	0.0753	0.0486	0.0023	0.1493	0.0657	0.0715	0.0529	0.1393	0.0023	
	Moyenne	0.5466	0.2628	0.2197	0.0988	0.1992	0.3199	0.1159	0.2748	0.3045	0.1998	
	Biais	0.0466	-0.0372	0.0197	-0.0012	-0.0008	0.0199	0.0159	-0.0252	0.0045	-0.0002	
1000	Écart type	0.1689	0.2117	0.1751	0.0331	0.2957	0.1602	0.2118	0.1748	0.2900	0.0336	
	MADE	0.1312	0.1708	0.1399	0.0267	0.2307	0.1258	0.1633	0.1400	0.2242	0.0272	
	MSE	0.03070	0.0461	0.0310	0.0010	0.0874	0.0260	0.0450	0.0311	0.0841	0.0011	
	Vraies Valeurs	0.5	0.3	0.2	0.1	0.2	0.3	0.1	0.3	0.3	0.2	

TABLE III.1 – Résultats d'estimation MCO pour le modèle M1

Modèle	n	Statistique	ω_1	α_{11}	β_{11}	ω_2	α_{22}	β_{22}
200	Moyenne	0.3522	0.1551	0.4911	0.6696	0.3092	0.2894	
	Biais	0.0523	-0.0449	-0.0089	0.1696	-0.0908	-0.0106	
	Écart type	0.1071	0.1389	0.0952	0.2857	0.1916	0.0832	
	MADÉ	0.0904	0.1174	0.0775	0.2366	0.1609	0.0667	
	MSE	0.0142	0.0213	0.0091	0.1104	0.0450	0.0070	
	Moyenne	0.3201	0.1878	0.4901	0.5668	0.3612	0.2979	
500	Biais	0.0201	-0.0122	-0.0099	0.0668	-0.0388	-0.0021	
	Écart type	0.0685	0.0914	0.0555	0.1520	0.1228	0.0528	
	MADÉ	0.0537	0.0726	0.0455	0.1285	0.0999	0.0421	
	MSE	0.0051	0.0085	0.0032	0.0276	0.0166	0.0028	
	Moyenne	0.3116	0.1918	0.4952	0.5369	0.3814	0.2967	
	Biais	0.0116	-0.0082	-0.0048	0.0369	-0.0186	-0.0033	
1000	Écart type	0.0502	0.0671	0.0440	0.1061	0.0809	0.0364	
	MADÉ	0.0398	0.0537	0.0353	0.0891	0.0671	0.0297	
	MSE	0.0027	0.0046	0.0020	0.0126	0.0069	0.0013	
	Vraies Valeurs	0.3	0.2	0.5	0.5	0.4	0.3	

TABLE III.2 – Résultats d'estimation MCO pour le modèle M2.

Modèle	n	Statistique	ω_1	α_{11}	β_{11}	β_{12}	ω_2	α_{22}	β_{21}	β_{22}
200	Moyenne	0.3602	0.1177	0.4887	0.4274	0.6763	0.2299	-0.0256	0.1834	
	Biais	0.0602	-0.0823	-0.0113	0.3274	0.1763	-0.1701	-0.3256	-0.0166	
	Écart type	0.2768	0.2832	0.0918	0.4290	0.3274	0.2880	0.4389	0.0823	
	MADÉ	0.2135	0.2191	0.0729	0.4255	0.2651	0.2523	0.4345	0.0662	
	MSE	0.0803	0.0870	0.0085	0.2913	0.1383	0.1119	0.2986	0.0071	
	Moyenne	0.3126	0.1529	0.4890	0.3587	0.5890	0.3021	0.0396	0.1996	
500	Biais	0.0126	-0.0471	-0.0110	0.2587	0.0890	-0.0979	-0.2604	-0.0004	
	Écart type	0.1730	0.2030	0.0578	0.3080	0.1880	0.2002	0.3137	0.0504	
	MADÉ	0.1300	0.1625	0.0475	0.3235	0.1518	0.1727	0.3323	0.0402	
	MSE	0.0301	0.0434	0.0035	0.1618	0.0433	0.0497	0.1663	0.0025	
	Moyenne	0.3070	0.1675	0.4983	0.3354	0.5608	0.3425	0.0657	0.1973	
	Biais	0.0070	-0.0325	-0.0017	0.2354	0.0608	-0.0575	-0.2343	-0.0027	
1000	Écart type	0.1194	0.1363	0.0403	0.2142	0.1414	0.1591	0.2240	0.0355	
	MADÉ	0.0945	0.1113	0.0318	0.2638	0.1143	0.1314	0.2709	0.0290	
	MSE	0.0143	0.0196	0.0016	0.1013	0.0237	0.0286	0.1050	0.0013	
	Vraies Valeurs	0.3	0.2	0.5	0.1	0.5	0.4	0.3	0.2	

TABLE III.3 – Résultats d'estimation MCO pour le modèle M3.

Modèle	n	Statistique	α_1	α_{11}	α_{12}	β_{11}	ω_2	α_{21}	α_{22}	β_{22}
200	Moyenne	Moyenne	0.6999	0.1304	0.2871	0.1872	0.5801	0.3111	0.1499	0.2907
		Biais	0.1999	-0.0696	-0.0129	-0.0128	0.2801	-0.0889	-0.0501	-0.0093
		Écart type	0.5677	0.3007	0.2301	0.0817	0.6040	0.3534	0.2395	0.0799
	MADE	MADE	0.4495	0.2356	0.1770	0.0651	0.4972	0.2770	0.1882	0.0638
		MSE	0.3622	0.0953	0.0531	0.0068	0.4432	0.1328	0.0599	0.0065
		Moyenne	0.6006	0.1581	0.2906	0.1982	0.4127	0.3668	0.1833	0.2961
M4	Biais	Biais	0.1006	-0.0419	-0.0094	-0.0018	0.1127	-0.0332	-0.0167	-0.0039
		Écart type	0.3396	0.2033	0.1394	0.0495	0.3374	0.2037	0.1471	0.0476
		MADE	0.2653	0.1616	0.1123	0.0392	0.2758	0.1600	0.1143	0.0385
	MSE	MSE	0.1254	0.0431	0.0195	0.0025	0.1265	0.0426	0.0219	0.0023
		Moyenne	0.5424	0.1804	0.2984	0.1978	0.3600	0.3806	0.1843	0.2999
		Biais	0.0424	-0.0196	-0.0016	-0.0022	0.0600	-0.0194	-0.0157	-0.0001
1000	Vraies Valeurs	Écart type	0.2252	0.1348	0.1018	0.0338	0.2359	0.1516	0.1043	0.0368
		MADE	0.1800	0.1073	0.0800	0.0266	0.1939	0.1205	0.0847	0.0294
		MSE	0.0525	0.0185	0.0104	0.0011	0.0593	0.0234	0.0111	0.0014
	Vraies Valeurs	0.5	0.2	0.3	0.2	0.3	0.4	0.2	0.3	

TABLE III.4 – Résultats d'estimation MCO pour le modèle M4.

On remarque des valeurs du biais faibles même pour des tailles d'échantillon plus petites ($n=200$). On observe aussi une tendance générale à la diminution du biais, de l'écart type, du MADE et du MSE lorsque la taille de l'échantillon augmente. Cela indique que l'estimation des paramètres devient plus précise et que les modèles sont plus performants avec davantage de données.

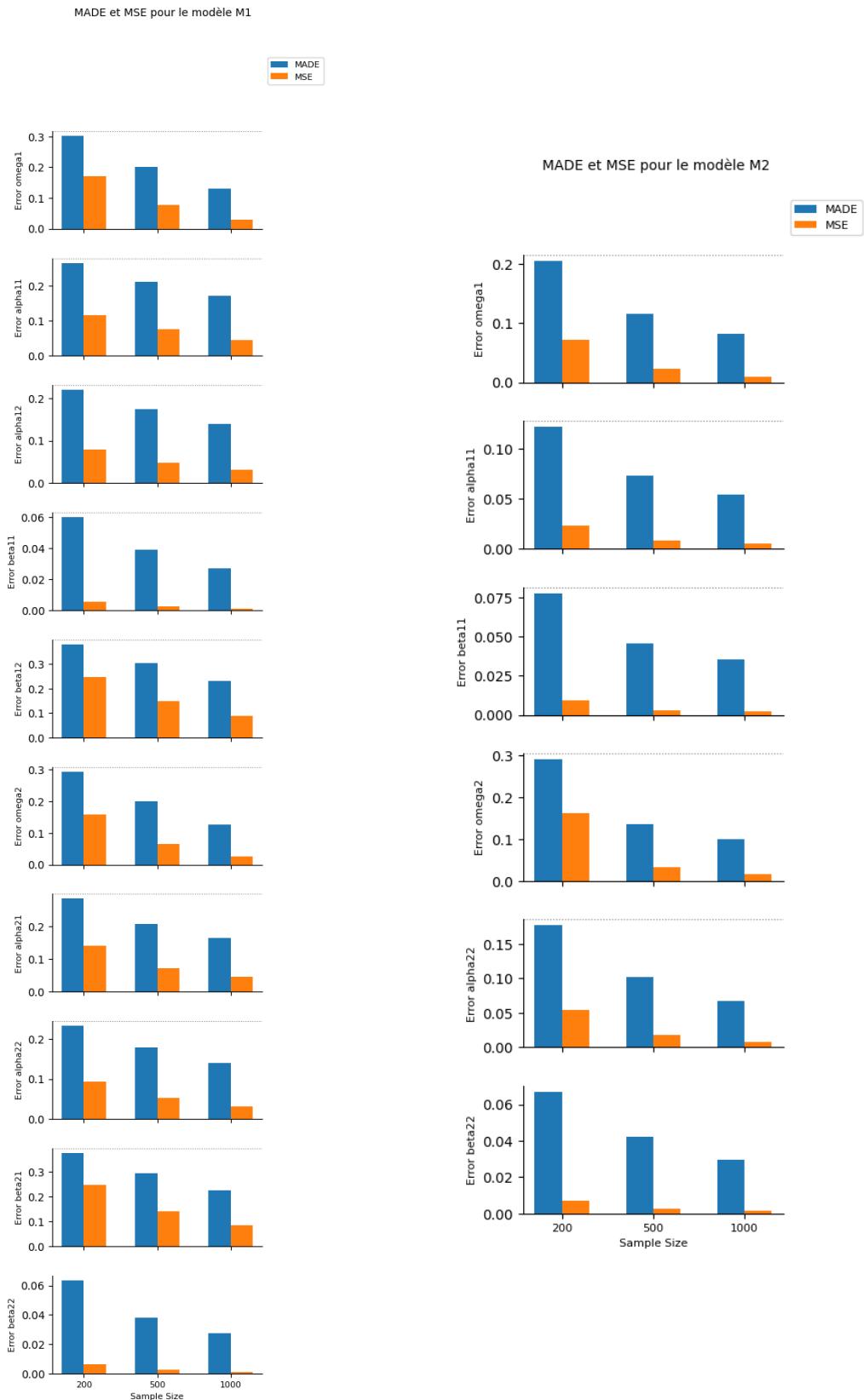


FIGURE III.3 – L'évaluation de MADE et MSE des estimateurs pour les modèles M1-M2

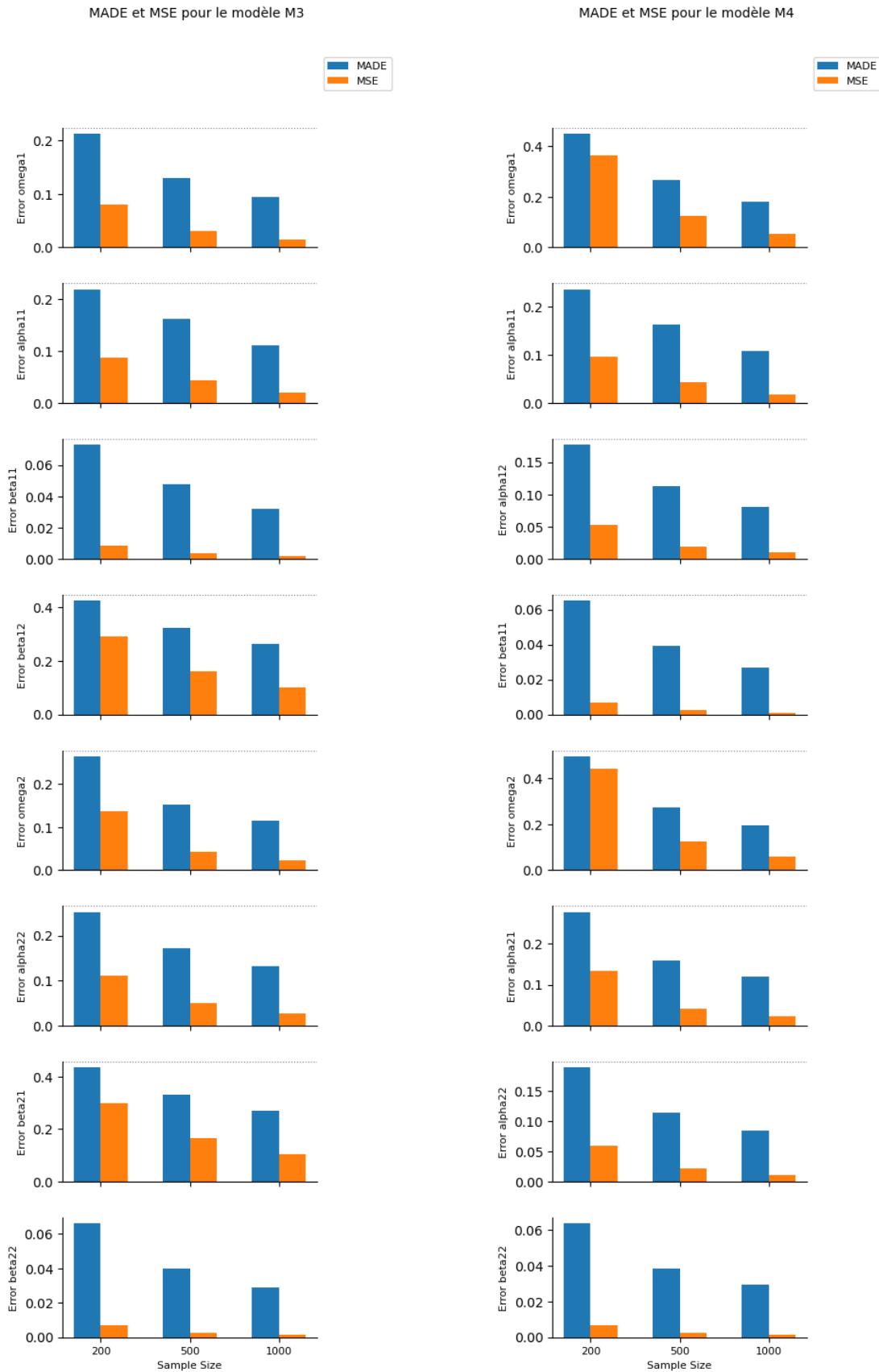


FIGURE III.4 – L'évaluation de MADE et MSE des estimateurs pour les modèles M3-M4

les figures III.3, III.4 représentent de manière graphique l'évolution des erreurs par rapport à la taille de l'échantillon.

Les résultats suggèrent que les modèles MCO peuvent fournir des estimations raisonnables des paramètres pour tous les modèles considérés. Cependant, la précision de l'estimation varie en fonction du modèle et du paramètre spécifique. La taille de l'échantillon joue un rôle crucial dans la précision de l'estimation, avec des estimations plus précises obtenues avec des échantillons plus importants.

Les estimateurs $\hat{\theta}_{MCO}$ pour chaque modèle sont maintenant utilisés comme valeurs initiales pour la méthode du maximum de vraisemblance. Comme observé par les résultats précédents, celles-ci s'approchent des vraies valeurs.

Les résultats de l'estimation par cette méthode sont présentés dans les tableaux III.5, III.6, III.7 et III.8.

Modèle	n	Statistique	α_{11}	α_{12}	β_{11}	β_{12}	α_{21}	α_{22}	β_{21}	β_{22}	δ
200	Moyenne	0.4654	0.3528	0.1728	0.0901	0.1970	0.2905	0.1502	0.2691	0.3024	0.1826
		Biais	-0.0346	0.0528	-0.0272	-0.0099	-0.0030	-0.0095	0.0502	-0.0309	0.0024
		Écart type	0.2103	0.1969	0.1647	0.0621	0.0749	0.2019	0.1528	0.1621	0.0744
	MADE	0.1742	0.1685	0.1416	0.0518	0.0591	0.1607	0.1221	0.1353	0.0578	0.0622
		MSE	0.0454	0.0416	0.0279	0.0039	0.0056	0.0409	0.0259	0.0272	0.0055
		Moyenne	0.4877	0.3211	0.1927	0.0945	0.1964	0.2970	0.1422	0.2580	0.3069
M1	Biais	-0.0123	0.0211	-0.0073	-0.0055	-0.0036	-0.0030	0.0422	-0.0420	0.0069	-0.0053
		Écart type	0.1644	0.1699	0.1463	0.0437	0.0467	0.1578	0.1395	0.1341	0.0427
		MADE	0.1320	0.1452	0.1242	0.0357	0.0374	0.1262	0.1127	0.1122	0.0347
	MSE	0.0272	0.0293	0.0215	0.0019	0.0022	0.0249	0.0212	0.0198	0.0019	0.0020
		Moyenne	0.4973	0.3128	0.1918	0.0967	0.1992	0.2843	0.1373	0.2680	0.3036
		Biais	-0.0027	0.0128	-0.0082	-0.0033	-0.0008	-0.0157	0.0373	-0.0320	0.0036
1000	Écart type	0.1275	0.1598	0.1352	0.0315	0.0310	0.1112	0.1276	0.1170	0.0309	0.0318
		MADE	0.1021	0.1345	0.1137	0.0254	0.0249	0.0877	0.1046	0.0974	0.0249
		MSE	0.0163	0.0257	0.0184	0.0010	0.0010	0.0126	0.0177	0.0147	0.0010
	Vraies Valeurs	0.5	0.3	0.2	0.1	0.2	0.3	0.1	0.3	0.3	0.2
											0.4

TABLE III.5 – Résultats d'estimation MV pour le modèle M1 avec des matrices non-diagonales.

Modèle	n	Statistique	ω_1	α_{11}	β_{11}	ω_2	α_{22}	β_{22}	δ
200	Moyenne	0.3121	0.1932	0.4924	0.2886	0.5284	0.3918	0.7967	
		Biais	0.0121	-0.0068	-0.0076	-0.0114	0.0284	-0.0082	0.0967
		Écart type	0.0856	0.1132	0.0806	0.0759	0.1938	0.1537	0.5173
	MADE	0.0684	0.0932	0.0640	0.0616	0.1502	0.1220	0.4295	
		MSE	0.0075	0.0129	0.0065	0.0059	0.0384	0.0237	0.2769
		Moyenne	0.3022	0.2005	0.4942	0.2989	0.5186	0.3879	0.6916
500	M2	Biais	0.0022	0.0005	-0.0058	-0.0011	0.0186	-0.0121	-0.0084
		Écart type	0.0568	0.0774	0.0491	0.0469	0.1306	0.1072	0.3730
		MADE	0.0448	0.0625	0.0398	0.0378	0.1018	0.0839	0.3071
	Moyenne	MSE	0.0032	0.0060	0.0024	0.0022	0.0174	0.0116	0.1392
		Biais	0.0011	0.0007	-0.0039	-0.0028	0.0106	-0.0041	-0.0021
		Écart type	0.0398	0.0557	0.0378	0.0329	0.0910	0.0704	0.2798
1000	MADE	0.0322	0.0451	0.0307	0.0264	0.0729	0.0564	0.2247	
		MSE	0.0016	0.0031	0.0014	0.0011	0.0084	0.0050	0.0783
		Vraies Valeurs	0.3	0.2	0.5	0.3	0.5	0.4	0.7

TABLE III.6 – Résultats d'estimation MV pour le modèle M2 avec des matrices diagonales.

Modèle	n	Statistique	ω_1	α_{11}	β_{11}	β_{12}	ω_2	α_{22}	β_{21}	β_{22}	δ
200	Moyenne	0.3179	0.1921	0.4896	0.1017	0.4896	0.4178	0.3056	0.1795	0.4934	
	Biais	0.0179	-0.0079	-0.0104	0.0017	-0.0104	0.0178	0.0056	-0.0205	-0.2066	
	Écart type	0.1309	0.1143	0.0809	0.0527	0.1628	0.1244	0.0738	0.0732	0.5240	
	MADÉ	0.1028	0.0942	0.0650	0.0421	0.1288	0.0992	0.0588	0.0610	0.4103	
	MSE	0.0175	0.0131	0.0067	0.0028	0.0266	0.0158	0.0055	0.0058	0.3173	
	Moyenne	0.3115	0.1966	0.4916	0.1009	0.4950	0.4041	0.3026	0.1954	0.5801	
500	Biais	0.0115	-0.0034	-0.0084	0.0009	-0.0050	0.0041	0.0026	-0.0046	-0.1199	
	Écart type	0.0796	0.0757	0.0474	0.0337	0.1138	0.0803	0.0447	0.0447	0.3598	
	MADÉ	0.0637	0.0594	0.0386	0.0271	0.0904	0.0633	0.0355	0.0352	0.2929	
	MSE	0.0065	0.0057	0.0023	0.0011	0.0130	0.0065	0.0020	0.0020	0.1438	
	Moyenne	0.3088	0.1932	0.4989	0.1008	0.5061	0.4005	0.3011	0.1968	0.6573	
	Biais	0.0088	-0.0068	-0.0011	0.0008	0.0061	0.0005	0.0011	-0.0032	-0.0427	
1000	Écart type	0.0553	0.0551	0.0354	0.0234	0.0816	0.0567	0.0344	0.0310	0.2846	
	MADÉ	0.0436	0.0450	0.0284	0.0187	0.0645	0.0457	0.0274	0.0251	0.2351	
	MSE	0.0031	0.0031	0.0013	0.0005	0.0067	0.0032	0.0012	0.0010	0.0828	
	Vraies Valeurs	0.3	0.2	0.5	0.1	0.5	0.4	0.3	0.2	0.7	

TABLE III.7 – Résultats d'estimation MV pour le modèle M3 avec A matrice diagonale et B non-diagonale.

Modèle	n	Statistique	α_1	α_{11}	α_{12}	β_{11}	ω_2	α_{21}	α_{22}	β_{22}	δ
200	Moyenne	Moyenne	0.4299	0.2553	0.3036	0.1830	0.3161	0.3806	0.2209	0.2844	-0.3568
		Biais	-0.0701	0.0553	0.0036	-0.0170	0.0161	-0.0194	0.0209	-0.0156	0.1432
		Écart type	0.2402	0.1850	0.1628	0.0749	0.2348	0.1887	0.1592	0.0727	0.5077
	MADE	MADE	0.2059	0.1499	0.1322	0.0608	0.1918	0.1542	0.1269	0.0587	0.4175
		MSE	0.0626	0.0373	0.0265	0.0059	0.0554	0.0360	0.0258	0.0055	0.2782
		Moyenne	0.4747	0.2223	0.2951	0.1953	0.3258	0.3866	0.2062	0.2948	-0.4774
M4	500	Biais	-0.0253	0.0223	-0.0049	-0.0047	0.0258	-0.0134	0.0062	-0.0052	0.0226
		Écart type	0.2082	0.1429	0.1064	0.0472	0.2081	0.1476	0.1166	0.0448	0.3609
		MADE	0.1700	0.1182	0.0847	0.0379	0.1684	0.1216	0.0927	0.0356	0.2898
	MSE	MSE	0.0440	0.0209	0.0114	0.0023	0.0440	0.0220	0.0136	0.0020	0.1308
		Moyenne	0.4790	0.2112	0.3048	0.1965	0.3214	0.3900	0.1976	0.2983	-0.4779
		Biais	-0.0210	0.0112	0.0048	-0.0035	0.0214	-0.0100	-0.0024	-0.0017	0.0221
1000	Écart type	0.1682	0.1069	0.0807	0.0319	0.1718	0.1158	0.0921	0.0338	0.2946	
		MADE	0.1368	0.0851	0.0633	0.0254	0.1368	0.0928	0.0761	0.0268	0.2356
		MSE	0.0287	0.0115	0.0065	0.0010	0.0300	0.0135	0.0085	0.0011	0.0873
	Vraies Valeurs	0.5	0.2	0.3	0.2	0.3	0.4	0.2	0.3	-0.5	

TABLE III.8 – Résultats d'estimation MV pour le modèle M4 avec A non-diagonale et B diagonale.

On observe globalement une diminution du biais, de l'écart type, du MADE et du MSE lorsque la taille de l'échantillon augmente. Cela indique une meilleure précision et performance des modèles avec davantage de données.

Les figures suivantes résument les données des tableaux d'estimation MCO et MV, incluant les médianes, les quartiles et les valeurs extrêmes pour les quatre modèles (M1-M4) pour n=200, 500, 1000 respectivement.

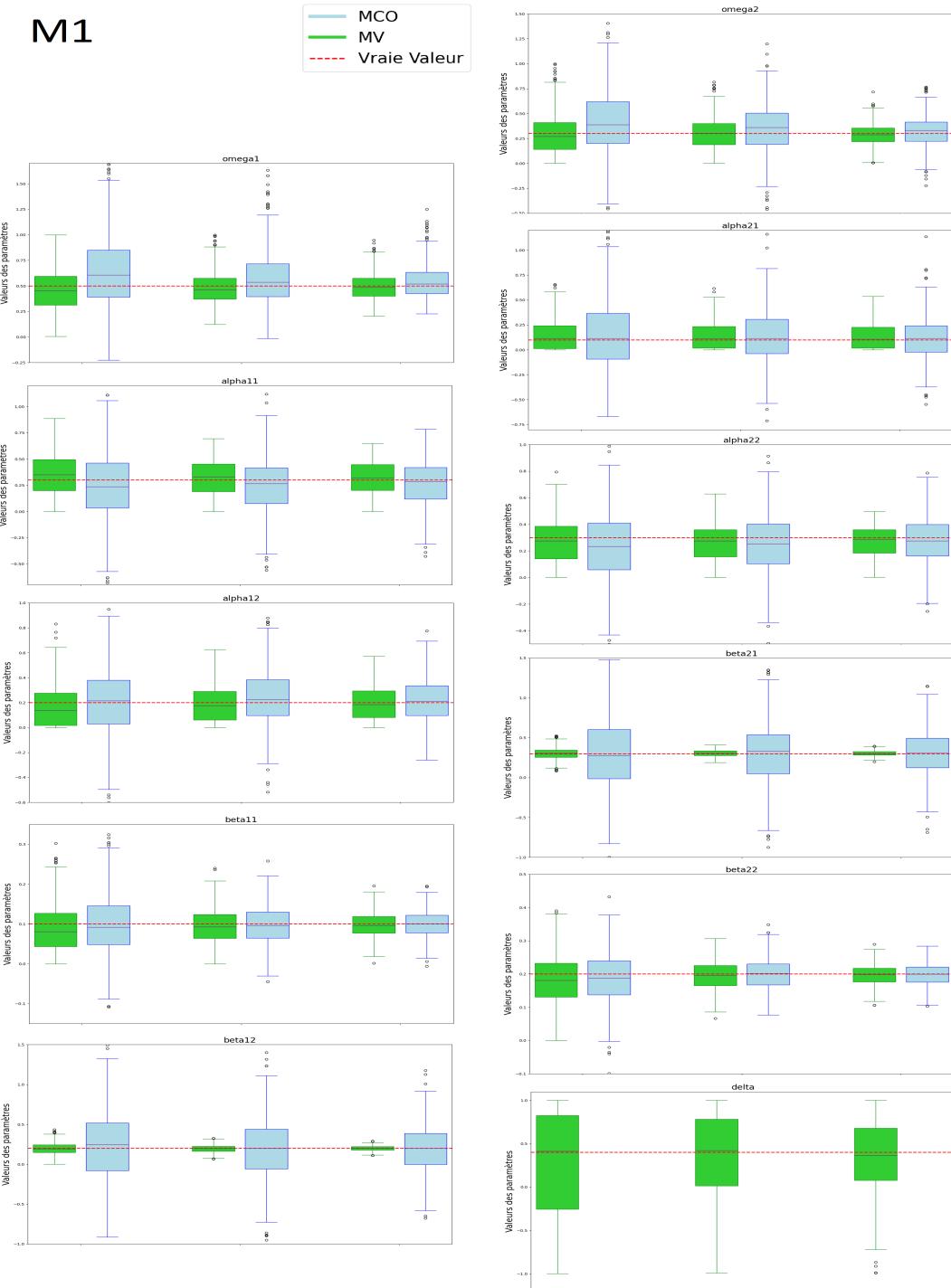


FIGURE III.5 – Boîte à moustaches M1.

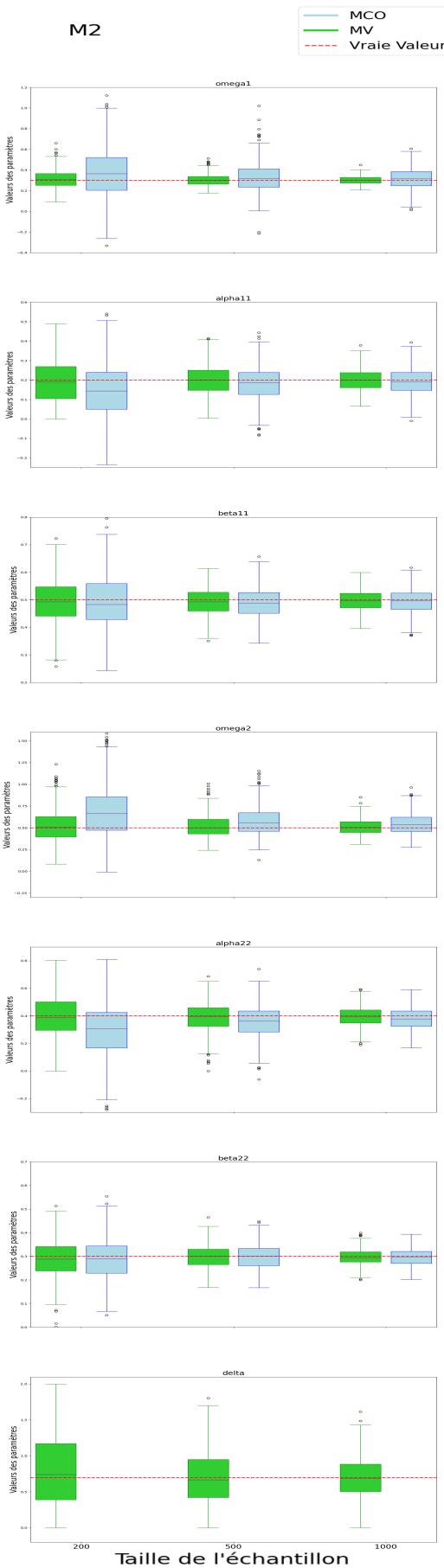


FIGURE III.6 – Boîte à moustaches M2.

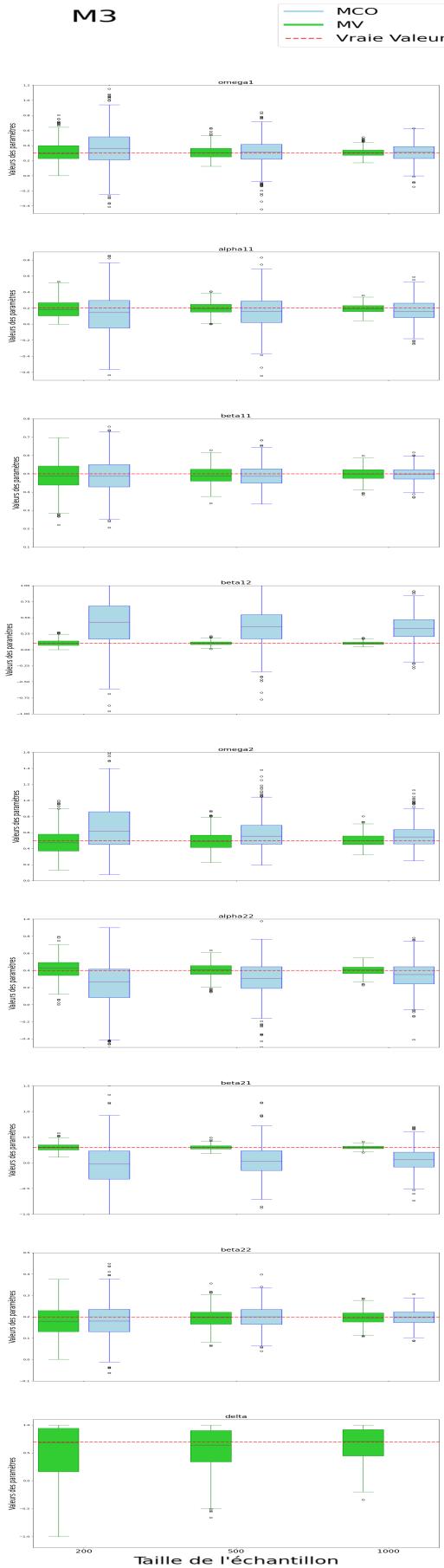


FIGURE III.7 – Boîte à moustaches M3.

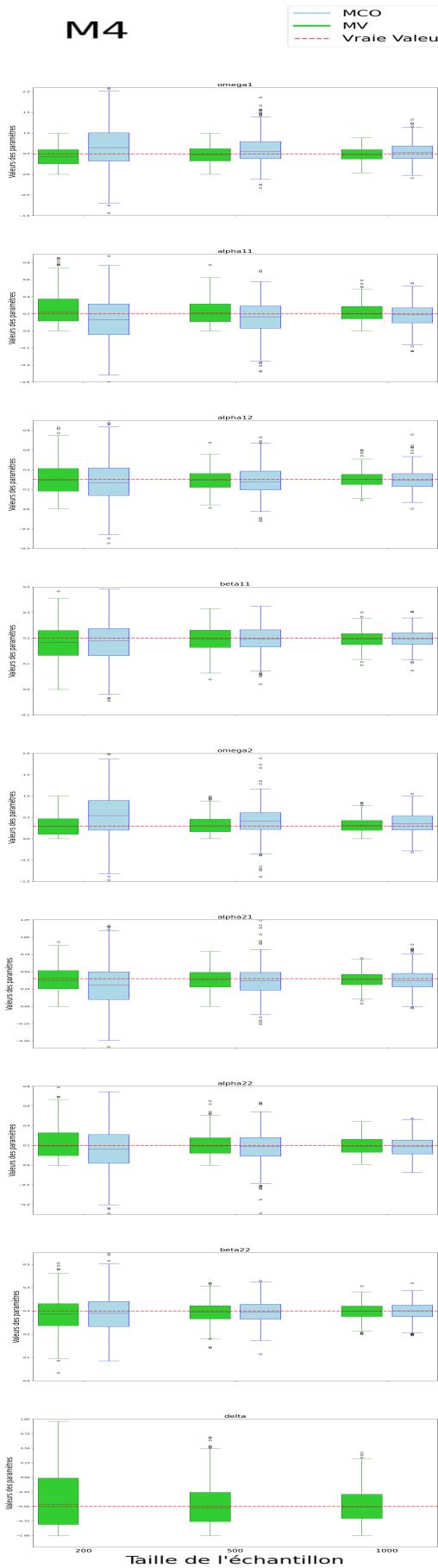


FIGURE III.8 – Boîte à moustaches M4.

à travers ces figures que les estimations gardent en général la vrai valeur dans l'intervalle du premier écart type, et ce dernier rétrécit lorsque la taille de l'échantillon augmente. On remarque aussi qu'avec cette augmentation, bien que pour certains cas les boîtes peuvent être décalées pour les tailles d'échantillon les plus faibles, elle deviennent centrées par rapport à la vraie valeur.

On constate que les estimateurs produits sont fiables (avec une nette amélioration lors de l'application de la méthode MV sur le modèle M3 par exemple), affichant généralement de faibles valeurs SD (écart-type), MADE et MSE. En augmentant la taille de l'échantillon, les estimateurs tendent à converger vers les vraies valeurs des paramètres.

Ces observations nous donnent une idée sur l'efficacité de l'estimation (celle-ci peut aussi être observée par l'erreur quadratique moyenne (MSE)), et le biais, remarquant enfin que ces indicateurs sont bons dans le cas qui nous importe le plus, celui de l'estimation par **MV**.

Pour comparer entre l'estimation par MCO et l'estimation par MV, le modèle M1 est choisi, dont la spécification est plus approprié pour la méthode par MCO car aucun paramètre n'est supposé nul à priori de l'estimation.

Les résultats d'estimation par MCO sont donnés par le tableau [Table III.1](#), et les résultats d'estimation par MV avec des valeurs initiales tirées aléatoirement (uniformément) pour chaque réplication sont donnés par [Table III.9](#) :

Modèle	n	Statistique	ω_1	α_{11}	α_{12}	β_{11}	β_{12}	ω_2	α_{21}	α_{22}	β_{21}	β_{22}	δ
200	Moyenne	0.4491	0.3656	0.2448	0.1274	0.2171	0.2912	0.2091	0.3329	0.2973	0.2038	0.2254	
	Biais	-0.0509	0.0656	0.0448	0.0274	0.0171	-0.0088	0.1091	0.0329	-0.0027	0.0038	-0.1746	
	Écart type	0.2492	0.2536	0.2754	0.1732	0.1779	0.2336	0.2571	0.2552	0.1690	0.1678	0.6449	
	MADÉ	0.2112	0.2129	0.2095	0.0933	0.1097	0.1890	0.1899	0.2019	0.1082	0.1081	0.5529	
	MSE	0.0647	0.0686	0.0778	0.0308	0.0319	0.0546	0.0780	0.0662	0.0286	0.0282	0.4464	
	Moyenne	0.4801	0.3725	0.2609	0.2355	0.3167	0.3586	0.2526	0.3364	0.3547	0.2774	0.2294	
M1	Biais	-0.0199	0.0725	0.0609	0.1355	0.1167	0.0586	0.1526	0.0364	0.0547	0.0774	-0.1706	
	Écart type	0.2357	0.2695	0.2706	0.2684	0.2478	0.2435	0.2994	0.2703	0.1801	0.2177	0.5495	
	MADÉ	0.1920	0.2231	0.2098	0.1736	0.1681	0.1924	0.2252	0.2064	0.1159	0.1349	0.4680	
	MSE	0.0560	0.0779	0.0769	0.0904	0.0751	0.0627	0.1130	0.0744	0.0354	0.0534	0.3310	
	Moyenne	0.4829	0.3668	0.2859	0.2450	0.2949	0.3776	0.2644	0.3316	0.3880	0.3088	0.2095	
	Biais	-0.0171	0.0668	0.0859	0.1450	0.0949	0.0776	0.1644	0.0316	0.0880	0.1088	-0.1905	
1000	Écart type	0.2172	0.2587	0.2787	0.2601	0.2266	0.2367	0.2936	0.2563	0.2140	0.2421	0.5343	
	MADÉ	0.1718	0.2085	0.2141	0.1691	0.1368	0.1791	0.2304	0.1911	0.1375	0.1505	0.4482	
	MSE	0.0475	0.0714	0.0851	0.0887	0.0603	0.0620	0.1132	0.0667	0.0535	0.0704	0.3218	
	Vraies Valeurs	0.5	0.3	0.2	0.1	0.2	0.3	0.1	0.3	0.3	0.2	0.4	

TABLE III.9 – Résultats d'estimation MV avec données initiales aléatoires pour le modèle M1.

En comparant ces résultats avec ceux obtenus par MCO, on remarque que la méthode MV offre généralement une meilleure précision d'estimation. Les biais et les écarts types sont généralement plus faibles avec la méthode MV.

Les Boîtes à moustaches pour comparer entre ces deux méthodes sont données par la figure suivante :

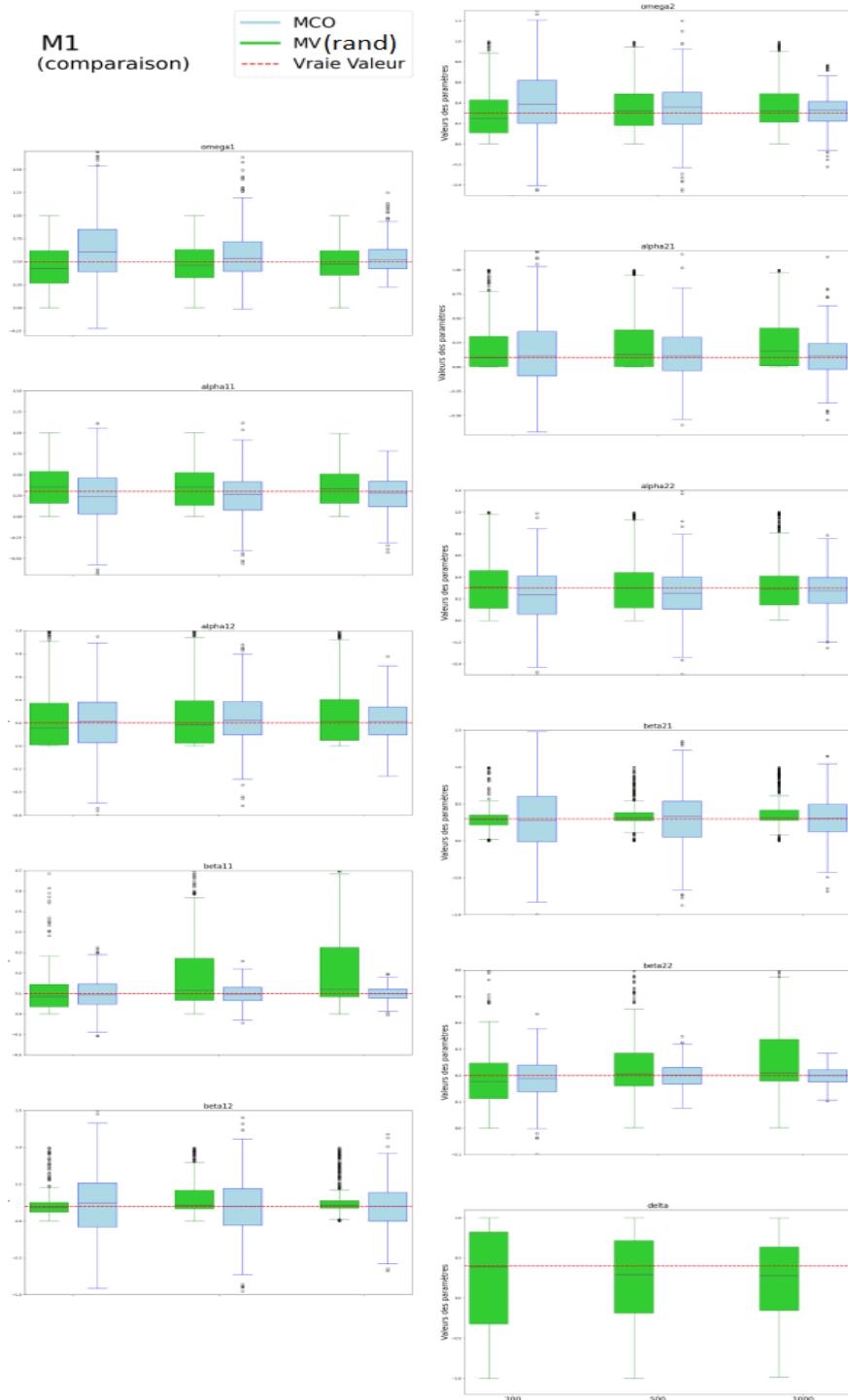


FIGURE III.9 – Boîte à moustaches M1.

En accord avec le tableau précédent, les *boîtes à moustaches* confirment que la précision des estimations par MCO (en bleu) et par MV avec valeurs initiales aléatoires (en vert) varie d'un paramètre à l'autre. Les boîtes bleues sont souvent plus centrées autour de la vraie valeur (ligne rouge pointillée) et plus étroites que les boîtes vertes.

La dispersion des estimations par MV avec valeurs initiales aléatoires est souvent plus importante, indiquant une plus grande variabilité et une convergence potentiellement plus difficile vers la vraie valeur. Ceci est visible par des boîtes vertes plus larges et des "moustaches" plus étendues.

On observe une amélioration de la précision des deux méthodes d'estimation avec l'augmentation de la taille de l'échantillon. Les boîtes deviennent plus étroites et plus centrées autour de la vraie valeur, notamment pour les échantillons de 1000 observations.

Pour comparer maintenant les résultats obtenus précédemment par la méthode MV utilisant des valeurs initiales aléatoires Table III.9 avec les ceux obtenus en combinant les deux méthodes MCO pour valeurs initiales et MV pour améliorer les résultat, on présente le graphe suivant :

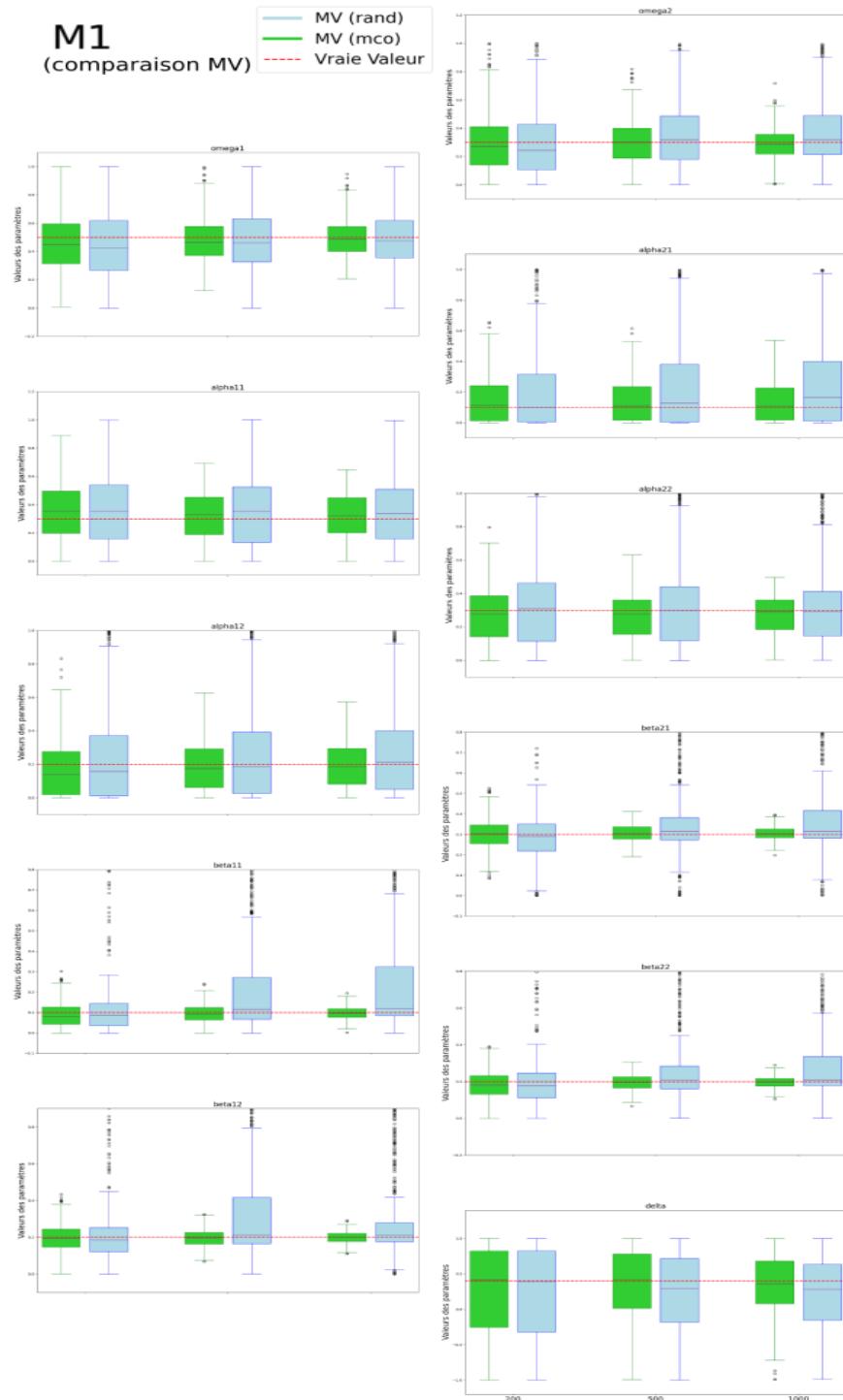


FIGURE III.10 – Boîte à moustaches M1.

Les boîtes à moustaches montrent clairement que l'utilisation des estimations MCO comme valeurs initiales pour la MV (boîtes vertes) améliore significativement la précision des estimations par rapport à l'utilisation de valeurs initiales aléatoires (boîtes bleues).

Les boîtes vertes sont généralement plus étroites et mieux centrées autour de la vraie valeur (ligne rouge pointillée) que les boîtes bleues pour la quasi-totalité des paramètres et des tailles d'échantillon. Cela indique un biais et une variance réduits pour l'approche MV (MCO).

L'approche MV (MCO) semble converger de manière plus stable et fiable vers les vraies valeurs des paramètres, même pour les petites tailles d'échantillon. Les boîtes bleues, en revanche, présentent souvent une plus grande variabilité et des valeurs aberrantes, suggérant une convergence plus difficile.

Ce graphique confirme l'importance cruciale du choix des valeurs initiales pour la performance de l'estimation par MV. L'utilisation des estimations MCO comme valeurs initiales pour la MV s'avère être une stratégie efficace pour améliorer la précision, la stabilité et la fiabilité des estimations pour le modèle M1.

Remarque III.1

Une autre approche est envisageable pour l'estimation par maximum de vraisemblance (MV) avec valeurs initiales aléatoires d'une trajectoire bivariée. Au lieu de moyenner les estimations des paramètres obtenues à partir de plusieurs exécutions avec des initialisations aléatoires différentes, cette méthode se concentre sur la recherche du meilleur jeu de valeurs initiales qui conduit à la vraisemblance la plus élevée.

Voici les étapes de cette approche :

1. Générer un vecteur de valeurs initiales dont la dimension est égale au nombre de paramètres à estimer.
2. Utiliser une routine d'optimisation numérique pour maximiser la vraisemblance (ou minimiser la fonction de -log-vraisemblance) et stocker la valeur maximale de la vraisemblance résultante.
3. Répéter les étapes 1 et 2 plusieurs fois avec différentes valeurs initiales aléatoires.
4. Identifier et conserver le jeu de valeurs initiales ayant produit la plus grande vraisemblance après optimisation.
5. Utiliser ces valeurs initiales conservées pour obtenir l'estimation finale des paramètres, au lieu de calculer une moyenne sur toutes les exécutions.

Cette approche met l'accent sur la recherche du maximum global de la vraisemblance en explorant l'espace des paramètres de manière plus efficace grâce à de multiples initialisations aléatoires. Cette approche serait d'autant plus utile lors de l'estimation sur données réelles où une seule réalisation de la trajectoire est disponible.

III.2 Application sur données réelles

Cette partie d'application concerne l'analyse conjointe des données **journalières** représentant le nombre d'accidents de la route **diurnes** et **nocturnes** dans la région de *Schiphol* aux Pays-Bas durant l'année 2001.

Les accidents nocturnes sont définis comme ceux survenant entre 22h00 et 06h00, les autres étant considérés comme des accidents diurnes.

Dans l'analyse des accidents ([PEDELI et KARLIS, 2011](#)), ces types d'accidents sont considérés comme ayant des comportements différents. Pendant la nuit, la nature de la circulation change (par exemple, davantage de personnes voyagent pour des activités de divertissement). Cependant, comme les deux types d'accidents partagent le même environnement, comme les conditions météorologiques ou les caractéristiques de la route, ils sont corrélés.

L'utilisation intuitive des séries temporelles pour l'analyse découle du fait que les données provenant de jours consécutifs sont typiquement corrélées car elles reflètent des conditions similaires. [PEDELI et KARLIS \(2011\)](#) indique également que les autocorrélations sont généralement faibles et d'ordre AR(1). Ignorer l'aspect temporel conduit à des inférences erronées ([BRIJS et al., 2008](#)). Quant à la modélisation bivariée, elle s'explique par le fait que ces mêmes conditions sont communes aux deux types d'accidents.

Les accidents diurnes et nocturnes ont des moyennes respectives de 7.29, 1.50 et des variances de 20.87, 1.86 respectivement, ce qui indique une *surdispersion*. La figure [III.11](#) illustre les données étudiées où on remarque à premier abord qu'une des séries (accidents diurnes) est clairement au dessus de l'autre, mais un comportement général assez similaire.

Ce jeu de donné à été utilisé par plusieurs auteurs dont : [BENTARZI et SADOUN, 2021](#); [KARLIS et al., 2008](#); [H. LIU, 2012](#); [MAGESA et al., 2023](#); [PEDELI et KARLIS, 2011](#).

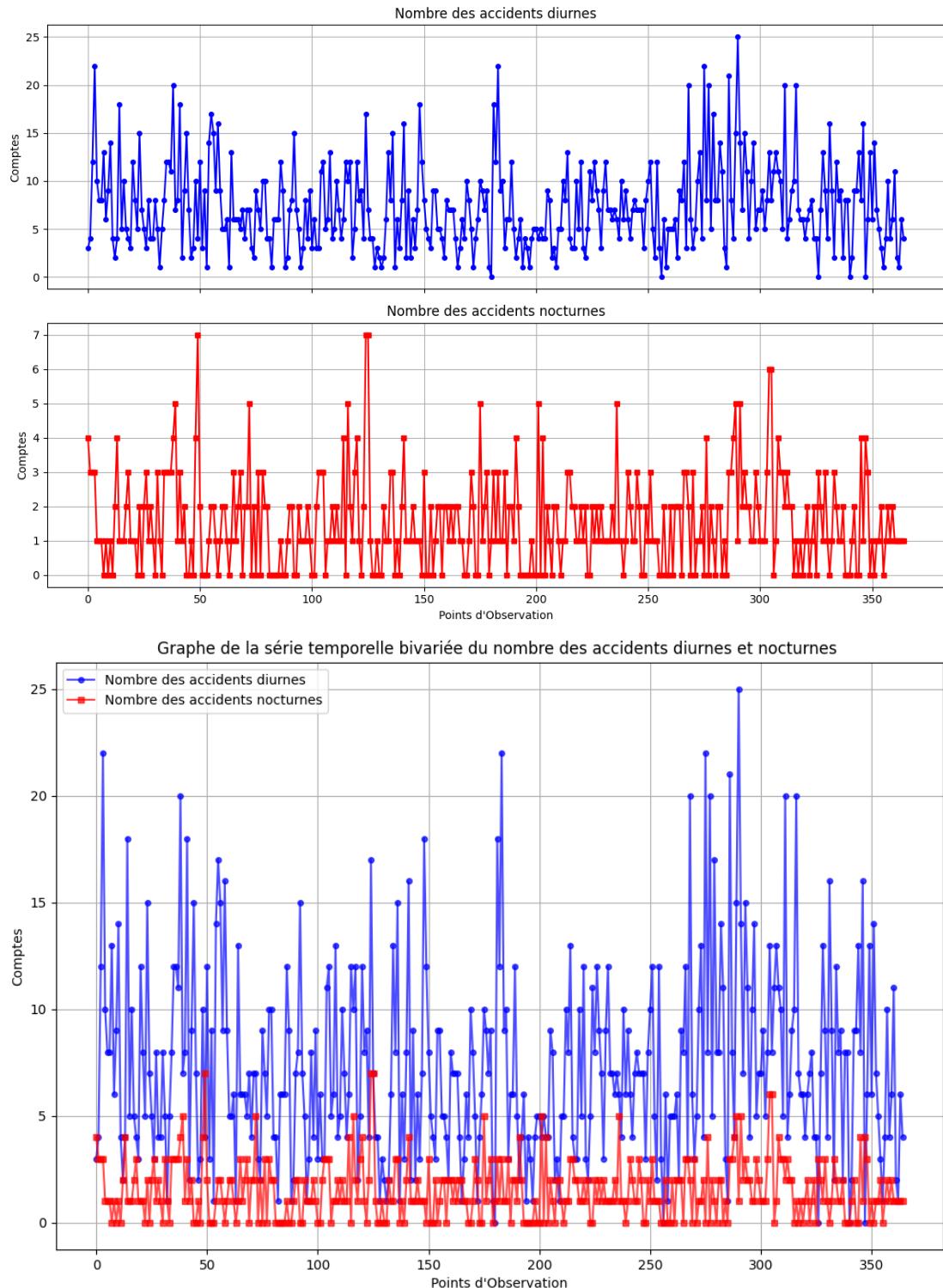


FIGURE III.11 – Graphes de la série temporelle bivariée des accidents diurnes et nocturnes.

Nous présentons ensuite les graphes de *Corrélation croisée* (CCF) représentant la corrélation entre les deux séries pour plusieurs retards, ainsi que les fonction d'autocorrélation et autocorrélation partielle (ACF, PACF) pour chacune des deux séries univariées.

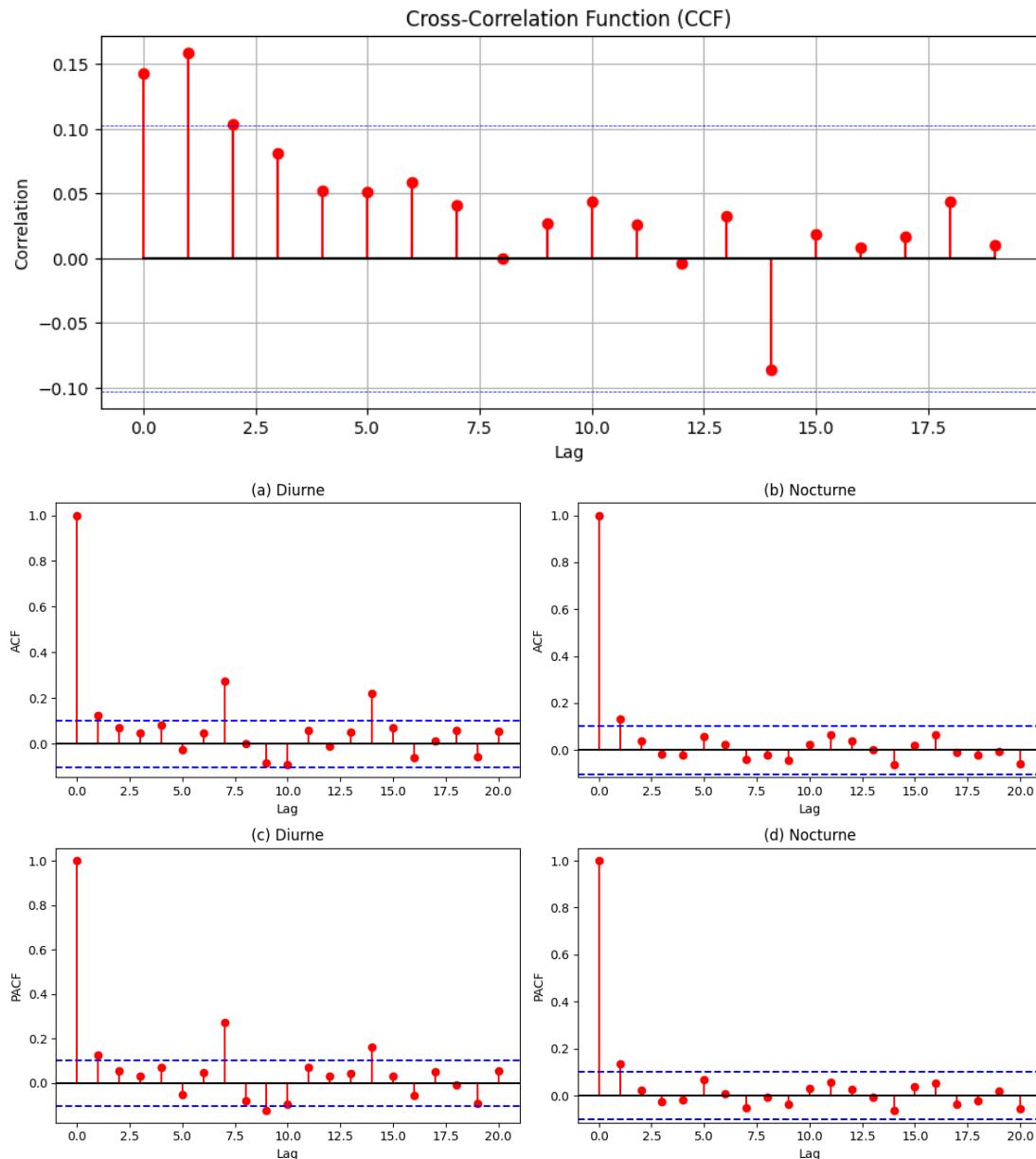


FIGURE III.12 – Les corrélations croisées, autocorrélations, et autocorrélations partielles

On remarque ici que pour la corrélation croisée les deux premiers retards sont les plus significatifs, tandis qu'en ce qui concerne les fonctions ACF et PACF, la série diurne a présente des piques significatifs aux retards 7 et 14, la série nocturne n'a une valeur ACF et PACF significative qu'au retard 1.

La forme et les piques des ACF et PACF nous indiquent aussi une **saisonalité hebdomadaire**.

Les propriétés de cette série et son historique d'utilisation par d'autres auteurs dans leurs travaux fait d'elle un bon candidat pour l'étude de notre modèle et son application sur des données pratiques.

Le tableau de contingence **Table III.2** montre la distribution conjointe des deux séries Diurne et Nocturne (Y_1, Y_2). On observe une concentration des observations sur et au dessus de la diagonale principale du tableau, et ce car le nombre d'accidents de type diurne est supérieur au nombre d'accident nocturne bien qu'il soient positivement corrélés et manifestent le même comportement.

$Y_1 \backslash Y_2$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	20	21	22	25
0	2	13	2	8	10	13	10	7	5	7	2	3	2	3	1	2	2	0	0	1	0	1	0
1	1	3	9	7	10	8	10	11	13	12	9	1	6	6	3	0	1	1	2	1	0	0	1
2	0	3	7	7	12	9	9	8	6	6	3	1	5	0	1	2	0	1	0	2	0	0	0
3	1	0	1	5	4	6	3	2	5	2	2	2	6	1	0	2	0	0	2	0	1	2	0
4	1	0	1	1	3	0	1	0	2	0	1	1	1	0	0	0	1	0	0	1	0	0	0
5	0	0	0	0	1	0	1	2	0	0	2	0	0	0	1	1	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
7	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0

TABLE III.10 – Tableau de contingence pour les données réelles avec Y_1 Diurne et Y_2 Nocturne

Modèle 1		
Paramètres	MCO	MV
ω_1	2.754	1.675
α_{11}	0.471	0.488
α_{12}	-0.091	0.637
β_{11}	0.078	0.062
β_{12}	-0.0006	0.419
ω_2	1.09	0.621
α_{21}	0.13	0.004
α_{22}	0.611	0.345
β_{21}	0.458	0.033
β_{22}	0.095	0.065
δ		1.656

Modèle 2		
Paramètres	MCO	MV
ω_1	2.710	3.429
α_{11}	0.528	0.403
β_{11}	0.093	0.126
ω_2	0.830	1.368
α_{22}	0.324	2.95×10^{-5}
β_{22}	0.111	0.107
δ		1.733

	Modèle 1	Modèle 2
Log-vraisemblance	-1724.4137	-1744.5682
AIC	3470.8275	3503.1364
BIC	3513.7264	3530.4357

Les valeurs des AIC et BIC pour les deux modèles sont calculés à partir des formules suivantes :

$$AIC = 2k - 2 \log(\mathcal{L})$$

$$BIC = k \log(n) - 2 \log(\mathcal{L})$$

Avec k est le nombre des paramètres du modèle, \mathcal{L} est la fonction de vraisemblance et n le nombre des observations. Ces valeurs nous indiquent que le modèle le plus approprié à choisir est le modèle 1.

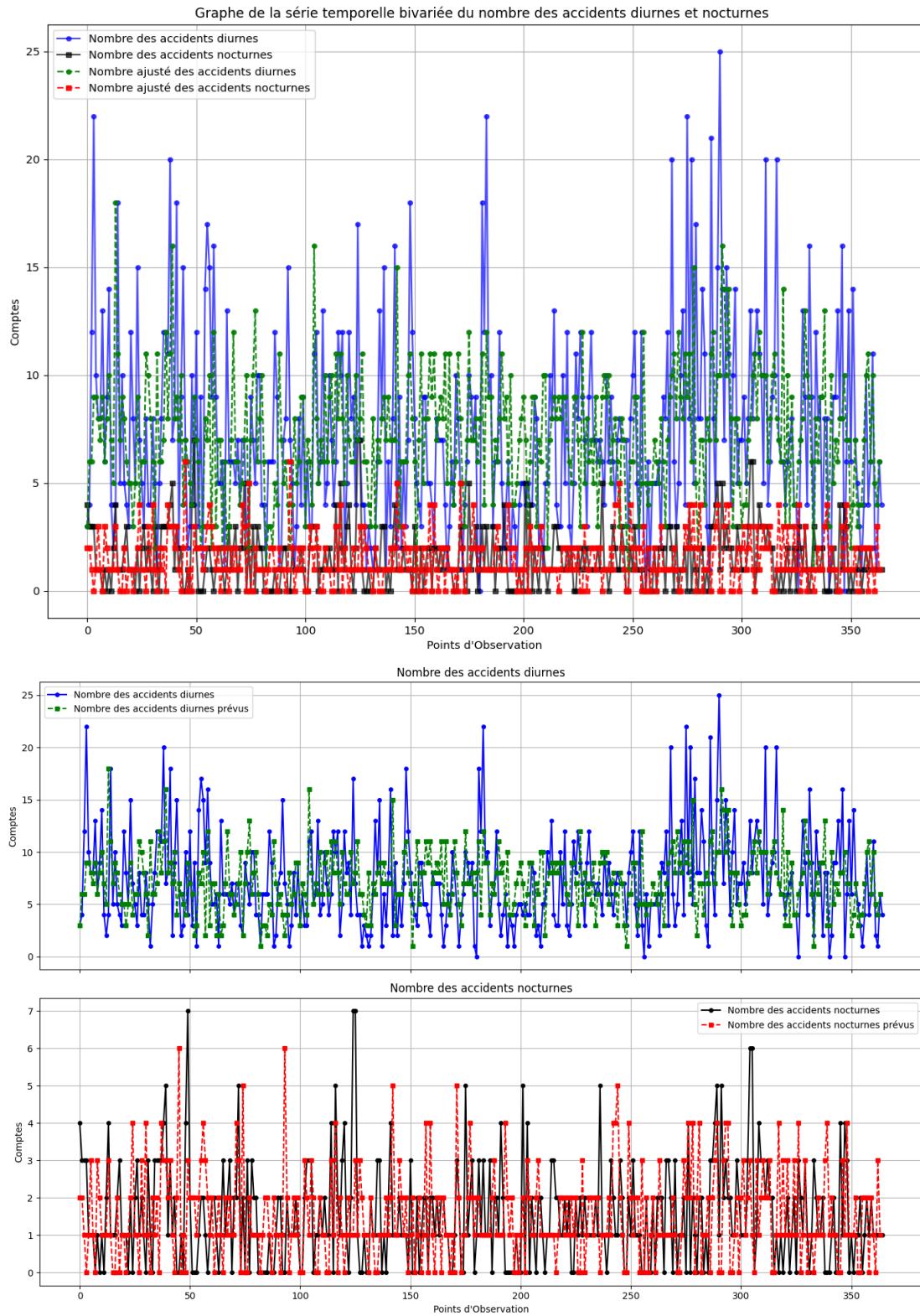


FIGURE III.13 – Graphes des prédictions (modèle 1)

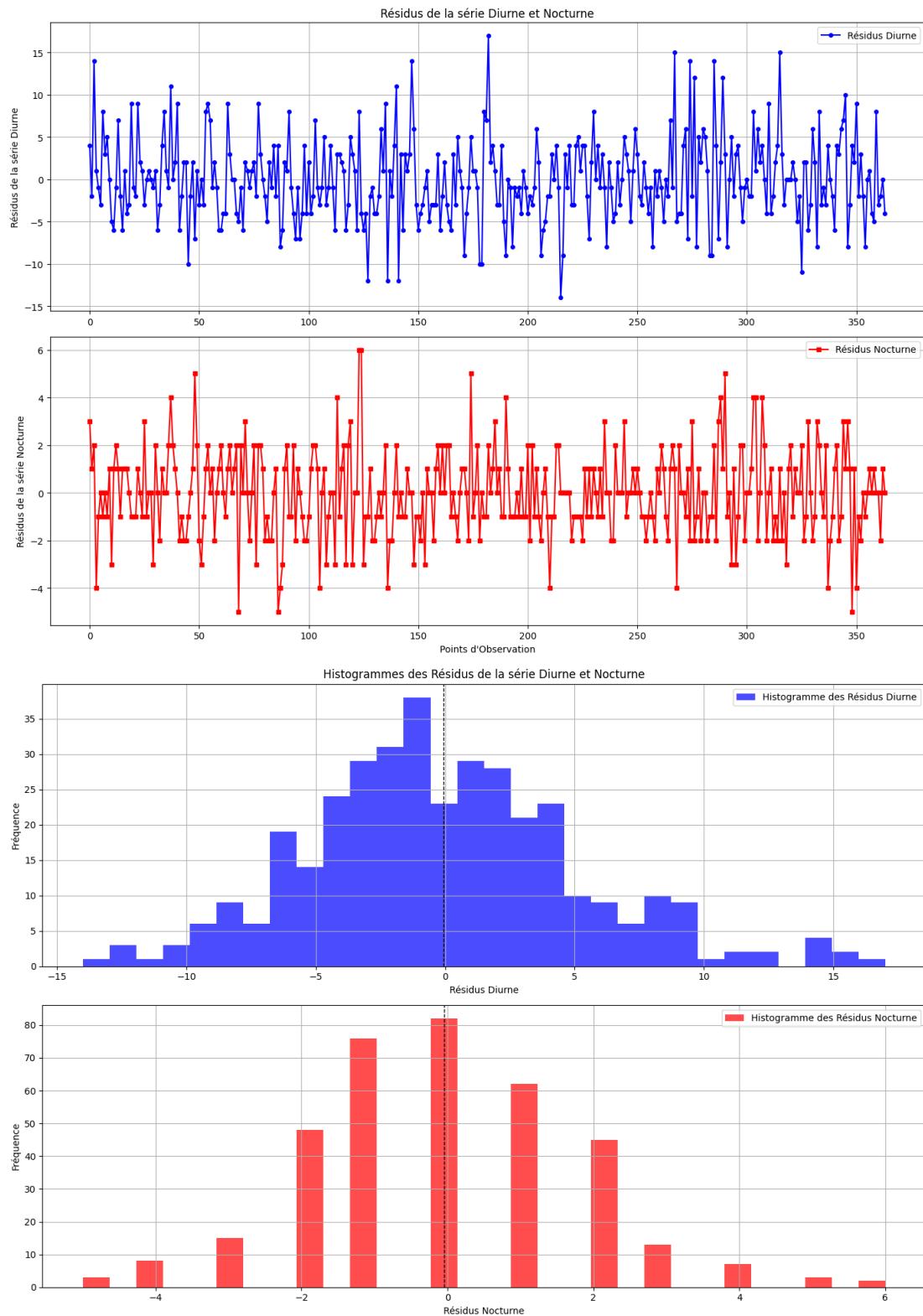


FIGURE III.14 – Graphes des résidus pour chaque composante de la série bivariée, et leur distribution(modèle 1)

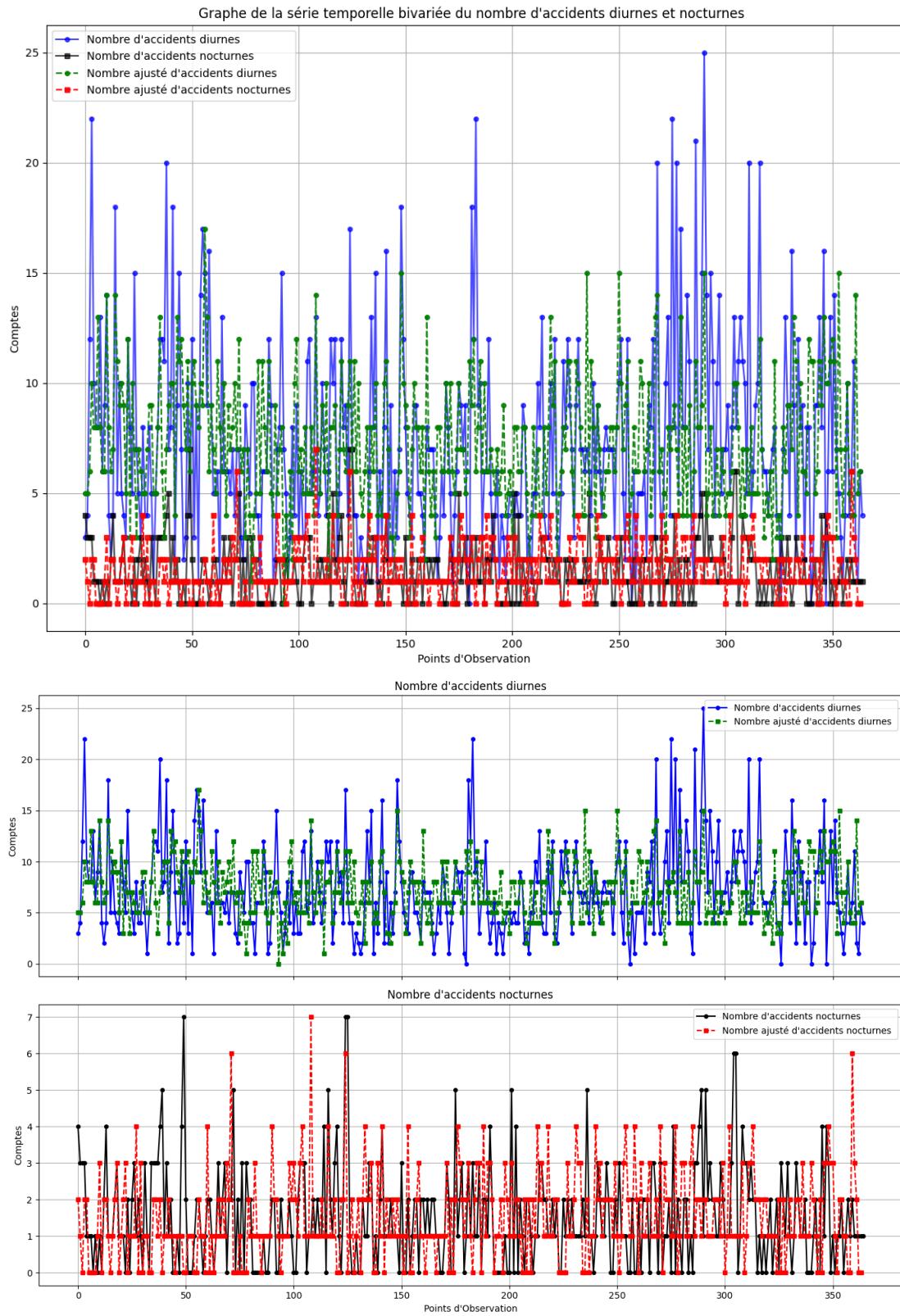


FIGURE III.15 – Graphes des prédictions(modèle 2)

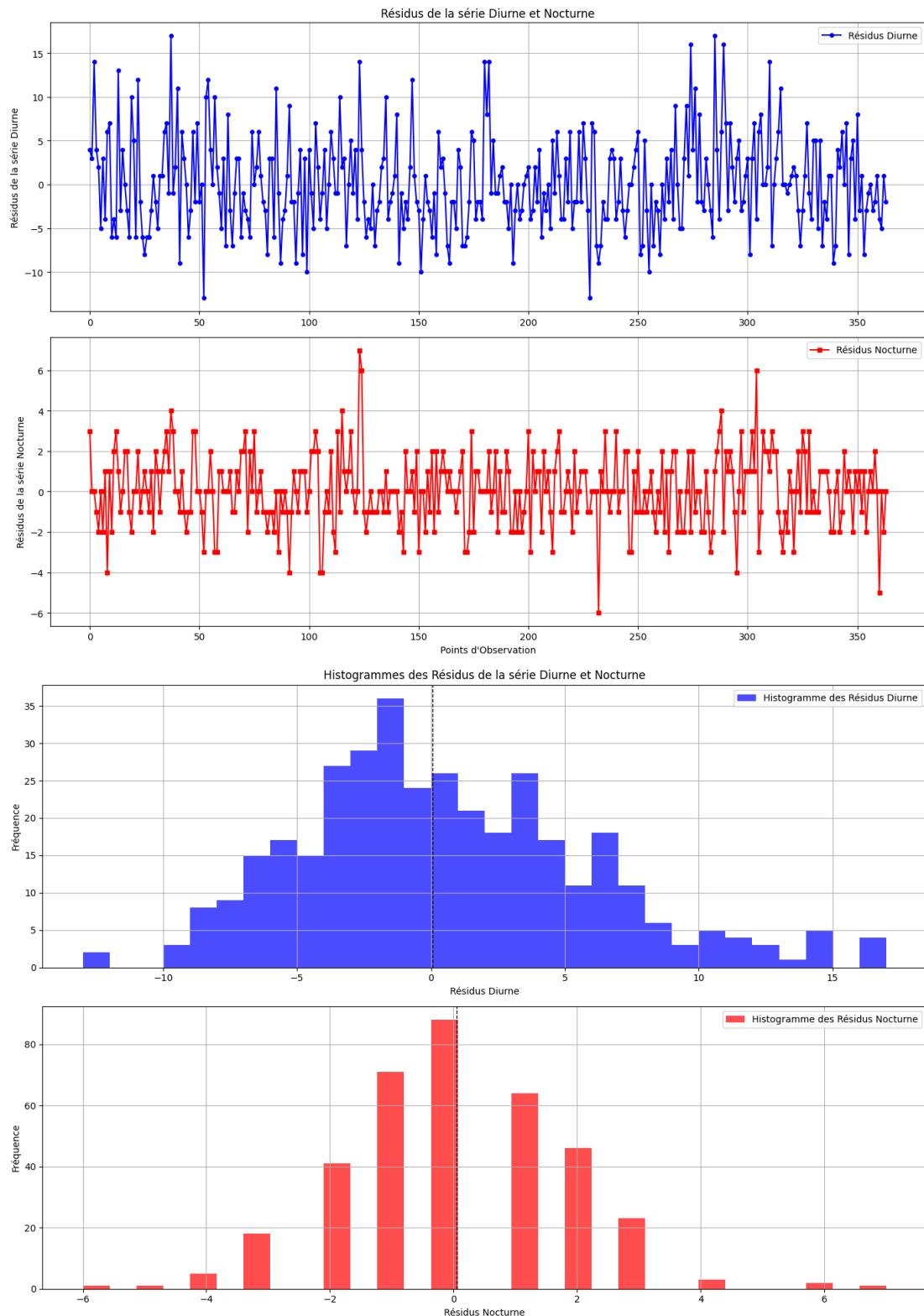


FIGURE III.16 – Graphes des résidus pour chaque composante de la série bivariée, et leur distribution(modèle 2)

Nous remarquons à travers les graphes superposés des séries réelles et séries ajustées

(prédition à 1 pas futur) que le modèle capture de façon acceptable les dynamique de nos données.

Nous remarquons aussi que pour ce qui est des deux modèles, les bruits sont bien centrés et que à partir de leur graphes montrent un comportement de bruit. De plus lorsqu'on observe leur distribution empirique, la série d'accident nocturne ressemble bien à une distribution normale.

Les fonctions *ACF* et *PACF* sont représentées par les graphes suivants :

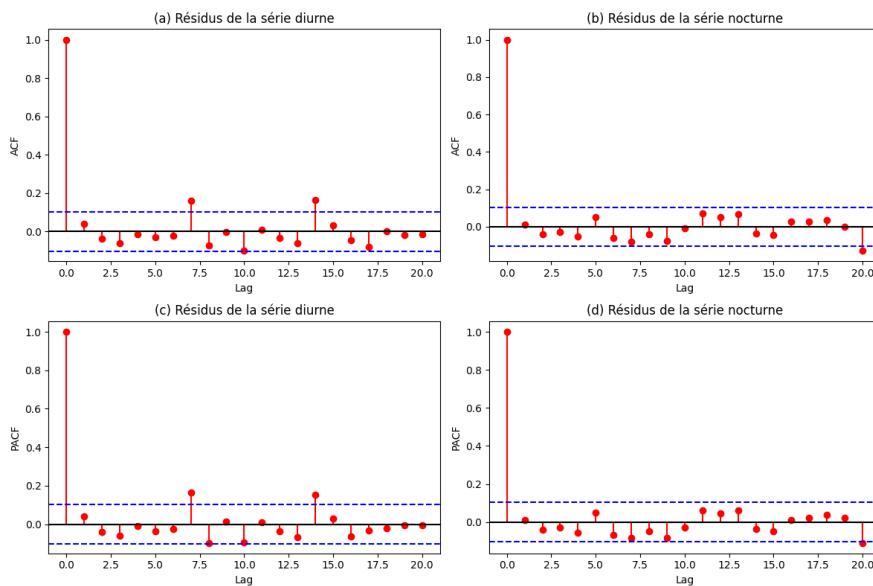


FIGURE III.17 – fonctions ACF et PACF pour le modèle 1

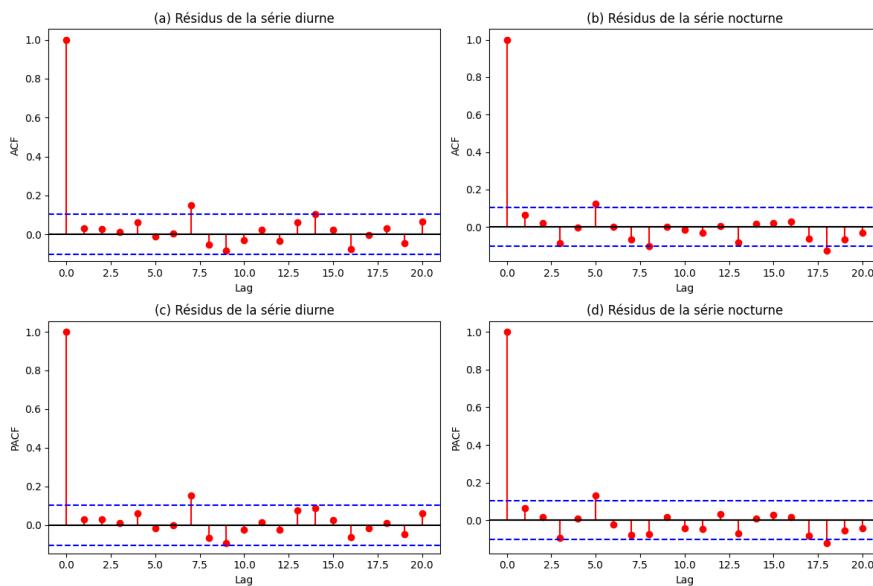


FIGURE III.18 – fonctions ACF et PACF pour le modèle 2

Ci dessous un tableau détaillant les erreurs des bruits pour les deux modèles et deux composantes de chacun.

	Modèle 1		Modèle 2	
	Résidus Diurnes	Résidus Nocturnes	Résidus Diurnes	Résidus Nocturnes
MADE	3.915	1.2958	3.7945	1.3342
MSE	25.5369	3.0273	23.9534	3.1808

TABLE III.11 – Résultats des MSE/MADE des résidus du Modèle 1 et Modèle 2 des deux séries résiduelles.

IV

CONCLUSION, DISCUSSION ET PERSPECTIVES

Ce travail explore plusieurs aspects de la modélisation des séries chronologiques de comptage dans le cadre multivarié en se focalisant surtout sur le cas bivarié.

La revue de littérature a permis d'avoir une idée globale sur l'état de l'art et les différentes approches employées dans l'analyse des données de comptage multivariées à l'aide d'une formulation INGARCH, situant ainsi le contexte de notre travail.

Ce mémoire de master traite du modèle bivarié BP-INGARCH. La formulation de ce modèle a été détaillée, incluant la distribution utilisée et sa construction, ainsi que les propriétés liées à la stationnarité et à l'ergodicité. De plus, les méthodes d'estimation existantes ont été décrites et appliquées.

Parmi les contributions de ce travail est une nouvelle approche pour l'estimation, utilisant la méthode des Moindres Carrés Ordinaires bloc par bloc, qui offre un estimateur équivalent à celui de la méthode des Moindres Carrés Généralisée.

L'étude de simulation a permis d'évaluer les résultats de cette nouvelle approche, démontrant sa performance, mesurée par l'erreur moyenne de déviation absolue (MADE) et l'erreur quadratique moyenne (MSE). Les simulations ont également montré l'utilité de ces résultats comme valeurs initiales lors de l'emploi des routines d'optimisation pour la méthode du Maximum de Vraisemblance. Cette dernière a aussi été traitée en simulation à travers des routines tel que `fmincon` sur Matlab (et aussi faite sur Python, donné en annexe [Section A.8](#)) et utilisée lors de l'application sur données réelles.

L'application sur des données réelles a permis de valider l'utilité pratique du modèle BP-INGARCH. Les résultats obtenus à travers une méthodologie classique ont montré que le modèle bivarié est capable de capturer efficacement les dynamiques complexes présentes dans les données, fournissant ainsi des prévisions et des résidus acceptables, comme illustré par les graphiques.

Néanmoins, on peut noter quelques limitations. La méthode MCO/MCG est adéquate à des spécifications de modèles où aucun paramètre n'est supposé nul à priori, bien qu'une adaptation peut être faite pour le cas où les deux matrices de coefficients sont

nulles, le cas où seulement l'une d'entre elles l'est n'est pas bien adapté.

Le modèle **BP-INGARCH** est limité au cas bivarié, une extension à partir de celui-ci vers le trivarié n'est pas possible naturellement car cela demande la construction d'une densité (au sens large) jointe avec une nouvelle méthode autre que le facteur multiplicatif, et les distributions de dimensions supérieures compliquent la manipulation du modèle et ses propriétés statistiques.

La majorité du travail a aussi été faite sur un ordre(1,1)

On remarque aussi que, bien que nous utilisons la méthode du Maximum de Vraisemblance, celle-ci manque d'une donnée initiale que nous devons choisir au préalable afin de pouvoir lancer une routine d'optimisation visant à maximiser la fonction de log-Vraisemblance.

Malgré l'existence de ces limitations, ce travail fournit aussi une opportunité de développer plusieurs extensions :

- On voudrait étendre les résultats à des ordres de retards supérieurs : $BP\text{-}INGARCH}(p,q)$ avec $p, q > 1$
- Il serait possible de sortir du cadre de la distribution Poisson *BP* particulière utilisée pour le modèle, soit pour arriver à une Poisson trivariée par exemple, ou généraliser la construction par *facteur multiplicatif* comme fait par **Cui et al. (2020)**.
- Changement de loi : Une nouvelle loi autre que Poisson pourrait être utilisée tel que la loi Binomiale Négative Bivariée (voir **KOCHERLAKOTA et KOCHERLAKOTA (2017)**) pour arriver à un modèle que l'on pourrait noter $NB\text{-}INGARCH}(p,q)$.
- Extension du modèle vers un modèle à changement de régime tel qu'un modèle à seuil (*Threshold*), un modèle avec des sauts (points de rupture) (*Change-point*), mélanges, et changement de régime Markovien

Nous allons nous focaliser dans un avenir proche sur deux des extensions citées précédemment, à savoir : le modèle BP-INGARCH à seuil qu'on note **SET-BP-INGARCH**, et le modèle BP-INGARCH ségmenté (modèle avec des points de rupture) qu'on peut noter **ChP-BP-INGARCH**

IV.1 Modèle Self excited Threshold BP-INGARCH

Le modèle **Self excited Threshold BP-INGARCH** peut être représenté somme suit pour une specification SET-BP-INGARCH(K, 1, 1) :

$$Z_t = \sum_{\kappa=1}^K Y_{\kappa,t} \mathbb{I}_{S(t)=\kappa}. \quad (\text{IV.1})$$

avec : $\mathbb{I}_{S(t)=\kappa}$ la fonction indicatrice ($\mathbb{I}_{S(t)=\kappa} = 1$ pour $S(t) = \kappa$ et nulle autrement),

S_t suite de variables aléatoires iid mais pas indépendante de $Z_{t-i}, i \geq 1$ qui indique le régime où on se trouve.

Z_t contient toutes les observations, séparées en partitions $Y_{\kappa,t}$

On définit alors :

$$\begin{aligned} Y_{\kappa,t} | \mathcal{F}_{t-1} &\sim BP(\lambda_{\kappa t,1}, \lambda_{\kappa t,2}, \delta_\kappa), \\ \lambda_{\kappa t} &= (\lambda_{\kappa t,1}, \lambda_{\kappa t,2})^T = \omega_\kappa + A_\kappa \lambda_{\kappa t-1} + B_\kappa Y_{\kappa t-1}. \end{aligned} \quad (\text{IV.2})$$

Avec $\kappa = \{1, 2, \dots, K\}$, K le nombre de régimes,

$$\omega_\kappa = \begin{pmatrix} \omega_{\kappa,1} \\ \omega_{\kappa,2} \end{pmatrix} \text{ et } A_\kappa = \begin{pmatrix} \alpha_{\kappa,11} & \alpha_{\kappa,12} \\ \alpha_{\kappa,21} & \alpha_{\kappa,22} \end{pmatrix}, B_\kappa = \begin{pmatrix} \beta_{\kappa,11} & \beta_{\kappa,12} \\ \beta_{\kappa,21} & \beta_{\kappa,22} \end{pmatrix}.$$

$$S(t) = \begin{cases} 1 & \text{si } \|Z_{t-1}\|_\infty \leq c_1 \\ 2 & \text{si } c_1 \leq \|Z_{t-1}\|_\infty \leq c_2 \\ \vdots & \\ K-1 & \text{si } c_{K-2} \leq \|Z_{t-1}\|_\infty \leq c_{K-1} \\ K & \text{si } \|Z_{t-1}\|_\infty > c_{K-1} \end{cases} \quad (\text{IV.3})$$

où les c_i avec $i \in \{1, \dots, K-1\}$ définissent les seuils pour lesquels le modèle change de régime,

$$Z_t = \begin{pmatrix} Z_{t,1} \\ Z_{t,2} \end{pmatrix},$$

et $\|Z_{t-1}\|_\infty = \max\{Z_{t-1,1}, Z_{t-1,2}\}$.

Nous allons plus tard traiter le cas particulier où $K = 2$ c'est à dire le modèle **SET-BPINGARCH(2, 1, 1)** avec :

$$S(t) = \begin{cases} 1 & \text{si } \|Z_{t-1}\|_\infty \leq c_1 \\ 2 & \text{si } \|Z_{t-1}\|_\infty > c_1 \end{cases} \quad (\text{IV.4})$$

Remarque IV.1 Nous pouvons choisir une autre norme vectorielle, tel que $\|\cdot\|_1$, où choisir l'une des composantes de Z_{t-1} pour dicter le changement de régime.

IV.2 Modèle BP-INGARCH Segmenté (Change Point BP-INGARCH)

Le modèle segmenté avec K sauts noté ChP-BP-INGARCH($K, 1, 1$) peut être défini par :

$$Z_t = \sum_{\kappa=1}^K Y_{\kappa,t} \mathbf{I}_{S(t)=\kappa} \quad (\text{IV.5})$$

$$\begin{aligned} Y_{\kappa,t} | \mathcal{F}_{t-1} &\sim BP(\lambda_{\kappa t,1}, \lambda_{\kappa t,2}, \delta_\kappa), \\ \lambda_{\kappa t} &= (\lambda_{\kappa t,1}, \lambda_{\kappa t,2})^T = \omega_\kappa + A_\kappa \lambda_{\kappa t-1} + B_\kappa Y_{\kappa t-1}. \end{aligned} \quad (\text{IV.6})$$

Avec $\kappa = \{1, 2, \dots, K\}$, $\omega_\kappa = \begin{pmatrix} \omega_{\kappa,1} \\ \omega_{\kappa,2} \end{pmatrix}$,

$$\text{et } A_\kappa = \begin{pmatrix} \alpha_{\kappa,11} & \alpha_{\kappa,12} \\ \alpha_{\kappa,21} & \alpha_{\kappa,22} \end{pmatrix}, B_\kappa = \begin{pmatrix} \beta_{\kappa,11} & \beta_{\kappa,12} \\ \beta_{\kappa,21} & \beta_{\kappa,22} \end{pmatrix}.$$

$$S(t) = \begin{cases} 1 & \text{si } t \leq \tau_1 \\ 2 & \text{si } \tau_1 \leq t \leq \tau_2 \\ \vdots & \\ K-1 & \text{si } \tau_{K-2} \leq t \leq \tau_{K-1} \\ K & \text{si } t > \tau_{K-1} \end{cases} \quad (\text{IV.7})$$

avec τ_i sont les points de rupture.

Pour le cas où on a un seul point de rupture ($K = 2$), ChPBP-INGARCH(2, 1, 1). On a :

$$S(t) = \begin{cases} 1 & \text{si } t \leq \tau^* \\ 2 & \text{si } t > \tau^* \end{cases} \quad (\text{IV.8})$$

A

ANNEXE

A.1 Preuve du Théorème II.1

Tout d'abord, notons que $\{\lambda_t\}$ possède au moins une distribution stationnaire (H. Liu, 2012 pour plus de détails). D'après (II.2), il est clair que $(I - A)^{-1}\omega$ est un état atteignable si $Y_{t-1} = Y_{t-2} = \dots = 0$ pour un certain $t \in \mathbb{N}$ suffisamment grand.

De plus, $\{\lambda_t, t \geq 1\}$ peut être représenté comme une fonction aléatoire itérée en suivant la notation utilisée par Wu et Shao (2004).

Pour faciliter l'investigation, la fonction aléatoire $f_u(\lambda)$ selon la loi de probabilité (II.1) est définie comme $f_u(\lambda) = \omega + A\lambda + BF_\lambda^{-1}(u)$, où $u = (u_1, u_2) \in [0, 1]^2$, $\lambda = (\lambda_1, \lambda_2)$, $F_\lambda^{-1}(u) = (F_{\lambda_1}^{-1}(u_1), F_{\lambda_2}^{-1}(u_2)) \in \mathbb{N}_0^2$, et $F^{-1}(u) = \inf\{t \geq 0 : F(t) \geq u\}$. Par conséquent, on peut voir que pour tout t , $\lambda_t = f_{U_t}(\lambda_{t-1})$, où $\{U_t, t \geq 1\}$ suit une distribution uniforme indépendante sur $[0, 1]^2$.

Pour établir l'unicité de cette distribution stationnaire, nous allons montrer que $\{\lambda_t\}$ est une e-chaîne. Rappelons qu'une chaîne de Markov est une e-chaîne si, pour toute fonction continue f à support compact définie sur $[0, \infty) \times [0, \infty)$ et pour tout $\epsilon > 0$, il existe un $\eta > 0$ tel que $|P_{x_1}^k f - P_{z_1}^k f| < \epsilon$, pour $\|x_1 - z_1\|_p < \eta$ et $\forall k \geq 1$ on a $x_1 = (x_{1,1}, x_{1,2})^T$ et $z_1 = (z_{1,1}, z_{1,2})^T$. $P_{x_1}^k f$ désigne l'espérance conditionnelle de $f(\lambda_k)$ sachant que $\lambda_0 = x_1$.

Sans perte de généralité, supposons que $|f| \leq 1$. Choisissons ϵ' et η suffisamment petits pour que $\epsilon' + 16\eta/(1 - \|A\|_p) < \epsilon$ et $|f(x_1) - f(z_1)| < \epsilon'$ dès que $\|x_1 - z_1\|_p < \eta$, pour un certain $p \in [1, \infty]$.

Pour $k = 1$, nous avons :

$$\begin{aligned}
 |P_{x_1}f - P_{z_1}f| &= \left| \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} [f(\omega + Ax_1 + B(m, n))p(m, n|x_1) - f(\omega + Az_1 + B(m, n))p(m, n|z_1)] \right| \\
 &\leq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p(m, n|x_1) |f(\omega + Ax_1 + B(m, n)) - f(\omega + Az_1 + B(m, n))| \\
 &\quad + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |p(m, n|x_1) - p(m, n|z_1)| |f(\omega + Az_1 + B(m, n))| \\
 &= I_1 + I_2,
 \end{aligned}$$

Où $p(m, n|x_1)$ et $p(m, n|z_1)$ sont les fonctions de masse des lois bivariées de Poisson $BP(x_{1,1}, x_{1,2}, \delta)$ et $BP(z_{1,1}, z_{1,2}, \delta)$ respectivement, données par l'équation (II.1).

En utilisant l'inégalité triangulaire et l'hypothèse $|f| \leq 1$, nous pouvons borner I_2 :

$$\begin{aligned}
 \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |p(m, n|x_1) - p(m, n|z_1)| &\leq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |\phi_{1\vee 2}| \left| \frac{x_{1,1}^m x_{1,2}^n}{m! n!} e^{-(x_{1,1}+x_{1,2})} - \frac{z_{1,1}^m z_{1,2}^n}{m! n!} e^{-(z_{1,1}+z_{1,2})} \right| \\
 &\leq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} 2 \left| \frac{x_{1,1}^m}{m!} e^{-x_{1,1}} - \frac{z_{1,1}^m}{m!} e^{-z_{1,1}} \right| \left| \frac{x_{1,2}^n}{n!} e^{-x_{1,2}} \right| \\
 &\quad + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} 2 \left| \frac{x_{1,2}^n}{n!} e^{-x_{1,2}} - \frac{z_{1,2}^n}{n!} e^{-z_{1,2}} \right| \left| \frac{z_{1,1}^m}{m!} e^{-z_{1,1}} \right| \\
 &\leq 2 \sum_{i=0}^{\infty} |p(i|x_{1,1}) - p(i|z_{1,1})| + 2 \sum_{i=0}^{\infty} |p(i|x_{1,2}) - p(i|z_{1,2})|.
 \end{aligned}$$

D'après la preuve du Lemme 6.4 de **WANG et al., 2014**, nous savons que : $\sum_{i=0}^{\infty} |p(i|x_1) - p(i|z_1)| \leq 2(1 - e^{-|x_1 - z_1|})$, où $p(i|x)$ est la fonction de masse de probabilité d'une loi de Poisson univariée d'intensité x évaluée en i .

De plus, comme $|x_{1,i} - z_{1,i}| \leq \|x_1 - z_1\|_1 \leq c_p \|x_1 - z_1\|_p$, pour $i = 1, 2$ et pour $1 \leq p \leq \infty$ avec $c_p = 2^{1-1/p} \leq 2$, on a :

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |p(m, n|x_1) - p(m, n|z_1)| \leq 8(1 - e^{-2\|x_1 - z_1\|_p}). \quad (\text{A.1})$$

Il s'ensuit de $|f| \leq 1$ que $I_2 \leq 8(1 - e^{-2\|x_1 - z_1\|_p})$. Quant à I_1 , puisque $\|\omega + Ax_1 + B(m, n) - (\omega + Az_1 + B(m, n))\|_p = \|A(x_1 - z_1)\|_p \leq \|A\|_p \|x_1 - z_1\|_p \leq \eta$, alors $I_1 \leq \varepsilon'$.

Par conséquent :

$$|P_{x_1}f - P_{z_1}f| \leq \varepsilon' + 8(1 - e^{-2\|x_1 - z_1\|_p}). \quad (\text{A.2})$$

Pour le cas $k = 2$, il découle de $E\{f(\lambda_2)|\lambda_0 = x\} = E\{E[f(\lambda_2)|\lambda_1]|\lambda_0 = x\}$ que :

$$\begin{aligned}
 |P_{x_1}^2 f - P_{z_1}^2 f| &= \left| \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} [p(m, n|x_1)P_{x_2}f - p(m, n|z_1)P_{z_2}f] \right| \\
 &\leq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p(m, n|x_1) |P_{x_2}f - P_{z_2}f| + \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |p(m, n|x_1) - p(m, n|z_1)| |P_{z_2}f|,
 \end{aligned}$$

où $x_2 = \omega + Ax_1 + B(m, n)$ et $z_2 = \omega + Az_1 + B(m, n)$. Puisque $\|x_2 - z_2\|_p = \|A(x_1 - z_1)\|_p \leq \|A\|_p\|x_1 - z_1\|_p \leq \eta$, il découle de (A.1) et (A.2) que :

$$\begin{aligned} |P_{x_1}^2 f - P_{z_1}^2 f| &\leq \epsilon' + 8(1 - e^{-2\|x_2 - z_2\|_p}) + 8(1 - e^{-2\|x_1 - z_1\|_p}) \\ &\leq \epsilon' + 8(1 - e^{-2\|A\|_p\|x_1 - z_1\|_p}) + 8(1 - e^{-2\|x_1 - z_1\|_p}). \end{aligned}$$

Par conséquent, par récurrence, on a pour tout $k \geq 1$:

$$\begin{aligned} |P_{x_1}^k f - P_{z_1}^k f| &\leq \epsilon' + 8 \sum_{s=0}^{k-1} (1 - e^{-2\|A\|_p^s\|x_1 - z_1\|_p}) \\ &\leq \epsilon' + 16 \sum_{s=0}^{\infty} \|A\|_p^s \|x_1 - z_1\|_p \\ &\leq \epsilon' + \frac{16\eta}{1 - \|A\|_p} \\ &\leq \epsilon, \end{aligned}$$

ce qui prouve que $\{\lambda_t\}$ est une e-chaîne. Par conséquent, il existe une unique distribution stationnaire pour $\{\lambda_t\}$.

La 2ème partie du théorème se démontre de manière similaire à la preuve de la Proposition 4.2.1 de [H. LIU, 2012](#).

A.2 Méthodologie de ZHU, 2011 (II.3.1.1)

Développons en un peu plus de détail la méthodologie entreprise par [ZHU \(2011\)](#) :

Le modèle d'origine utilisé est un INGARCH de distribution binomiale négative **NBINGARCH**, la forme **NBINGARCH(1,1)** défini comme suit est celle utilisée :

$$\begin{aligned} X_t | \mathcal{F}_{t-1} &\sim \mathcal{NB}(r, p_t) \\ \frac{1 - p_t}{p_t} &= \lambda_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}. \end{aligned}$$

En utilisant le Lemme II.1 une écriture **ARMA(1,1)** que le modèle d'origine satisfait est tirée :

$$(X_t - \mu) - a(X_{t-1} - \mu) = e_t + b e_{t-1}, \quad (\text{A.3})$$

avec $a = r\alpha_1 + \beta_1$, $b = -\beta_1$ et e_t un bruit blanc.

On sait que ([ZHU, 2011 remarque 2, page 3](#)) :

$$\mu = \frac{r\alpha_0}{1 - r\alpha_1 - \beta_1},$$

on tire alors une nouvelle écriture équivalente :

$$\begin{aligned} X_t - \mu - aX_{t-1} + a\mu &= e_t + be_{t-1}, \\ X_t &= \mu + aX_{t-1} - a\mu + e_t + be_{t-1}, \\ X_t &= \mu(1-a) + aX_{t-1} + e_t + be_{t-1}, \\ X_t &= \frac{r\alpha_0}{(1-a)}(1-a) + aX_{t-1} + e_t + be_{t-1}, \\ X_t &= r\alpha_0 + aX_{t-1} + e_t + be_{t-1}. \end{aligned}$$

Passons maintenant vers la procédure d'estimation par **MCC** (CLS) :

- La première étape consiste à ajuster sur un autorégressif d'ordre supérieur, estimer ses paramètres et enfin obtenir une estimation des erreurs \hat{e}_t : Posons $Y_t = X_t - \sum_{i=1}^n \frac{X_i}{n}$ et ajustons par un modèle **AR**(p^*) :

$$\begin{aligned} Y_t &= \sum_{i=1}^{p^*} a_i Y_{t-i}, \\ \hat{e}_t &:= Y_t - \sum_{i=1}^{p^*} \hat{a}_i Y_{t-i}. \end{aligned}$$

- La deuxième étape consiste à les insérer dans l'équation que nous avons obtenu précédemment et de minimiser la fonction objective :

$$\sum_{t=1}^n (X_t - r\alpha_0 - aX_{t-1} - b\hat{e}_{t-1})^2,$$

Après minimisation, nous aurons une estimation des paramètres $\hat{\alpha}_0$, \hat{a} , \hat{b} , qui permet d'obtenir une estimation des paramètres du modèle d'origine **NBINGARCH**(1,1) par les relations :

$$\begin{aligned} \hat{\alpha}_1 &= \frac{\hat{a} + \hat{b}}{r} \\ \hat{\beta}_1 &= -\hat{b} \end{aligned}$$

A.3 Estimation d'un VAR par MCO/MCG :

La formulation générale du **VAR(p)** est donnée par comme :

$$\Phi(L)X_t = \varepsilon_t,$$

où $\Phi(L) = I - \sum_{i=1}^p \Phi_i L^i$.

(X_t le vecteur d'observations N -dimensionnel à l'instant t , et ε_t un bruit (de même dimension N), et chaque Φ_i , ($i \in \{1, \dots, p\}$) une matrice $N \times N$.)

La j-ème équation du **VAR(p)** se décompose en :

$$\underline{X}_j = \begin{pmatrix} X_{j1} \\ X_{j2} \\ \vdots \\ X_{jT} \end{pmatrix} = \begin{pmatrix} X'_1 \\ X'_2 \\ \vdots \\ X'_T \end{pmatrix} \Psi_j + \varepsilon_j,$$

avec \underline{X}_j un vecteur $(T \times 1)$ contenant les T observations de la composante j , et $X'_t = (X_{1t-1}, X_{2t-1}, \dots, X_{Nt-1}, X_{1t-2}, \dots, X_{Nt-2}, \dots, X_{1t-p}, \dots, X_{Nt-p})$ de dimension $(1 \times Np)$.

ou encore :

$$\underline{X}_j = \underline{\underline{X}} \Psi_j + \varepsilon_j.$$

avec :

$$\underline{\underline{X}} = \begin{pmatrix} X'_1 \\ X'_2 \\ \vdots \\ X'_T \end{pmatrix}, \varepsilon_j = \begin{pmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \vdots \\ \varepsilon_{jT} \end{pmatrix},$$

$$\Psi_j = \begin{pmatrix} \phi_{1j}^1 \\ \phi_{1j}^2 \\ \vdots \\ \phi_{1j}^N \\ \phi_{2j}^1 \\ \vdots \\ \phi_{pj}^N \end{pmatrix}.$$

Ψ_j est un vecteur vertical de taille $(Np \times 1)$ où ϕ_{ij}^k représente l'élément de la matrice de coefficients du retard i Φ_i , à la ligne j et colonne k .

et $\underline{\underline{X}}$ une matrice de taille $(T \times Np)$

Après l'empilement des N équations, on a le **VAR** suivant :

$$\begin{pmatrix} \underline{X}_1 \\ \underline{X}_2 \\ \vdots \\ \underline{X}_N \end{pmatrix} = \begin{pmatrix} X_{11} \\ X_{12} \\ \vdots \\ X_{1T} \\ X_{21} \\ \vdots \\ X_{NT} \end{pmatrix} = \begin{pmatrix} \underline{\underline{X}} & 0 & \dots & 0 \\ 0 & \underline{\underline{X}} & \dots & 0 \\ & & \ddots & \\ & & & \underline{\underline{X}} \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \vdots \\ \Psi_N \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1T} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{NT} \end{pmatrix} \quad (\text{A.4})$$

On observe que le vecteur de variables à expliquer est de taille $(NT \times 1)$, la matrices des variables explicatives est diagonale par blocs de taille $(NT \times N^2p)$, et le vecteur de coefficients Ψ est de taille (N^2p) .

La matrice Var-Covar des erreurs indique la présence d'hétéroscédasticité et d'autocorrélation ce qui pose un problème lors de l'application des MCO (les estimateurs sont sans biais mais ne sont plus de variance minimale). On passe alors à l'utilisation de la méthode MCG qui fournit un estimateur **BLUE**.

La matrice var-cov des erreurs (de taille $(NT \times NT)$) :

$$\left(\begin{array}{cccc} (\sigma_{11}) & 0 & \dots & 0 \\ 0 & (\sigma_{11}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_{11}) \\ (\sigma_{21}) & 0 & \dots & 0 \\ 0 & (\sigma_{21}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_{21}) \end{array} \right) \left(\begin{array}{cccc} (\sigma_{12}) & 0 & \dots & 0 \\ 0 & (\sigma_{12}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_{12}) \\ (\sigma_{22}) & 0 & \dots & 0 \\ 0 & (\sigma_{22}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_{22}) \end{array} \right) \dots \dots \left(\begin{array}{cccc} (\sigma_{NN}) & 0 & \dots & 0 \\ 0 & (\sigma_{NN}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_{NN}) \end{array} \right)$$

Une autre écriture pour la matrice var-cov est donnée par :

$$\mathbf{V}(\epsilon) = \Sigma_\epsilon \otimes I = \Omega$$

où $\Sigma_\epsilon = \alpha_{ij}$.

Cependant, la structure diagonale par blocs de la matrice des variables explicatives permet la mise en œuvre de l'estimation par les Moindres Carrés Ordinaires (MCO) bloc par bloc. Il est important de noter que le théorème de Zellner démontre l'équivalence entre l'estimation de chacune des N équations individuellement par MCO et l'estimation du modèle complet par MCG. (prouvé en annexe [Section A.4](#))

On a alors :

$$\hat{\Psi}_{MCO} = \begin{pmatrix} (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T Y_1 \\ (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T Y_2 \\ \vdots \\ (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T Y_N \end{pmatrix}$$

A.4 Preuve de l'équivalence entre MCG et MCO bloc par bloc (A.3), (Proposition II.1)

Comme nous l'avons vu précédemment, la matrice de variance-covariance des résidus suggère l'application théorique des Moindres Carrés Généralisés (MCG). Cependant, la structure diagonale par blocs de la matrice des variables explicatives permet la mise en œuvre de l'estimation par les Moindres Carrés Ordinaires (MCO) bloc par bloc.

Le théorème de Zellner démontre l'équivalence entre l'estimation de chacune des N équations individuellement par MCO et l'estimation du modèle complet par MCG **LARDIC et MIGNON, 2002.**

Pour un modèle défini par : $Y = Xa + b$, où ϵ est un bruit blanc.

L'estimateur des MCO est donné par : $\hat{a}_{MCO} = (X^T X)^{-1} X^T Y$.

L'estimateur des MCG est donné par : $\hat{a}_{MCG} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} Y$. Où Ω désigne la matrice var-cov de ϵ .

Dans notre cas, on a :

$$X = \begin{pmatrix} \underline{\underline{X}} & 0 & \dots & 0 \\ 0 & \underline{\underline{X}} & \dots & 0 \\ & & \ddots & \\ & & & \underline{\underline{X}} \end{pmatrix} = I \otimes \underline{\underline{X}}$$

$$\begin{aligned} X^T \Omega^{-1} X &= (I \otimes \underline{\underline{X}}^T)(\Sigma_{\epsilon}^{-1} \otimes I)(I \otimes \underline{\underline{X}}) \\ &= \Sigma_{\epsilon} \otimes \underline{\underline{X}}^T \underline{\underline{X}} \end{aligned}$$

avec $\Omega^{-1} = \Sigma_{\epsilon}^{-1} \otimes I$.

En déduit que :

$$(X^T \Omega^{-1} X)^{-1} = \Sigma_{\epsilon} \otimes \underline{\underline{X}}^T \underline{\underline{X}}$$

$$\begin{aligned} X^T \Omega^{-1} Y &= (I \otimes \underline{\underline{X}}^T)(\Sigma_{\epsilon}^{-1} \otimes I)Y \\ &= (\Sigma_{\epsilon} \otimes \underline{\underline{X}}^T)Y \end{aligned}$$

On a donc :

$$\begin{aligned} \hat{a}_{MCG} &= \Sigma_{\epsilon} \otimes (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} (\Sigma_{\epsilon} \otimes \underline{\underline{X}}^T)Y \\ &= I \otimes (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T Y \end{aligned}$$

D'où :

$$\begin{aligned}\hat{a}_{MCG} &= \begin{pmatrix} (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T & 0 & \dots & 0 \\ 0 & (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T & \dots & 0 \\ & & \ddots & \\ & & & (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix} \\ &= \begin{pmatrix} (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T Y_1 \\ (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T Y_2 \\ \vdots \\ (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T Y_N \end{pmatrix}\end{aligned}$$

On aboutit à l'estimateur des MCO équation par équation (bloc par bloc).

A.5 Algorithme pour générer des échantillons de la distribution de Poisson bivariée

En suivant les étapes de l'algorithme proposé par [Cui et Zhu \(2018\)](#), on peut générer un échantillon de la distribution bivarié $\text{BP}(\lambda_1, \lambda_2, \delta)$ ([I.1.1.1.2](#))

Puisque nous connaissons la fonction de masse conjointe et marginale de $\text{BP}(\lambda_1, \lambda_2, \delta)$, l'étape clé est de savoir comment générer des échantillons i.i.d. de (Y_1, Y_2) . Notons que $Y_1 \sim P(\lambda_1)$ et que la fonction de masse conditionnelle de $(Y_2|Y_1)$ est :

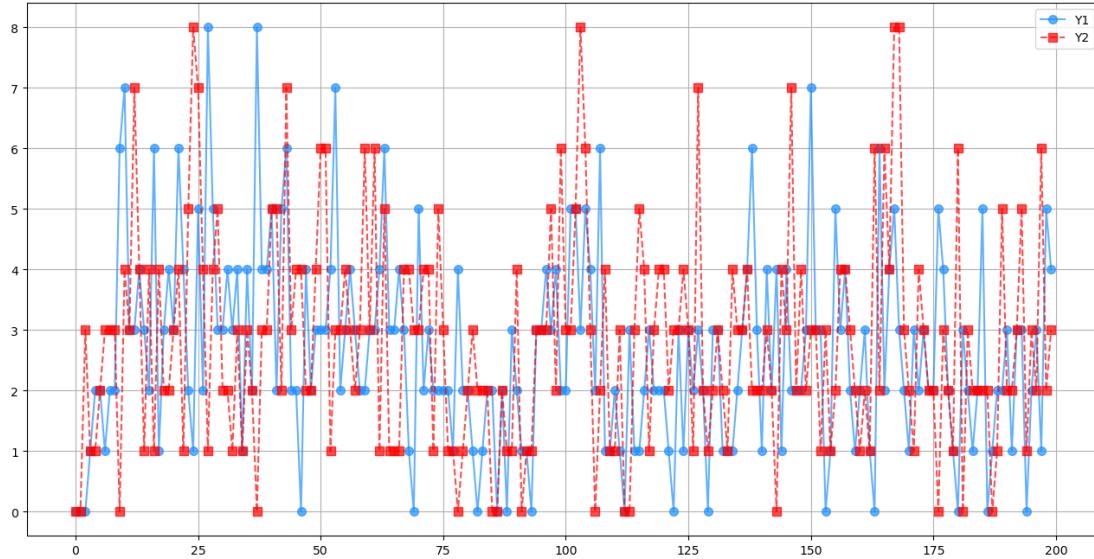
$$P(y_2|y_1) := P(Y_2 = y_2|Y_1 = y_1) = \frac{\lambda_2^{y_2}}{y_2!} e^{-\lambda_2} [1 + \delta(e^{-y_1} - e^{-c\lambda_1})(e^{-y_2} - e^{-c\lambda_2})], \quad y_2 = 0, 1, 2, \dots$$

Ainsi, la méthode d'échantillonnage conditionnel et de transformation inverse est utilisée avec la fonction de masse cumulative suivante :

$$F_{y_2|y_1} = \sum_{i=1}^{\infty} P(Y_2 = i|Y_1 = y_1) = \sum_{i=1}^{\infty} \frac{\lambda_2^i}{i!} e^{-\lambda_2} [1 + \delta(e^{-y_1} - e^{c\lambda_1})(e^{-i} - e^{c\lambda_2})],$$

- **étape 1** : On commence par des valeurs pour λ_1, λ_2 et δ .
- **étape 2** : On génère l'observation $y_1 \sim \mathcal{P}(\lambda_1)$.
- **étape 3** : On génère un nombre aléatoire $u \sim \mathbf{U}(0, 1)$.
- **étape 4** : Si $u < F_{y_2|y_1}(0)$ alors $y_2 = 0$.
- **étape 5** : S'il existe un k , tel que $F_{y_2|y_1}(k-1) \leq u < F_{y_2|y_1}(k)$, alors $y_2 = k$.

La figure ([A.1](#)) illustre les réalisations des deux variables aléatoires Y_1 et Y_2 composantes de la variable aléatoire bivariée $Y \sim BP(5, 3, 1)$ avec distributions marginales $Y_i \sim \mathcal{P}(\lambda_i)$.

FIGURE A.1 – Un échantillon de la distribution bivarié BP .

A.6 Ajustement du VAR(p^*)

Le tableau Table A.1 résume les résultats d'estimation des données centrées (le VAR(p^*)) par utilisation du package `statsmodels` et par programmation directe de la méthode MCO/MCG :

"statsmodel" package		
	Equation Y1	Equation Y2
$\beta_{1,1}$	0.076756	0.343100
$\beta_{1,2}$	0.233846	0.164855
$\beta_{2,1}$	0.113633	0.126374
$\beta_{2,2}$	0.205832	0.094280
$\beta_{3,1}$	-0.064629	0.079058
$\beta_{3,2}$	0.090207	0.043254

programmation MCO/MCG		
	Equation Y1	Equation Y2
$\beta_{1,1}$	0.07675628	0.34309996
$\beta_{1,2}$	0.23384587	0.16485549
$\beta_{2,1}$	0.11363265	0.12637447
$\beta_{2,2}$	0.20583171	0.0942801
$\beta_{3,1}$	-0.06462857	0.07905843
$\beta_{3,2}$	0.09020714	0.0432539

TABLE A.1 – Résultats pour les équations Y1 et Y2

A.7 Adaptation de la méthode MCO pour le cas diagonal

On a pour un processus $\{Y_t\}$ avec $t = \{1, T$ satisfaisant l'équation équation II.2 pour le cas diagonal où les deux matrices A et B sont diagonales, la formulation mathématique suivante :

$$\begin{aligned} Y_t | \mathcal{F}_{t-1} &\sim \mathbf{BP}(\lambda_{t,1}, \lambda_{t,2}, \delta), \\ \lambda_t &= (\lambda_{t,1}, \lambda_{t,2})^T = \omega + \mathbf{A}\lambda_{t-1} + \mathbf{B}Y_{t-1}, \end{aligned}$$

où $A = \begin{pmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{pmatrix}$ et $B = \begin{pmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{pmatrix}$.

$$\begin{aligned}\lambda_{t,1} &= \omega_1 + \alpha_{11}\lambda_{t-1,1} + \beta_{11}Y_{t-1,1}, \\ \lambda_{t,2} &= \omega_2 + \alpha_{22}\lambda_{t-1,2} + \beta_{22}Y_{t-1,2}.\end{aligned}\tag{A.5}$$

L'écriture VARMA de ce modèle est donnée par :

$$Y_t = \omega + \Phi Y_{t-1} + e_t + \Theta e_{t-1}$$

$$\begin{aligned}Y_{t,1} &= \omega_1 + \phi_{11}Y_{t-1,1} + \theta_{11}e_{t-1,1} \\ Y_{t,2} &= \omega_2 + \phi_{22}Y_{t-1,2} + \theta_{22}e_{t-1,2}\end{aligned}\tag{A.6}$$

Avec $\Phi = \begin{pmatrix} \phi_{11} & 0 \\ 0 & \phi_{22} \end{pmatrix} = \begin{pmatrix} \beta_{11} + \alpha_{11} & 0 \\ 0 & \beta_{22} + \alpha_{22} \end{pmatrix}$, $\Theta = \begin{pmatrix} \theta_{11} & 0 \\ 0 & \theta_{22} \end{pmatrix} = \begin{pmatrix} -\alpha_{11} & 0 \\ 0 & -\alpha_{22} \end{pmatrix}$.

Comme fait en [Section II.3.1.2](#) on aura deux vecteurs indépendants de $Y'_{i,t}$ et $e_{t,i}$, $i = 1, 2$:

$$\begin{aligned}Y'_{1,t} &= (1, Y_{t-1,1}, e_{t-1,1}) \\ Y'_{2,t} &= (1, Y_{t-1,2}, e_{t-1,2})\end{aligned}$$

Et $\Psi_1 = (\omega_1, \phi_{11}, \theta_{11})^T$, $\Psi_2 = (\omega_2, \phi_{22}, \theta_{22})^T$.

On aura deux équations VAR(p) :

$$\begin{aligned}Y_{t,1} &= Y'_{1,t} \Psi_1 + e_{t,1} \\ Y_{t,2} &= Y'_{2,t} \Psi_2 + e_{t,2}\end{aligned}\tag{A.7}$$

$$e_j = (e_{t,1}, e_{t,2})^T.$$

On pose

$$\underline{Y}_1 := \begin{pmatrix} Y_{1,1} \\ Y_{2,1} \\ \vdots \\ Y_{T,1} \end{pmatrix}, \text{ alors } \underline{\underline{Y}}_1 := \begin{pmatrix} Y'_{1,1} \\ Y'_{1,2} \\ \vdots \\ Y'_{1,T} \end{pmatrix} = \begin{pmatrix} 1 & Y_{0,1} & e_{0,1} \\ 1 & Y_{1,1} & e_{1,1} \\ \vdots & \vdots & \vdots \\ 1 & Y_{T-1,1} & e_{T-1,1} \end{pmatrix}$$

$$\underline{Y}_2 := \begin{pmatrix} Y_{1,2} \\ Y_{2,2} \\ \vdots \\ Y_{T,2} \end{pmatrix}, \text{ alors } \underline{\underline{Y}}_2 := \begin{pmatrix} Y'_{2,1} \\ Y'_{2,2} \\ \vdots \\ Y'_{2,T} \end{pmatrix} = \begin{pmatrix} 1 & Y_{0,2} & e_{0,2} \\ 1 & Y_{1,2} & e_{1,2} \\ \vdots & \vdots & \vdots \\ 1 & Y_{T-1,2} & e_{T-1,2} \end{pmatrix}$$

On aura :

$$\begin{aligned}\underline{Y}_1 &= \underline{\underline{Y}}_1 \Psi_1 + e_1 \\ \underline{Y}_2 &= \underline{\underline{Y}}_2 \Psi_2 + e_2\end{aligned}\tag{A.8}$$

avec \underline{Y}_i de dimension $(T \times 1)$ et $\underline{\underline{Y}}_i$ de dimension $(T \times 3)$ pour $i = 1, 2$.

La matrice Var-Cov $(2T \times 2T)$ des erreurs indique l'absence d'auto-corrélation, et d'hétéroscédasticité dans les séries des composantes à part.

On peut alors appliquer une estimation MCO équation par équation et on aura des estimateurs sans biais et de variance minimale.

La matrice var-cov des erreurs :

$$\begin{pmatrix} \begin{pmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{11} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{11} \end{pmatrix} & \mathbf{0}_{(T \times T)} \\ \mathbf{0}_{(T \times T)} & \begin{pmatrix} \sigma_{22} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{22} \end{pmatrix} \end{pmatrix}$$

Sans passage par MCG. L'estimateur des MCO de Ψ_i , pour $i = 1, 2$ est donnée par :

$$\Psi_i = (\underline{\underline{Y}}_i^T \underline{\underline{Y}}_i)^{-1} \underline{\underline{Y}}_i^T \underline{Y}_i$$

A.8 Résultats d'estimation MV en Python

En utilisant la librairie SciPy sur Python, et la fonction `optimize.minimize`, les résultats suivants pour le modèle **M1** $n = 1000$ et $r = 250$ réplications est donné.

θ^0	0.5	0.3	0.2	0.1	0.2
La moyenne	0.509	0.290	0.210	0.096	0.199
Biais	-0.0090	0.0099	-0.0099	0.0043	0.0005
Écart-type	0.129	0.160	0.136	0.031	0.031

(a) Paramètres 1 à 5

θ^0	0.3	0.1	0.3	0.3	0.2
La moyenne	0.305	0.111	0.283	0.305	0.198
Biais	-0.0055	-0.0113	0.0171	-0.0053	0.0021
Écart-type	0.117	0.127	0.119	0.031	0.032

(b) Paramètres 6 à 10

θ^0	0.4
La moyenne	0.332
Biais	0.0679
Écart-type	0.482

(c) Paramètre 11

TABLE A.2 – Résultats d'estimation sur Python

L'adaptation de cette routine à notre situation demanderait un travail de développement supplémentaire, ce qui n'est pas réalisable dans le temps imparti.

BIBLIOGRAPHIE

- AKNOUCHE, A., & DEMMOUCHE, N. (2019). Ergodicity conditions for a double mixed Poisson autoregression. *Statistics & Probability Letters*, 147, 6-11.
- AL-OSH, M. A., & ALZAID, A. A. (1987). First-order integer-valued autoregressive (INAR (1)) process. *Journal of Time Series Analysis*, 8(3), 261-275.
- ALZAID, A. A., & AL-OSH, M. (1990). An integer-valued pth-order autoregressive structure (INAR (p)) process. *Journal of Applied Probability*, 27(2), 314-324.
- ARMILLOTTA, M., & FOKIANOS, K. (2023). Nonlinear network autoregression. *The Annals of Statistics*, 51(6), 2526-2552.
- BENTARZI, M., & SADOUN, M. (2021). Efficient estimation in periodic INARp model : Nonparametric innovation distributions case. *Journal of Statistical Planning and Inference*, 211, 340-361.
- BLUNDELL, R., GRIFFITH, R., & WINDMEIJER, F. (2002). Individual effects and dynamics in count data models. *Journal of econometrics*, 108(1), 113-131.
- BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307-327.
- BRIJS, T., KARLIS, D., & WETS, G. (2008). Studying the effect of weather conditions on daily crash counts using a discrete time-series model. *Accident Analysis & Prevention*, 40(3), 1180-1190.
- CAMPBELL, J. (1934). The Poisson correlation function. *Proceedings of the Edinburgh Mathematical Society*, 4(1), 18-26.
- CHESNEAU, C., KACHOUR, M., & KARLIS, D. (2015). On some distributions arising from a generalized trivariate reduction scheme. *Statistical Methodology*, 25, 36-50.
- CUI, Y., LI, Q., & ZHU, F. (2020). Flexible bivariate Poisson integer-valued GARCH model. *Annals of the Institute of Statistical Mathematics*, 72(6), 1449-1477.
- CUI, Y., & ZHU, F. (2018). A new bivariate integer-valued GARCH model allowing for negative cross-correlation. *Test*, 27, 428-452.
- DEBALY, Z. M., & TRUQUET, L. (2019). Stationarity and moment properties of some multivariate count autoregressions. *arXiv preprint arXiv:1909.11392*.
- DIOP, M. L., DIOP, A., & DIONGUE, A. K. (2016). A mixture integer-valued GARCH model. *Revstat-Statistical Journal*, 14(3), 245-271.
- DIOP, M. L., & KENGNE, W. (2017). Testing Parameter Change in General Integer-Valued Time Series. *Journal of Time Series Analysis*, 38(6), 880-894.
- DOUKHAN, P., FOKIANOS, K., & RYNKIEWICZ, J. (2021). Mixtures of nonlinear Poisson autoregressions. *Journal of Time Series Analysis*, 42(1), 107-135.
- FERLAND, R., LATOUR, A., & ORAICHI, D. (2006). Integer-valued GARCH process. *Journal of time series analysis*, 27(6), 923-942.
- FOKIANOS, K. (2021). Multivariate count time series modelling. *Econometrics and Statistics*.
- FOKIANOS, K., FRIED, R., KHARIN, Y., & VOLOSHKO, V. (2022). Statistical analysis of multivariate discrete-valued time series. *Journal of multivariate analysis*, 188, 104805.

- FOKIANOS, K., STØVE, B., TJØSTHEIM, D., & DOUKHAN, P. (2020). Multivariate count autoregression.
- FOKIANOS, K., & TJØSTHEIM, D. (2011). Log-linear Poisson autoregression. *Journal of multivariate analysis*, 102(3), 563-578.
- FRANCQ, C., & ZAKOIAN, J.-M. (2022). Testing the existence of moments for GARCH processes. *Journal of econometrics*, 227(1), 47-64.
- HOLGATE, P. (1964). Estimation for the bivariate Poisson distribution. *Biometrika*, 51(1-2), 241-287.
- JOHNSON, N. L., KOTZ, S., & BALAKRISHNAN, N. (1997). *Discrete multivariate distributions* (T. 165). Wiley New York.
- KANG, J., & LEE, S. (2014). Parameter change test for Poisson autoregressive models. *Scandinavian Journal of Statistics*, 41(4), 1136-1152.
- KARLIS, D. (2016). Models for multivariate count time series. *Handbook of discrete-valued time series*, 407, 424.
- KARLIS, D., SERMAIDIS, G., & BRIJS, T. (2008). Discrete valued time series models for examining weather effects in daily accident counts. *Stat. Modell.*
- KAWAMURA, K. (1976). The structure of trivariate Poisson distribution. *Kodai Mathematical Seminar Reports*, 28(1), 1-8.
- KOCHERLAKOTA, S., & KOCHERLAKOTA, K. (2017). *Bivariate discrete distributions*. CRC Press.
- LAKSHMINARAYANA, J., PANDIT, S. N., & SRINIVASA RAO, K. (1999). On a bivariate Poisson distribution. *Communications in Statistics-Theory and Methods*, 28(2), 267-276.
- LARDIC, S., & MIGNON, V. (2002). *Econom'etrie des s'eries temporelles macro'economiques et financi'eres*. Economica.
- LATOUR, A. (1997). The multivariate GINAR (p) process. *Advances in Applied Probability*, 29(1), 228-248.
- LEE, S., KIM, D., & KIM, B. (2023). Modeling and inference for multivariate time series of counts based on the INGARCH scheme. *Computational Statistics & Data Analysis*, 177, 107579.
- LEE, Y., & LEE, S. (2019). CUSUM test for general nonlinear integer-valued GARCH models : Comparison study. *Annals of the Institute of Statistical Mathematics*, 71, 1033-1057.
- LEE, Y., LEE, S., & TJØSTHEIM, D. (2018). Asymptotic normality and parameter change test for bivariate Poisson INGARCH models. *Test*, 27, 52-69.
- LIU, H. (2012). *Some models for time series of counts*. Columbia University.
- LIU, M., ZHU, F., LI, J., & SUN, C. (2023). A systematic review of INGARCH models for integer-valued time series. *Entropy*, 25(6), 922.
- LIU, Y., & LUGER, R. (2009). Efficient estimation of copula-GARCH models. *Computational Statistics & Data Analysis*, 53(6), 2284-2297.
- LIVSEY, J., LUND, R., KECHAGIAS, S., & PIPIRAS, V. (2018). Multivariate integer-valued time series with flexible autocovariances and their application to major hurricane counts. *The Annals of Applied Statistics*, 12(1), 408-431.
- LOUKAS, S., & PAPAGEORGIOU, H. (1991). On a trivariate Poisson distribution. *Applications of Mathematics*, 36(6), 432-439.
- MAGESA, G., SINKWEMBE, E. E., SHABAN, N., & NGAILO, T. (2023). Bivariate Discrete Time Series Model for Assessing the effects of Rainfall and Temperature on Road Accidents : The case of Morogoro and Pwani Regions in Tanzania. *Scientific African*, 19, e01522.
- MAO, H., ZHU, F., & CUI, Y. (2020). A generalized mixture integer-valued GARCH model. *Statistical Methods & Applications*, 29, 527-552.
- MEYN, S. P., & TWEEDIE, R. L. (2012). *Markov chains and stochastic stability*. Springer Science & Business Media.

- OUZZANI, F., & BENTARZI, M. (2021). On mixture periodic Integer-Valued ARCH models. *Communications in Statistics-Simulation and Computation*, 50(12), 3931-3957.
- PAUL, M., HELD, L., & TOSCHKE, A. M. (2008). Multivariate modelling of infectious disease surveillance data. *Statistics in medicine*, 27(29), 6250-6267.
- PEDELI, X., & KARLIS, D. (2011). A bivariate INAR (1) process with application. *Statistical modelling*, 11(4), 325-349.
- RAVISHANKER, N., VENKATESAN, R., & HU, S. (2016). Dynamic models for time series of counts with a marketing application. *Handbook of discrete-valued time series*, 425-446.
- SARABIA ALEGRIA, J. M., GOMEZ D'ENIZ, E., et al. (2008). Construction of multivariate distributions : a review of some recent results.
- SCOTTO, M. G., WEISS, C. H., & GOUVEIA, S. (2015). Thinning-based models in the analysis of integer-valued time series : a review. *Statistical Modelling*, 15(6), 590-618.
- SONG, P. X.-K., FAN, Y., & KALBFLEISCH, J. D. (2005). Maximization by parts in likelihood inference. *Journal of the American Statistical Association*, 100(472), 1145-1158.
- TEICHER, H. (1954). On the multivariate Poisson distribution. *Scandinavian Actuarial Journal*, 1954(1), 1-9.
- WANG, C., LIU, H., YAO, J.-F., DAVIS, R. A., & LI, W. K. (2014). Self-excited threshold Poisson autoregression. *Journal of the American Statistical Association*, 109(506), 777-787.
- WEISS, C. H. (2018). *An introduction to discrete-valued time series*. John Wiley & Sons.
- WU, W. B., & SHAO, X. (2004). Limit theorems for iterated random functions. *Journal of Applied Probability*, 41(2), 425-436.
- ZHU, F. (2011). A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis*, 32(1), 54-67.
- ZHU, F., LI, Q., & WANG, D. (2010). A mixture integer-valued ARCH model. *Journal of Statistical Planning and inference*, 140(7), 2025-2036.